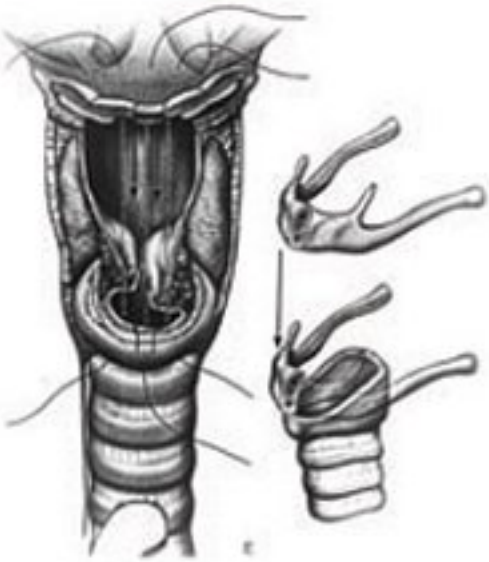
Ballenger's Otorhinolaryngology Head and Neck Surgery

Sixteenth Edition



James B. Snow Jr, MD
John Jacob Ballenger, MD

Ballenger's

Otorhinolaryngology Head and Neck Surgery

Sixteenth Edition

James B. Snow Jr, MD
Professor Emeritus
University of Pennsylvania
Philadelphia, Pennsylvania

Former Director National Institute on Deafness and
Other Communication Disorders
National Institutes of Health
Bethesda, Maryland

John Jacob Ballenger, MD
Associate Professor
Department of Otolaryngology–Head and Neck Surgery
Northwestern University
Chicago, Illinois

Chief Emeritus
Division of Otolaryngology–Head and Neck Surgery
Evanston Hospital
Evanston, Illinois

BC Decker Inc
P.O. Box 620, L.C.D. 1
Hamilton, Ontario L8N 3K7
Tel: 905-522-7017; 800-568-7281
Fax: 905-522-7839; 888-311-4987
E-mail: info@bcdecker.com
www.bcdecker.com

Previously published by Williams & Wilkins 1996 under ISBN 0-683-00315-1.

02 03 04 05/GSA/9 8 7 6 5 4 3 2 1

ISBN 1-55009-197-2
Printed in Spain

Sales and Distribution

United States
BC Decker Inc
P.O. Box 785
Lewiston, NY 14092-0785
Tel: 905-522-7017; 800-568-7281
Fax: 905-522-7839; 888-311-4987
E-mail: info@bcdecker.com
www.bcdecker.com

Canada
BC Decker Inc
20 Hughson Street South
P.O. Box 620, LCD 1
Hamilton, Ontario L8N 3K7
Tel: 905-522-7017; 800-568-7281
Fax: 905-522-7839; 888-311-4987
E-mail: info@bcdecker.com
www.bcdecker.com

Foreign Rights
John Scott & Company
International Publishers' Agency
P.O. Box 878
Kimberton, PA 19442
Tel: 610-827-1640
Fax: 610-827-1671
E-mail: jsco@voicenet.com

Japan
Igaku-Shoin Ltd.
Foreign Publications Department
3-24-17 Hongo
Bunkyo-ku, Tokyo, Japan 113-8719
Tel: 3 3817 5680
Fax: 3 3815 6776
E-mail: fd@igaku-shoin.co.jp

U.K., Europe, Scandinavia, Middle East
Elsevier Science
Customer Service Department
Foots Cray High Street
Sidcup, Kent
DA14 5HP, UK
Tel: 44 (0) 208 308 5760
Fax: 44 (0) 181 308 5702
E-mail: cservice@harcourt.com

Singapore, Malaysia,Thailand, Philippines, Indonesia, Vietnam, Pacific Rim, Korea
Elsevier Science Asia
583 Orchard Road
#09/01, Forum
Singapore 238884
Tel: 65-737-3593
Fax: 65-753-2145

Australia, New Zealand
Elsevier Science Australia
Customer Service Department
STM Division
Locked Bag 16
St. Peters, New South Wales, 2044
Australia
Tel: 61 02 9517-8999
Fax: 61 02 9517-2249
E-mail: stmp@harcourt.com.au
www.harcourt.com.au

Mexico and Central America
ETM SA de CV
Calle de Tula 59
Colonia Condesa
06140 Mexico DF, Mexico
Tel: 52-5-5553-6657
Fax: 52-5-5211-8468
E-mail: editoresdetextosmex@prodigy.net.mx

Argentina
CLM (Cuspide Libros Medicos)
Av. Córdoba 2067 - (1120)
Buenos Aires, Argentina
Tel: (5411) 4961-0042/(5411) 4964-0848
Fax: (5411) 4963-7988
E-mail: clm@cuspide.com

Brazil
Tecmedd
Av. Maurílio Biagi, 2850
City Ribeirão Preto – SP – CEP: 14021-000
Tel: 0800 992236
Fax: (16) 3993-9000
E-mail: tecmedd@tecmedd.com.br

Contents

Foreword

This edition of *Ballenger's Otorhinolaryngology Head and Neck Surgery* continues in the tradition of being a comprehensive compendium of current knowledge in the fields of otology/neurotology, rhinology, facial plastic and reconstructive surgery, pediatric otorhinolaryngology, laryngology, head and neck surgery, and bronchoesophagology. The informational content for the 16th edition has been reorganized into 67 chapters with new sections on facial plastic and reconstructive surgery and pediatric otorhinolaryngology. More than half of the authors are new to the book, and another third of the chapters have been redesigned to address major developments in the topical areas. The content reflects the central responsibility of the otorhinolaryngologist in treating patients with diseases affecting the senses of smell, taste, and balance and with disorders of human communication affecting hearing, voice, speech, and language. The presentation of basic science underlying clinical otorhinolaryngology head and neck surgery has been expanded, and chapters on clinical topics have been structured to provide comprehensive coverage to meet the needs of practitioners in the twenty-first century.

The 64 senior authors have been chosen not only for their contribution to new knowledge to their topics through highly regarded research but also for their intellectual leadership of the specialty and, thereby, are truly authoritative in the subject matter of their chapters.

Both the basic science and clinical chapters stress the important role of molecular biology in understanding the pathogenesis of disease. Throughout the book, the genetic basis of disease is emphasized. This book provides students, residents, fellows, and practitioners with the understanding necessary to participate fully in the coming role of molecular biology in therapy.

James B. Snow, Jr, MD

Preface

Ballenger's Otorhinolaryngology Head and Neck Surgery is in its 94th year in this sixteenth edition. It has been in continuous use since 1908, when my great-uncle, William Lincoln Ballenger, brought out the first edition. My father, Howard Charles Ballenger, succeeded my great-uncle as editor, and in 1957, I succeeded my father. I am greatly pleased to have James B. Snow Jr. as coeditor. This reorganized and revitalized version is much to the credit of his leadership and forward-looking enthusiasm. We are delighted to have BC Decker Inc as the book's new publisher.

This new edition, as in the past, is enriched by the outstanding contributions of 118 distinguished scholars and practitioners, all of whom have written on their special interests. I extend my gratitude to each of them as well as their medical artists and illustrators, whose work clarifies the text. The contributors are drawn mainly from the United States but from several other countries as well.

The general format is to present an accurate picture of normal anatomy and physiology as a basis for the understanding of altered conditions that underlie diseases and their symptoms and signs. I believe that this book will be of great value to all studying this specialty and to other specialists seeking an authoritative reference.

John Jacob Ballenger, MD

Contributors

Omid Abbassi, MD, PhD
Department of Otorhinolaryngology and Communicative Sciences
Baylor College of Medicine
Houston, Texas

Peter W. Alberti, MB, PhD
Department of Otolaryngology
University of Toronto
Toronto, Ontario

Daniel G. Becker, MD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

John Jacob Ballenger, MD
Department of Otolarynology–Head and Neck Surgery
Northwestern University
Chicago, Illinois

Peter C. Belafsky, MD, PhD
Department of Otolaryngology
Bowman Gray School of Medicine
Winston-Salem, North Carolina

Joel M. Bernstein, MD, PhD
Department of Otolaryngology
State University of New York
Buffalo, New York

John L. Boone, MD
Department of Otolaryngology
Camp Pendleton Naval Hospital
Camp Pendleton, California

Bruce Brockstein, MD
Department of Medicine
Northwestern University
Chicago, Illinois

Steven M. Bromley, MD
Department of Neurology
University of Pennsylvania
Philadelphia, Pennsylvania

Brian B. Burkey, MD
Department of Otolaryngology
Vanderbilt University
Nashville, Tennessee

William R. Carroll, MD
Department of Otolaryngology–Head and Neck Surgery
University of Alabama
Birmingham, Alabama

Ara A. Chalian, MD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

Richard A. Chole, MD, PhD
Department of Otolaryngology
Washington University
St. Louis, Missouri

Daniel I. Choo, MD
Department of Otolaryngology
University of Cincinnati
Cincinnati, Ohio

Ellen S. Deutsch, MD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

Laurence J. DiNardo, MD
Department of Otolaryngology–Head and Neck Surgery
Virginia Commonwealth University
Richmond, Virginia

Robert A. Dobie, MD
Division of Extramural Research
National Institute on Deafness and Other Communication Disorders
Bethesda, Maryland

Paul J. Donald, MD
Department of Otolaryngology
University of California
Sacramento, California

Richard L. Doty, PhD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

Jose N. Fayad, MD
Department of Otolaryngology
Columbia University
New York, New York

Richard R. Gacek, MD
Department of Otolaryngology–Head and Neck Surgery
University of South Alabama
Mobile, Alabama

Mark R. Gacek, MD
Department of Neurology
University of South Alabama
Mobile, Alabama

Ray O. Gustafson, MD
Department of Otolaryngology
Mayo Medical School
Rochester, Minnesota

James W. Hall III, PhD
Department of Communicative Sciences and Disorders
University of Florida
Gainesville, Florida

Lee A. Harker, MD
Department of Surgery
Creighton University
Omaha, Nebraska

Jeffrey P. Harris, MD, PhD
Division of Otolaryngology–Head and Neck Surgery
University of California
San Diego, California

James M. Hartman, MD
Department of Otolaryngology
Washington University
St. Louis, Missouri

Gerald B. Healy, MD
Department of Otology and Laryngology
Harvard Medical School
Boston, Massachusetts

Robert A. Hendrix, MD
Department of Surgery
Nash General Hospital
Rocky Mount, North Carolina

Richard Hoistad, MD
Department of Otolaryngology–Head and Neck Surgery
Northwestern University
Chicago, Illinois

Stephen W. Hone, MD
Department of Otolaryngology
University of Iowa
Iowa City, Iowa

Ian N. Jacobs, MD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

Margaret M. Jastreboff, PhD
Department of Otolaryngology
Emory University
Atlanta, Georgia

Pawel J. Jastreboff, PhD, ScD
Department of Otolaryngology
Emory University
Atlanta, Georgia

Herman A. Jenkins, MD
Department of Otolaryngology
University of Colorado
Denver, Colorado

Tae Hoon Jinn, MD
Division of Otolaryngology–Head and Neck Surgery
Loma Linda University
Loma Linda, California

John K. Joe, MD
Department of Surgery
Memorial Sloan-Kettering Cancer Center
New York, New York

Jacob Johnson, MD
Department of Otolaryngology–Head and Neck Surgery
University of California
San Francisco, California

Timothy T. K. Jung, MD, PhD
Division of Otolaryngology–Head and Neck Surgery
Loma Linda University
Loma Linda, California

William M. Keane, MD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

Elizabeth M. Keithley, PhD
Division of Otolaryngology–Head and Neck Surgery
University of California
San Diego, California

David W. Kennedy, MD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

Janardan D. Khandekar, MD
Department of Medicine
Northwestern University
Chicago, Illinois

Young-Ho Kim, MD, PhD
Department of Otolaryngology
Yonsei University
Seoul, Korea

Charles P. Kimmelman, MD
Department of Otolaryngology
Cornell University
New York, New York

Karen Iler Kirk, PhD
Department of Otolaryngology
University of Indiana
Indianapolis, Indiana

Robert I. Kohut, MD
Department of Otolaryngology
Bowman Gray School of Medicine
Winston-Salem, North Carolina

James A. Koufman, MD
Department of Otolaryngology
Bowman Gray School of Medicine
Winston-Salem, North Carolina

Dennis H. Kraus, MD
Department of Otolaryngology
Cornell University
New York, New York

Stephen Y. Lai, MD, PhD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

Anil K. Lalwani, MD
Department of Otolaryngology–Head and Neck Surgery
University of California
San Francisco, California

Paul R. Lambert, MD
Department of Otolaryngology
University of South Carolina
Charleston, South Carolina

Andrew P. Lane, MD
Department of Otolaryngology–Head and Neck Surgery
Johns Hopkins University
Baltimore, Maryland

M. Samantha Lewis, MA
Department of Communicative Sciences and Disorders
University of Florida
Gainesville, Florida

David Litman, MD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

David J. Lim, MD
House Ear Institute
Los Angeles, California

Brenda L. Lonsbury-Martin, PhD
Department of Otolaryngology
University of Colorado
Denver, Colorado

Louis D. Lowry, MD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

Frank E. Lucente, MD
Department of Otolaryngology
State University of New York
Brooklyn, New York

Christy L. Ludlow, PhD
Laryngeal and Speech Section
National Institute of Neurological Disorders and Stroke
Bethesda, Maryland

Anne E. Luebke, PhD
Department of Otolaryngology
University of Miami
Miami, Florida

Valerie J. Lund, MS, FRCS, FRCS (Ed)
Department of Otolaryngology
University College of London
London, England

Rodney P. Lusk, MD
Department of Otolaryngology
Washington University
St. Louis, Missouri

Mahmood F. Mafee, MD
Department of Radiology
University of Illinois
Chicago, Illinois

Eric A. Mann, MD, PhD
Laryngeal and Speech Section
National Institute of Neurological Disorders and Stroke
Bethesda, Maryland

Glen K. Martin, PhD
Department of Otolaryngology
University of Colorado
Denver, Colorado

Elizabeth G. Miller, RD
Clinical Nutrition Services
Virginia Commonwealth University
Richmond, Virginia

John H. Mills, PhD
Department of Otolaryngology
University of South Carolina
Charleston, South Carolina

Fred D. Minifie, PhD
Department of Speech and Hearing Sciences
University of Washington
Seattle, Washington

Richard T. Miyamoto, MD
Department of Otolaryngology
University of Indiana
Indianapolis, Indiana

Kris S. Moe, MD
Department of Surgery
University of California
San Diego, California

C. Elliott Morgan, DMD, MD
Department of Otolaryngology–Head and Neck Surgery
University of Alabama
Birmingham, Alabama

Fabiola Müller, Dr. habil rer nat
School of Medicine
University of California
Davis, California

Robert M. Naclerio, MD
Department of Otolaryngology
University of Chicago
Chicago, Illinois

H. Bryan Neel III, MD, PhD
Department of Otolaryngology
Mayo Clinic and Medical School
Rochester, Minnesota

Selcuk Onart, MD
Department of Otorhinolaryngology
Uludag University
Bursa, Turkey

Frank G. Ondrey, MD, PhD
Department of Otolaryngology
University of Minnesota
Minneapolis, Minnesota

Ronan O'Rahilly, MD, DSc
School of Medicine
University of California
Davis, California

Lisa A. Orloff, MD
Department of Surgery
University of California
San Diego, California

Robert H. Ossoff, MD
Department of Otolaryngology
Vanderbilt University
Nashville, Tennessee

Simon C. Parisier, MD
Department of Otolaryngology
Mount Sinai School of Medicine
New York, New York

Philip Passalaqua, MD
Department of Otolaryngology
Mount Sinai School of Medicine
New York, New York

Jayant M. Pinto, MD
Department of Otolaryngology
University of Chicago
Chicago, Illinois

James S. Reilly, MD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

Lou Reinisch, PhD
Department of Physics and Astronomy
University of Canterbury
Christchurch, New Zealand

David Reiter, MD, DMD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

John S. Rhee, MD
Department of Otolaryngology
Medical College of Wisconsin
Milwaukee, Wisconsin

Kristina W. Rosbe, MD
Department of Otolaryngology–Head and Neck Surgery
University of California
San Francisco, California

Marc R. Rosen, MD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

Lee D. Rowe, MD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

Michael Ruckenstein, MD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

Allen F. Ryan, PhD
Division of Otolaryngology–Head and Neck Surgery
University of California
San Diego, California

Leonard P. Rybak, MD, PhD
Department of Otolaryngology
University of Southern Illinois
Springfield, Illinois

Clarence T. Sasaki, MD
Section of Otolaryngology
Yale University
New Haven, Connecticut

Cecelia E. Schmalbach, MD
Department of Otorhinolaryngology
University of Michigan
Ann Arbor, Michigan

Mitchell K. Schwaber, MD
St. Thomas Neuroscience Institute
Nashville Hearing and Balance Center
Nashville, Tennessee

Samir Shah, MD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

Neil T. Shepard, PhD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

Richard J. H. Smith, MD
Department of Otolaryngology–Head and Neck Surgery
University of Iowa
Iowa City, Iowa

Joseph C. Sniezek, MD
Department of Otolaryngology
Tripler Army Medical Center
Honolulu, Hawaii

David Solomon, MD, PhD
Department of Neurology
University of Pennsylvania
Philadelphia, Pennsylvania

Jeffery Staab, MD
Department of Psychiatry
University of Pennsylvania
Philadelphia, Pennsylvania

Scott E. Strome, MD
Department of Otolarygnology
Mayo Medical School
Rochester, Minnesota

Jonathan M. Sykes, MD
Department of Otolaryngology
University of California
Sacramento, California

M. Eugene Tardy Jr, MD
Department of Otolaryngology–Head and Neck Surgery
University of Illinois
Chicago, Illinois

Steven A. Telian, MD
Department of Otorhinolarygnology
University of Michigan
Ann Arbor, Michigan

Lawrence W. C. Tom, MD
Department of Otorhinolaryngology–Head and Neck Surgery
University of Pennsylvania
Philadelphia, Pennsylvania

John Touliatos, MD
Department of Otolaryngology
University of Southern Illinois
Springfield, Illinois

Galdino E. Valvassori, MD
Departments of Radiology and Otolaryngology
University of Illinois
Chicago, Illinois

Carter Van Waes, MD, PhD
Head and Neck Surgery Branch
National Institute on Deafness and Other Cummunication Disorders
Bethesda, Maryland

Jeffrey Vierra, MD
Department of Medicine
State University of New York
Brooklyn, New York

Phillip A. Wackym, MD
Department of Otolaryngology
Medical College of Wisconsin
Milwaukee, Wisconsin

Robert A. Weisman, MD
Department of Surgery
University of California
San Diego, California

Richard J. Wiet, MD
Department of Otolaryngology–Head and Neck Surgery
Northwestern University
Chicago, Illinois

Peak Woo, MD
Department of Otolaryngology
Mount Sinai School of Medicine
New York, New York

Simon K. Wright, MD
Department of Otolaryngology
University of Minnesota
Minneapolis, Minnesota

Doris K-W. Wu, PhD
Laboratory of Molecular Biology
National Institute on Deafness and Other Communication Disorders
Rockville, Maryland

Jane Y. Yang, MD
Department of Otolaryngology
Thomas Jefferson University
Philadelphia, Pennsylvania

CHAPTER 1

Anatomy of the Auditory and Vestibular Systems

Richard R. Gacek, MD, Mark R. Gacek, MD

The temporal bone is a complex portion of the skull base that contains the labyrinth with its nerve supply (cranial nerve VIII) but also other cranial nerves such as the facial, trigeminal, vagus, glossopharyngeal, spinal accessory, and hypoglossal nerves. A thorough knowledge of the gross and microscopic anatomy[1,2] of the temporal bone (TB) and the physiology of the labyrinthine sense organs is essential for the specialist who strives for accuracy in diagnosis and precision in surgery of the TB. This knowledge is gained first from dissection of cadaveric whole TB specimens but is greatly enhanced by study of prepared histologic sections from normal and pathologic TB.

OSTEOLOGY

Four major components of the temporal bone (TB) contribute to the skull base: the squamous, mastoid, tympanic, and petrous.

The *squamous* portion of the TB provides attachment for the temporalis muscle, which is bounded inferiorly by the temporal line (Figure 1–1). The temporal line provides an external landmark for the floor of the middle cranial fossa. The zygomatic process projects forward from the lower portion of this bone, and together they form the anterior border of the mandibular fossa, which receives the condyle of the mandible.

The *tympanic* portion of the TB is an incomplete cylindrical portion of the TB that, together with the squamosal portion, forms the medial part of the external auditory canal. This portion of the external auditory canal is 2 cm in length by 1 cm in diameter. Its anterior boundary is the posterior limit of the mandibular fossa; medially, its border is the tympanic membrane. The posterior part fuses with the mastoid component of the TB at the tympanomastoid suture. Failure in development of this part of the TB is responsible for congenital aural atresia, a form of conductive hearing loss correctable by surgery.

The major portion of the TB formed by the *mastoid* portion attributes its large size to extensive pneumatization. The mastoid process projects posteriorly and inferiorly behind the external auditory meatus and serves as the attachment for the sternocleidomastoid muscle. A deep groove in its inferior aspect houses the posterior belly of the digastric muscle, which is innervated by the facial nerve. The superior surface of the mastoid compartment is formed by a thin plate of bone known as the tegmen mastoidea. Posteriorly, it forms the anterior plate of the posterior cranial fossa and is indented by a groove for the sigmoid sinus. The superior and inferior petrosal sinuses travel medially along the superior and inferior aspects of this part of the TB.

The *petrous* portion of the TB forms its medial part inferior to the middle cranial fossa; posteriorly,

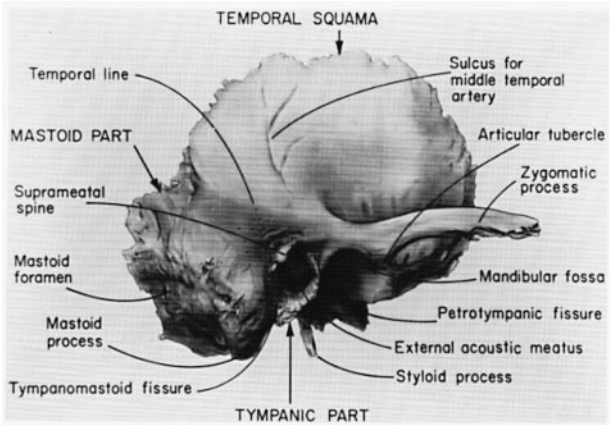


FIGURE 1–1. Right temporal bone, lateral view. Reproduced with permission from Anson BJ and Donaldson JA.[2]

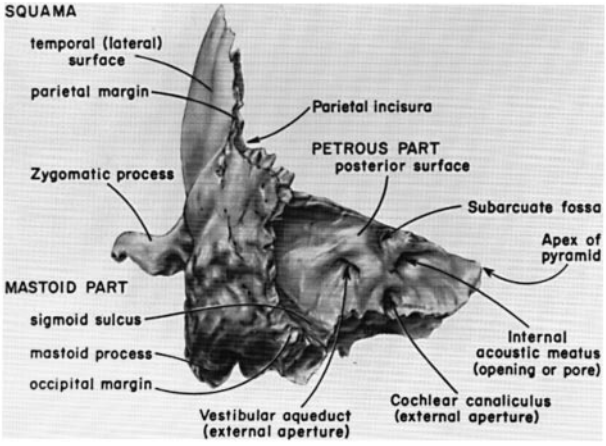


FIGURE 1–2. Left temporal bone, posterolateral view. Reproduced with permission from Anson BJ and Donaldson JA.[2]

it forms the anterior surface of the posterior cranial fossa (Figure 1–2). The superior surface of the petrous bone is highlighted by the prominence of the superior semicircular canal, a landmark in surgery within the middle cranial fossa. Anterior to this portion of the petrous bone is the hiatus for the greater superficial petrosal nerve, which joins with the geniculate ganglion of the facial nerve. In some temporal bones, this hiatus is enlarged, and the geniculate ganglion may be exposed in the middle cranial fossa. Anterior and medial to this region is a concave area for the semilunar ganglion of the trigeminal nerve. On the posterior surface of the petrous bone are several important landmarks. The most obvious aperture is the internal auditory meatus (canal) that transmits the seventh and eighth cranial nerves as well as the labyrinthine artery or loop of the anterior inferior cerebellar artery. The lateral end (fundus) of the internal auditory canal (IAC) is divided horizontally by the falciform crest.[1–3] The superior compartment contains the facial nerve anteriorly and the superior division of the vestibular nerve posteriorly. The inferior compartment transmits the cochlear nerve anteriorly and the inferior division of the vestibular nerve posteriorly. The endolymphatic sac may be found in a depression covered by a bony shelf (operculum) anterior to the sigmoid groove. It narrows down into the vestibular aqueduct as the endolymphatic duct. The depression for the semilunar ganglion and the fifth cranial nerve on the anterior surface of the petrous bone also carries the sixth cranial nerve through a dural canal referred to as Dorello's canal. These two nerves may be involved in inflammatory or neoplastic processes that occupy the petrous apex and are responsible for the clinical syndrome known as Gradenigo's syndrome (fifth cranial nerve pain, diplopia from lateral rectus muscle palsy, and otorrhea).

AUDITORY SYSTEM

External Ear

The external or outer ear is that portion of the ear that is lateral to the tympanic membrane (Figure 1–3). It consists of the external auditory canal as well as the auricle and cartilaginous portion of the ear. The auricle is a semicircular plate of elastic cartilage characterized by a number of ridges or grooves. The major ridges of the auricle are the helix and antihelix, the tragus and antitragus, which surround the concha, which is the scaphoid depression posterior to the external auditory meatus. The cartilage of the external auditory meatus is continuous with that of the outer portion of the ear canal and auricle.

The external auditory canal is made up of a cartilaginous extension of the auricle in its outer half and the mastoid and tympanic portion of the TB in its medial half. It is bounded medially by the tympanic membrane and is lined with skin that is thin with little subcutaneous tissue medially but laterally contains numerous hair follicles and ceruminous and sebaceous glands. The bony external auditory canal averages 3½ cm in length, with a diameter of 1 cm. The tympanic membrane is composed of three layers: the outer squamous cell epithelial layer, the medial mucosal layer facing the middle ear, and the fibrous layer or tunica propria, forming the substance of the tympanic membrane.[3] The fibrous layer gives the tympanic membrane its shape and consistency. Radial fibers of the tunica propria insert into the manubrium, circumferential fibers providing strength without interfering with vibration, whereas tangential fibers reinforce the architecture of the tympanic membrane. These physical characteristics are important for the vibratory characteristics necessary for sound transmission.

The tympanic membrane is identified by a prominent landmark, the manubrium of the malleus, which is limited superiorly by its lateral or short process and inferiorly by a rounded end

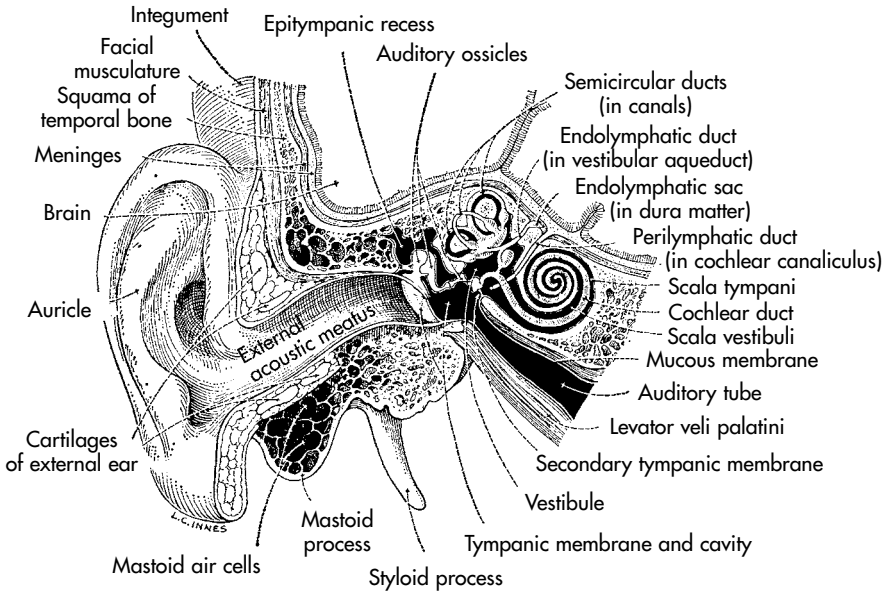


FIGURE 1–3. General relationships of parts of the ear (semidiagrammatic). Reproduced with permission from Anson BJ, McVay CB. Surgical anatomy. 5th ed. Philadelphia (PA): WB Saunders; 1981.

referred to as the umbo. The umbo forms the deep apex of the conical shape formed by the tympanic membrane. The tympanic membrane is incomplete superiorly, where it lacks a fibrous layer in the portion superior to the short process of the manubrium.[4] Since it lacks a fibrous layer, this portion is called the pars flaccida (Shrapnell's membrane). The major or inferior portion of the tympanic membrane is referred to as the pars tensa.

Middle Ear

The space between the tympanic membrane and the bony capsule of the labyrinth in the petrous portion of the TB contains the ossicular chain with its associated muscles, the aperture of the eustachian tube, and the vascular system. The tympanic cavity is divided into the epitympanic, mesotympanic, and hypotympanic regions. The *hypotympanic* portion is that portion of the middle ear that lies inferior to the aperture of the eustachian tube and the round window niche (RWN). This portion of the middle ear contains various bony trabeculae and the bony covering of the jugular bulb. This bony surface may be dehiscent, exposing the jugular bulb in the hypotympanic region. Inferiorly, a small channel (the inferior tympanic canaliculus) transmits Jacobson's nerve (a branch of cranial nerve IX).

The *mesotympanic* portion of the middle ear is limited superiorly by the horizontal portion of the facial canal and inferiorly by the RWN. This region contains the oval and round windows, the stapes bone, the stapedius muscle posteriorly, and the canal for the tensor tympani muscle anteriorly. The oval window is kidney bean shaped with a convex superior rim and a concave inferior rim. In the oval window, the footplate of the stapes bone is held in place by the annular ligament. The RWN forms a deep recess often covered with various mucous membrane configurations that obscure the round window membrane (RWM). The RWM is a fibrous membrane covered with a layer of mucosa that is roughly kidney bean shaped, with a major component anterior and inferior and a minor component located posteriorly and horizontally in the RWN. Posteriorly in the mesotympanum there are two bony recesses of clinical importance. The recess lateral to the vertical segment of the facial canal is called the facial recess. The space medial to the facial canal is called the sinus tympani (Figure 1–4). These two recesses are important clinically as they frequently harbor chronic middle ear infection and must be controlled in surgery. The facial recess also provides access to the middle ear space and RWN in those procedures in which the ear canal wall is preserved (ie, intact canal wall mastoidectomy, cochlear implantation). A bony projection from the facial canal (pyramidal eminence) contains the tendon of the stapedius muscle before its insertion into the neck of the stapes bone. The most anterior portion of the middle ear space is called the *protympanum* and is bordered superiorly by the orifice of the eustachian tube and anteriorly by the canal for the internal carotid artery (see Figure 1–4).

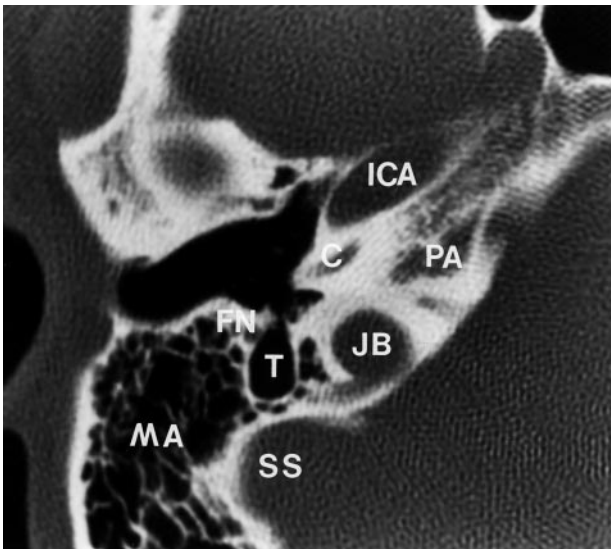


FIGURE 1–4. This axial computed tomographic scan of the temporal bone illustrates a normal mastoid cell system (MA), the horizontal segment of the internal carotid artery (ICA), the jugular bulb (JB), the sigmoid sinus (SS), and a nonpneumatized petrous apex (PA). C = basal turn of the cochlea; FN = facial nerve; T = sinus tympani.

The *epitympanum* is the portion of the middle ear that is limited superiorly by the bony roof of the middle ear called the tegmen tympani. This bony landmark is continuous posteriorly as the tegmen mastoidea. The medial wall of the epitympanum is formed by the bony prominence of the lateral and superior semicircular canal ampullae as well as the epitympanic portion of the facial (fallopian) canal. The head and neck of the malleus and its articulation with the body and short process and a portion of the long process of the incus occupy most of the space in the epitympanum. These two ossicular masses are held in place by ligaments anteriorly and posteriorly to provide an axis of rotation for the ossicular chain (Figure 1–5). The epitympanic space communicates posteriorly through a narrow opening called the aditus ad antrum to the central mastoid tract of the mastoid cavity. Anteriorly, the epitympanum is separated at the cochleariform process from an anterior epitympanic cell of variable size by a bony and mucous membrane barrier, which may completely or incompletely separate the two compartments. This anterior epitympanic space is formed by pneumatization from the protympanum (see Figure 1–5). The anterior epitympanic space is also important surgically as it may contain inflammatory tissue (ie, cholesteatoma) that has extended from the protympanum.

Auditory Ossicles Sound pressure energy is transmitted from the tympanic membrane across the middle ear space by the ossicular chain comprised of the malleus, incus, and stapes (Figure 1–6). The head of the malleus and body of the incus function as a unit suspended by ligaments in the epitympanum. The tip of the long process of the incus articulates at a right angle with the head of the stapes so that the sound energy transmission initiated by medial displacement of the tympanic membrane is carried by the parallel displacement of the elongate processes of the malleus and incus to the head, crura, and footplate of the stapes (see Figure 1–6). Since the surface area of the tympanic membrane is larger than that of the stapes footplate by a ratio of 25 to 1, the sound pressure density in the oval window and the inner ear fluids is similarly increased. Maintaining this ratio by various reconstructive methods constitutes

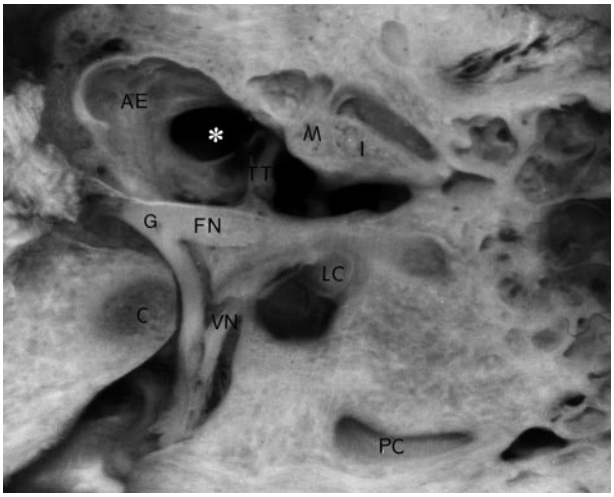


FIGURE 1–5. This horizontal cut through a celloidin-embedded temporal bone illustrates the relationship of the facial nerve (FN) to the superior division of the vestibular nerve (VN) in the internal auditory canal. The axis of rotation of the head of the malleus (M) and body of the incus (I) with their ligamentous attachments in the epitympanum is shown. AE = anterior epitympanic space ventilated into the protympanum (*); LC = lateral canal crista and ampulla; PC = posterior semicircular canal; C = endosteum of the cochlea (basal turn); TT = tensor tympani tendon; G = geniculate ganglion.

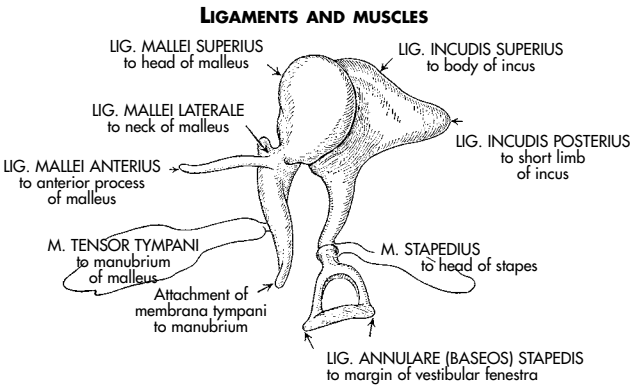


FIGURE 1–6. Auditory ossicles: their adult form and muscles and ligaments. Reproduced with permission from Anson BJ, editor. Morris' human anatomy. 12th ed. New York: McGraw-Hill; 1966.

an important principle in middle ear surgery. The stapes therefore acts in a piston-like fashion in the oval window. The stapes bone is shaped like a stirrup with a head, neck, and footplate or base. The crura are bowed, the posterior one more so than the anterior, and fused with the footplate, which is formed from both otic capsule and periosteal bone. These auditory ossicles are controlled to some degree by two middle ear muscles, the tensor tympani and the stapedius. The tensor tympani muscle is housed in a bony semicanal in the anterior mesotympanum just superior to the orifice of the eustachian tube (Figure 1–7). The muscle converges posteriorly into a tendinous segment that is anchored at the cochleariform process and turns abruptly lateralward to insert on the neck of the malleus. The tensor tympani is innervated by a branch of the fifth cranial nerve. Its motoneurons are located centrally in the parvocellular division of the trigeminal motor nucleus, and its action causes the drumhead to be pulled medially, thus raising the resonant frequency of the sound conduction system.

The stapedius muscle arises within either its own or the fallopian canal and is accompanied by the motor portion of the facial nerve. It converges superiorly and anteriorly to form the stapedius tendon, which emerges through the pyramidal eminence to insert at the neck of the stapes. The stapedius muscle is innervated by a branch of the seventh nerve, and its motoneurons are located in the brainstem in the interface between the facial nucleus and the lateral superior olivary nucleus. Contraction of the stapedius muscle displaces the stapes posteriorly and attenuates sound transmitted by the ossicular chain. Since reflex contraction of the stapedius muscle is activated by sound, it is regarded as a protective mechanism for the cochlea.

Eustachian Tube The eustachian tube is an essential communication between the nasopharynx and the middle ear (Figure 1–8). It is responsible for pneumatization of the middle ear and the mastoid and for maintaining normal pressure between the middle ear and the atmosphere. It represents the pharyngeal extension of the first branchial arch and extends from the lateral wall of the nasopharynx. The skeleton of the medial three-fourths of the eustachian tube is cartilage that is surrounded by soft tissue, adipose tissue, and respiratory epithelium. The cartilage of the eustachian tube, which is hook shaped on cross-section, is stabilized and displaced

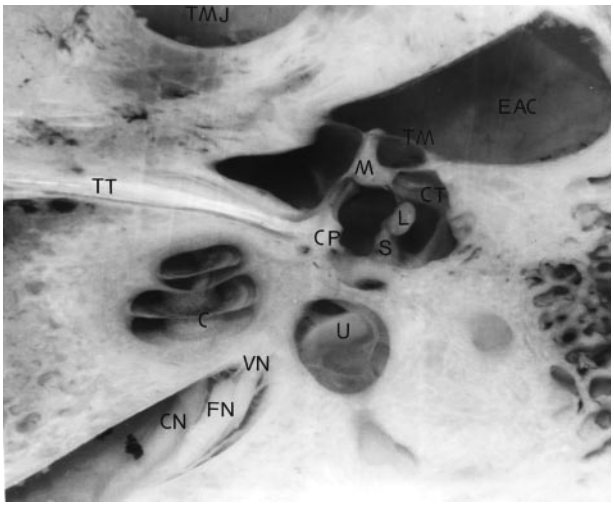


FIGURE 1–7. A more inferior cut through the same temporal bone as in Figure 1–5 demonstrates the cochlea (C), the utricular macula and its nerve (U), the cochlear nerve (CN), the facial nerve (FN), and vestibular nerves (VN) in the internal auditory canal. The muscle and tendon of the tensor tympani muscle (TT) overlie the cochlea as it turns laterally in the cochleariform process (CP) to attach near the neck of the malleus (M). The articulation of the long process (L) of the incus with the stapes head (S) can be seen in the mesotympanum. The chorda tympani nerve (CT) passes between the malleus and incus. TM = tympanic membrane; EAC = external auditory canal; TMJ = temporomandibular joint space.

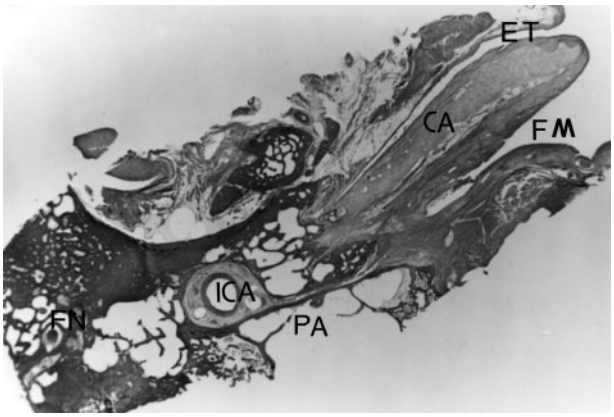


FIGURE 1–8. Low-power horizontal section through the eustachian tube (ET) as it passes into the protympanum. CA = cartilage of the ET; FM = fossa of Rosenmüller; ICA = internal carotid artery; FN = facial nerve; PA = petrous apex air cells.

by contraction of the tensor veli palatini and levator veli palatini muscles on swallowing or yawning. The eustachian tube is thereby opened, allowing for pressure equalization. The lining epithelium of the cartilaginous portion is similar to that of the pharynx with pseudostratified columnar cell epithelium and many mucous glands. Posterior to the union of the cartilaginous and osseous portion of the eustachian tube where the isthmus is located, the mucosa undergoes transition to cuboidal or low columnar cell epithelium similar to the tympanic cavity epithelium. Neoplastic compression of the eustachian tube lumen near its pharyngeal orifice (Rosenmüller's fossa) will cause fluid to fill the middle ear space (serous otitis media), usually in an adult patient (see Figure 1–8). Investigation of this occult region by endoscopy, radiologic imaging, and biopsy is necessary in such instances.

Nerve Supply of the External and Middle Ear The auricle and the external auditory canal receive the sensory nerve branches from the fifth nerve via the auriculotemporal nerve and the greater and lesser auricular nerves. Branches from the glossopharyngeal and vagus nerves also contribute to this innervation. The branch of the vagus nerve is referred to as Arnold's nerve, which travels in the posterior part of the ear canal in the posterior part of the tympanomastoid suture. When this nerve is stimulated, it produces a cough reflex as when the external auditory canal is being cleaned with an instrument. It may also participate in heralding a neoplastic or infectious process in distant regions of the aerodigestive tract also innervated by the vagus nerve (ie, larynx, hypopharynx) when pain is referred to the ear.

The main innervation to the middle ear space is through the tympanic plexus and Jacobson's nerve, which receives a major contribution from the glossopharyngeal nerve through the inferior tympanic canaliculus. This nerve travels in a bony sulcus or canal over the promontory along with the inferior tympanic artery anterior to the oval window and finally anteriorly to become the lesser superficial petrosal nerve. This nerve ultimately carries the fibers of the preganglionic neurons of the ninth nerve to the otic ganglion, where they synapse with postganglionic neurons and are carried over the auriculotemporal nerve to the parotid gland. Sympathetic fibers from the carotid plexus also contribute to the tympanic plexus. The chorda tympani nerve, which is a sensory branch of the facial nerve, will be discussed in the section on the facial nerve.

Mastoid Compartment The air cell system of the mastoid bone represents an extension of the air compartment in the middle ear from the first pharyngeal pouch. This process occurs in development of the TB and may result in a variable degree of pneumatization in the mastoid compartment (see Figure 1–4). Recurrent infection in the middle ear and mastoid has been identified as a factor that may limit the extent of pneumatization of the mastoid air cell system, whereas absence of such infection may favor full development of the air cell system. The air cells in the mastoid compartment extend from the aditus ad antrum in the epitympanum to the central mastoid tract (antrum) from which further extension in several directions may occur.[5] The *posterior superior* cell tract extends medially at the level of the superior semicircular canal toward the petrous apex (PA), and the *posterior medial* cell tract extends toward the PA at the level of the posterior semicircular canal. The *supralabyrinthine* cell system extends medially superior to the labyrinth, whereas the *retrofacial* cell system extends posteriorly and inferiorly along the bony ear canal to pneumatize the mastoid tip. These cell tracts may vary considerably and are important for the surgeon to know as a

guide in tracing infection into deep recesses of the mastoid compartment, particularly the PA.

Cell tracts arising from the middle ear space also are important for the surgeon. These cell tracts, particularly those that may lead to air cell development in the PA, are those that course inferior to the labyrinth or those that extend around the canal of the internal carotid artery (see Figure 1–8). Extensive pneumatization in the development of the PA may create air cells that can become isolated when the cell tract is obliterated by bone or fibrous tissue leading to the formation of cholesterol cysts over a period of many years. Over time, these cysts erode the surrounding bone and may reach considerable size in early or late adulthood. Compression of the trigeminal nerve and the sixth nerve near the PA may present a clinical picture similar to Gradenigo's syndrome, which is the clinical manifestation of petrous apicitis.

The PA is of special interest because of a variety of lesions that may involve this region. The clinical manifestation may be subtle and requires a high index of suspicion to pursue the diagnosis.[6] Imaging (both computed tomography and magnetic resonance imaging) is especially sensitive to the identification of pathology in the PA. Since the composition of the PA can include air cells (see Figure 1–8), bone marrow (Figure 1–4), the internal carotid artery (Figure 1–4), and the cartilage of the foramen lacerum, the list of lesions that occur here is lengthy. Cholesterol cysts and congenital cholesteatomas are the most common; however, infection, bone marrow neoplasms, cartilage tumors, metastatic malignancies, neurogenic tumors, and aneurysms of the internal carotid artery have been reported. Clinical signs of a progressive lesion in the PA relate to nearby structures: eustachian tube obstruction, facial pain or anesthesia, and lateral rectus muscle palsy.

Inner Ear

The petrous portion of the TB houses the labyrinth with its attendant sensory structures responsible for auditory and balance function. Within the bony labyrinth is contained the membranous labyrinth, which represents a continuous series of epithelial lined tubes and spaces of the inner ear containing endolymph and the sense organs of hearing and balance. The membranous labyrinth can be divided into three regions that are interconnected: the pars superior or the vestibular labyrinth with the exception of the saccule, the pars inferior (cochlea and the saccule), and the endolymphatic duct and sac. All of the sense organs of the labyrinth have in common that they contain hair cells with rigid cilia and are innervated by afferent and efferent neurons.[7,8] Displacement of the cilia of the hair cells is responsible for opening potassium and calcium channels that initiate the electrical potential within the hair cell that is then leaked into the afferent neuron and carried to the brainstem.

Cochlea The cochlear duct, the auditory portion of the labyrinth, extends approximately 35 mm.[9] The cochlear duct and associated sensory and supportive structures assume the form of a spiral similar to a snail shell of 2½ to 2¾ turns (Figure 1–9). This allows the long cochlear duct to be contained in a small space. A cross-section of a cochlear turn (Figure 1–10) demonstrates the essential structures in this sense organ. The scala media or cochlear duct con-

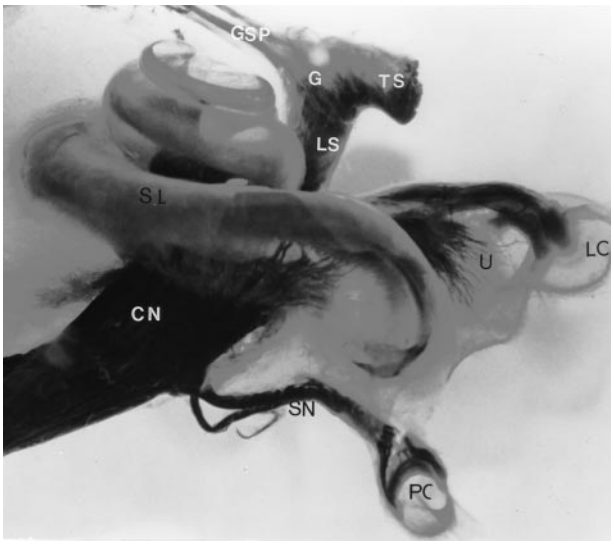


FIGURE 1–9. Photograph of a dissection of the human labyrinth and its nerve supply that demonstrates the 2½ turns of the cochlear duct and the spiral ligament (SL). U = utricular nerve and macula; LC = lateral duct ampulla; PC = posterior duct ampulla; SN = singular nerve. Facial nerve: labyrinthine segment (LS) and tympanic segment (TS). G = geniculate ganglion; GSP = greater superficial petrosal nerve; CN = cochlear nerve. Reproduced with permission from Gacek RR. Membranous inner ear. Ann Otol Rhinol Laryngol 1961;70:974–5.

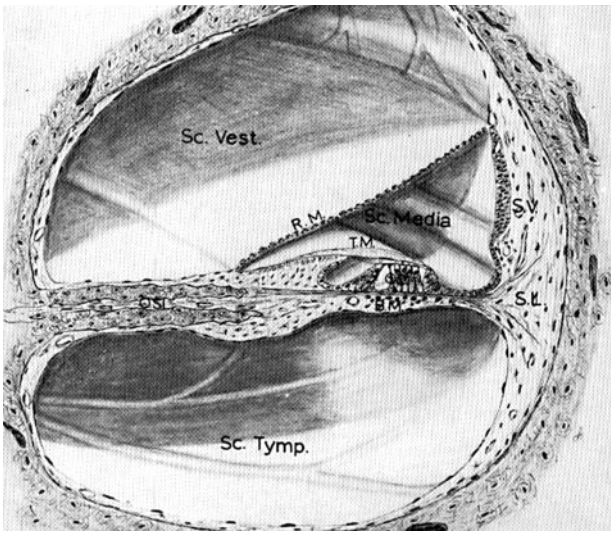


FIGURE 1–10. Section (diagrammatic) through the cochlea. Sc. Vest. = scala vestibuli; R.M. = Reissner's membrane; Sc. Media = scala media; T.M. = tectorial membrane; O.C. = organ of Corti; B.M. = basilar membrane; S.V. = stria vascularis; S.L. = spiral ligament; O.S.L. = osseous spiral lamina; Sc. Tymp. = scala tympani.

taining endolymph is triangular in shape in cross-section. The basilar membrane forms the horizontal limb of the triangle, Reissner's membrane, the superior limb, and the stria vascularis with spiral ligament on the vertical side. The cochlear duct is filled with a fluid referred to as endolymph, whereas the fluid in the scala vestibuli and scala tympani is perilymph. Perilymph of the two scalae communicates through the helicotrema at the apex of the cochlea. All of the structures of the cochlear duct and, particularly, the basilar membrane have a morphologic gradient whereby the width of the basilar membrane is narrowest at the basal end and widest at the apex. The spiral ligament and epithelial elements in the organ of Corti also have a morphologic gradient from base to apex (see Figure 1–9). This morphologic gradient, to a large degree, determines the location of maximal stimulation of the basilar membrane and inner hair cells by a given tone or frequency that is introduced to the inner ear. In this way, high frequencies are located at the base and low frequencies at the apex, with the frequency scale laid out in an orderly fashion over the remainder of the basilar membrane. The cochlear duct ends in a blind pouch (cecum) that is located near the RWM.

Perilymph of the scala vestibuli fills the vestibule under the stapes footplate. This perilymphatic compartment extends up the scala vestibuli of the cochlea and communicates with the perilymph in the scala tympani, which extends down the cochlea to terminate at the RWM. The perilymphatic compartment also communicates with the subarachnoid space through the periotic duct by way of the cochlear aqueduct that is filled with a trabecular meshwork of connective tissue capable of allowing some exchange of cerebrospinal fluid and perilymph. However, perilymph is primarily formed by filtration from the vascular network in the spiral ligament.

The organ of Corti is a complex sense organ that contains inner and outer hair cells and supporting cells resting on the basilar membrane, with the ciliated ends of the hair cells protruding into or near a covering structure, the tectorial membrane (Figure 1–11). The apical portions of the hair cells are anchored in the cuticular plate, with the stereocilia (usually 100 to 150 per cell) protruding through the cuticular plate.[7] The stereocilia of the outer hair cells make contact with the tectorial membrane, whereas the stereocilia of the inner hair cells lie free in the endolymphatic space inferior to the tectorial membrane. There are a single row of inner hair cells and three to five rows of outer hair cells. These cells differ morphologically in that the inner hair cells are more flask shaped and tightly surrounded by supporting cells and have stereocilia that are arranged in a linear fashion, whereas the outer hair cells are columnar and incompletely surrounded by phalangeal or supporting cells lying free in the perilymph of the organ of Corti.[10] The stereocilia of the outer hair cells form an inverted "W,"

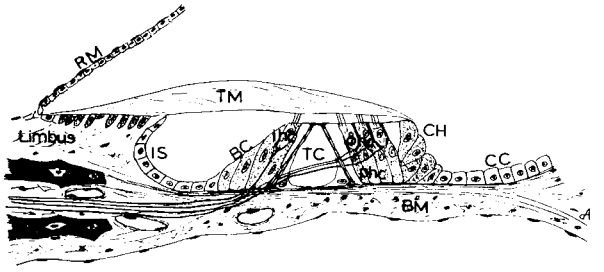


FIGURE 1–11. Detail of organ of Corti. RM = Reissner's membrane; TM = tectorial membrane; IS = inner sulcus; BC = border cells; ihc = inner hair cells; TC = tunnel of Corti; ohc = outer hair cells; phc = phalangeal cells; CH = cells of Hensen; CC = cells of Claudius; BM = basilar membrane.

and a basal body representing a rudimentary kinocilium is located on the spiral ligament side of the ciliary tuft. The inner hair cells are supported by interphalangeal cells, whereas the outer hair cells are supported by Deiters' cells inferiorly and laterally by Hensen's cells. The tectorial membrane is anchored medially at the limbus and attached to the Hensen's cells laterally by a fibrous net. The basilar membrane and tectorial membrane are displaced vertically by the traveling wave created by sound energy delivered to the oval window. Since the fulcrum of these two structures is separate, they will slide horizontally when stimulated, resulting in a shearing action between the tectorial membrane and the cuticular plate. The resultant displacement of stereocilia initiates an electrical event in the hair cell. The organ of Corti contains approximately 15,500 hair cells, with about 3,500 of them being inner hair cells and 12,000 being outer hair cells. These hair cells are innervated by afferent and efferent neurons in a complex but orderly manner. The afferent neurons to the auditory sense organ are bipolar neurons referred to as spiral ganglion cells that are located in Rosenthal's canal of the bony modiolus. Approximately 30,000 spiral ganglion cells innervate the human organ of Corti (see Figure 1–11). The spiral ganglion takes the form of clusters of ganglion cells throughout the extent of the length of the cochlea. Ninety to 95% of the spiral ganglion neurons are type I neurons, which are large and myelinated and project a single dendrite directly to an inner hair cell.[10] Approximately 10 to 20 type I spiral ganglion cells innervate one inner hair cell. These form the major afferent input from stimulation of the organ of Corti. Type I ganglion cells degenerate readily following injury to the dendrite. The remaining 5% of afferent neurons in the spiral ganglion are type II ganglion cells, which are smaller and unmyelinated and have very thin distal processes. The dendrites of these type II neurons cross the tunnel space along its floor enveloped by pillar cell processes and form spiral bundles between Deiters' cells. These dendrites then course apically between Deiters' cells to innervate several to many outer hair cells per type II dendrite.[9] A type II ganglion survives following injury to its dendrite. The axons of the spiral ganglion cells project to the cochlear nucleus complex, which has anteroventral and posteroventral divisions of the ventral cochlear nucleus and the dorsal cochlear nucleus. Each type I afferent neuron bifurcates and also sends a trifurcating branch to the dorsal cochlear nucleus in an orderly fashion according to frequency.[9] Apical turn neurons terminate in the most medial portion of the nuclear complex, whereas the basal turn neurons terminate laterally. Remaining frequency projections are ordered between these two regions of the cochlear nucleus. The central termination of the type II ganglion cells is not known largely because the small caliber axons are difficult to trace for long distances.

This frequency organization of the auditory pathway characterizes the remainder of the afferent pathway from end-organ to cortex. Another feature of the afferent auditory pathway is that the numbers of neurons involved at the various nuclear way stations undergo a progressive increase from cochlear nucleus to the cortex.[11] Although there are 30,000 spiral ganglion cells in the monkey auditory nerve, 88,000 neurons are found in one cochlear nucleus in the primate. One superior olivary complex contains 34,000 neurons, whereas the nucleus of the lateral lemniscus has 38,000 neurons, and at the inferior colliculus level there are almost 400,000 neurons on each side and at the medial geniculate body 500,000. The auditory cortex has approximately 10 million neurons.

A brief description of the afferent auditory pathway follows (Figure 1–12). The cells of the dorsal cochlear nucleus project axons to the dorsal acoustic striae, which crosses the midline and ascends in the contralateral lateral lemniscus to terminate in the dorsal nucleus of the lateral lemniscus and the inferior colliculus, particularly its inferior half. The cell bodies of the ventral cochlear nucleus project axons to the ipsilateral accessory and main superior olivary nuclei and to the medial dendrites of the contralateral accessory olive. The neurons of the accessory olive have bipolar dendrites arranged horizontally. This arrangement is favorable to receive input from projections of both cochlear nuclei. As such, it is an important nuclear way station for determining sound localizaton. Some fibers of the intermediate and ventral cochlear striae travel beyond the superior olivary complex and enter the contralateral lateral lemniscus to terminate in the inferior colliculus. The superior olive is thought to function as both a relay station for the auditory pathway and as a reflex center. The best-known reflex mediated through the superior olive is the stapedius reflex. Stapedius muscle motoneurons are

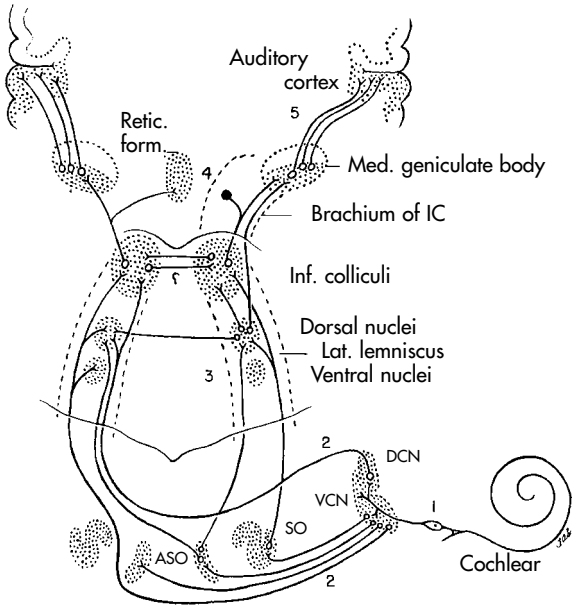


FIGURE 1–12. Diagram of the neuronal linkage that serves the afferent auditory pathway from one cochlea. Numerals indicate order of neuron units in the pathway. VCN = ventral cochlear nucleus; DCN = dorsal cochlear nucleus; SO = lateral superior olivary nucleus; ASO = accessory superior olivary nucleus; IC = inferior colliculus; Retic. form. = reticular formation. Reproduced with permission from Gacek RR. Neuroanatomy of the auditory system. In: Tobias JV, editor. Foundations of modern auditory theory. Vol 2. New York: Academic Press; 1972. p. 242.

located in the interface between the superior olivary and the facial nerve nuclei. The accessory superior olive projects bilaterally in the lateral lemnisci to terminate in the dorsal nuclei of the lateral lemnisci and the inferior colliculi. The lateral superior olive projects homolaterally in the lateral lemniscus to terminate in the dorsal nucleus of the lateral lemniscus and also in the inferior colliculus. No neurons from the superior olivary nuclei project beyond the inferior colliculus.

Projections from the inferior colliculus are primarily to the medial geniculate body. However, some projections to the medial geniculate body are received from the nuclei of the lateral lemniscus. All ascending neurons terminate in the medial geniculate body so that the final projection pathway to the auditory cortex, which is a major one, is from the medial geniculate body to the auditory cortex. Furthermore, the only commissural or interconnections between the two sides of the auditory pathway are at the superior olivary level, the level of the nuclei of the lateral lemniscus, and the inferior colliculus. No commissural projections are present superior to the inferior colliculus.

The ascending auditory pathway, although comprised of four to five neurons in the linkage from end-organ to auditory cortex and having an increasing volume of neural units active at each level, nevertheless is precisely organized according to the frequency scale and project bilaterally but predominantly in a contralateral pathway to the auditory cortex.

Efferent Auditory Pathways Paralleling the afferent auditory pathway is a descending pathway originating in the auditory cortex and terminating in the end-organ (Figure 1–13). This pathway does not involve as many neurons as the ascending or afferent pathway but has the feature of extensive ramification and formation of many terminals, which contact a large number of neurons. Nevertheless, the efferent

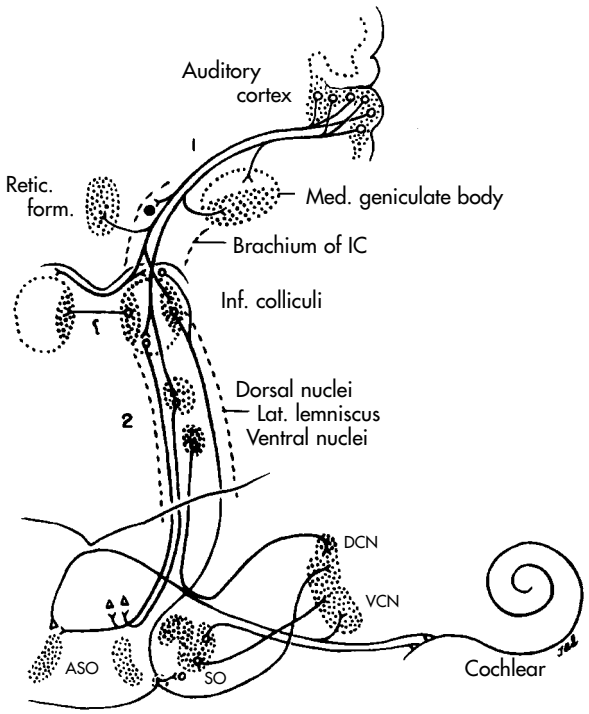


FIGURE 1–13. Diagram of the descending (efferent) auditory pathway. IC = inferior colliculus; DCN = dorsal cochlear nucleus; VCN = ventral cochlear nucleus; ASO = accessory superior olivary nucleus; SO = lateral superior olivary nucleus. Reproduced with permission from Gacek RR. Neuroanatomy of the auditory system. In: Tobias JV, editor. Foundations of modern auditory theory. Vol 2. New York: Academic Press; 1972. p. 253.

pathway does not make as many neural contacts in the auditory nuclei as the ascending pathway. The descending auditory pathway originates in the auditory cortex and initially projects to the inferior colliculus and the dorsal nucleus of the lateral lemniscus, with some termination in the medial geniculate body and the reticular formation.

The next neuron in the descending chain is located in peripheral regions of the inferior colliculus and the nuclei of the lateral lemniscus. These projections terminate in brainstem neurons that give rise to efferent neurons projecting to divisions of the cochlear nucleus and to neurons that give rise to the third and final neuron in the descending auditory pathway, the olivocochlear bundle, which innervates the organ of Corti.[12,13]

The olivocochlear bundle has both an ipsilateral and a contralateral limb or component (Figure 1–14). The neurons, which number approximately 1,000 to 3,000, arise from neurons located within and near the superior olivary complex. The contralateral limb of the olivocochlear bundle forms a major part of the efferent bundle, accounting for approximately three-fourths of the number of efferent neurons projecting to the organ of Corti in one ear. These axons arise from small neurons located near the accessory olivary nucleus,[14] ascend in the brainstem, and cross the midline at the level of the facial genu below the floor of the fourth ventricle. They are joined by the smaller ipsilateral component, which arises from similar olivary and periolivary neurons of the ipsilateral superior olivary nucleus before joining the contralateral limb as it enters the vestibular nerve root. As it leaves the brainstem in the vestibular nerve, the bundle gives off collateral branches to the ventral cochlear nucleus and exits the vestibular nerve within the IAC just distal to the saccular ganglion. It then enters Rosenthal's canal, where it travels perpendicular to the spiral ganglion cells and their dendrites, forming the intraganglionic or juxtaganglionic spiral bundle. The fibers from the efferent bundle are then given off regularly as they ascend the cochlea. These fibers penetrate the habenulae perforatae in the osseous spiral lamina along with afferent dendrites to enter the organ of Corti. The differential termination of the efferent neurons from the ipsilateral and the contralateral limbs is as follows: the fibers from the ipsilateral efferent component terminate on type I afferent dendrites and their terminals below the inner hair cell. The contralateral limb of the efferent

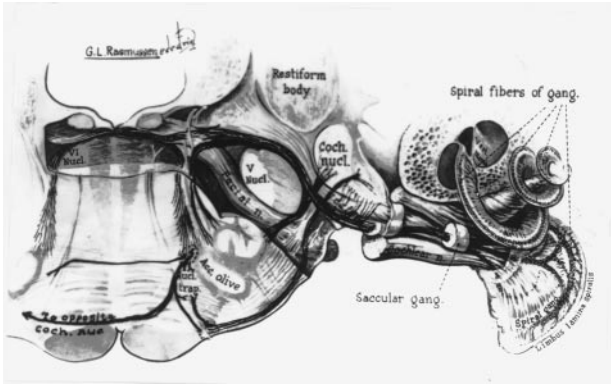


FIGURE 1–14. Drawing of the olivocochlear efferent pathway (from Rasmussen). See text for a description.

pathway crosses the tunnel space and ramifies extensively to terminate on several outer hair cells in the first and second rows in a wider area than the inner hair cells at that region. The efferent innervation to outer hair cells is most extensive in the basal turn and decreases as the apex is reached.[15] This decrease is most noticeable in the outermost rows of outer hair cells of the organ of Corti. Whereas the inner hair cell type I neuron innervation provides the major afferent input to the cochlear nucleus, the efferent innervation of the outer hair cells is thought to alter mechanically the resistance at the outer hair cell level by contractile changes in the length of the outer hair cells. In this way, the sensitivity to sound stimulation of type I–innervated hair cells is modified. The predominant effect on auditory nerve transmission by stimulation of the efferent pathway has been to suppress the action potential in the auditory nerve.[16] It is thought that the outer hair cells with their type II afferent innervation and the efferent innervation sharpen the characteristic frequency discrimination of the organ of Corti. It is also probable that the hair cells, with their spontaneous and induced transition in length, may be responsible for otoacoustic emissions, which are small electrical potentials recorded from the ear.[17] In some patients in whom the inhibitory effect of the efferent system is lost (ie, surgery, infection), the facilitated otoacoustic emissions may be perceived as ringing in the ear (tinnitus).

VESTIBULAR SYSTEM

The sense organs of the vestibular system are of two types, the cristae and the maculae (Figures 1–15 and 1–16). These sense organs have two types of hair

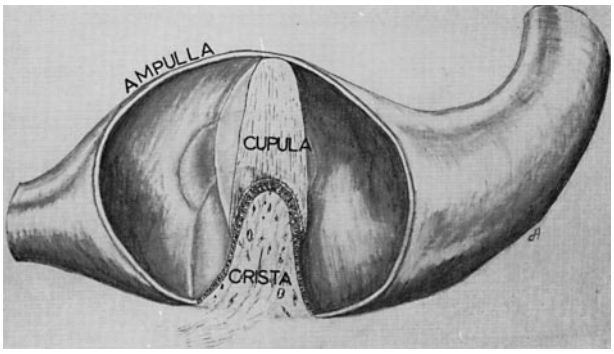


FIGURE 1–15. Section (diagrammatic) through the ampulla of a semicircular duct.

cells,[8,18] type I and type II, which are secured in the neural epithelium by supporting cells (see Figure 1–16). The type I hair cells are flask shaped and are enclosed in a large calyx-type ending by one or two large-diameter afferent neurons. Type II hair cells are cylindrical in shape and contacted by bouton-shaped endings from both the afferent and efferent systems. These afferent neurons are myelinated but small compared with those that innervate type I hair cells. The crista is a ridge of neuroepithelial cells that traverse the ampullated end of each membranous duct. The ampulla contains the sense organ for the detection of angular acceleration and deceleration[19] (see Figure 1–15). The hair cells in the crista ampullaris are arranged with the type I hair cells near the crest of the crista and type II hair cells near the slopes.[7,18] It is noteworthy that the hair cells in cristae of the semicircular ducts are oriented in a single direction based on the alignment of their ciliary bundle.[7,18] The ciliary bundle of each vestibular hair cell consists of 100 to 150 stereocilia and a single kinocilium, which is located to one side of the stereocilia (see Figure 1–16). In the case of the horizontal duct, the kinocilium of each hair cell is located on the utricular side of the crista. In the vertical cristae, the kinocilia are located toward the nonampullated ends or away from the utricle (Figure 1–17). Sitting on top of the crista and reaching to the roof of the ampullary wall, even attached to it, is a gelatinous partition called the cupula, composed of mucopolysaccharides possessing an equal density to the endolymph, which surrounds it (see Figure 1–15). It is the displacement of the cupula that initiates through the hair cells an action potential in the neurons contacting the hair cells.[19,20] When the deflection of the cupula is toward the kinocilium, there is a depolarization of the hair cell neuron unit with an increase in the resting vestibular nerve action potential, whereas a deflection away from the kinocilium produces hyperpolarization or a decrease in the resting discharge of the vestibular neuron.[7] The semicircular ducts in each labyrinth are oriented to each of three planes in space, with the lateral ducts recording angular acceleration in a horizontal plane, whereas the two vertical ducts record movement in the two vertical planes of space. To induce

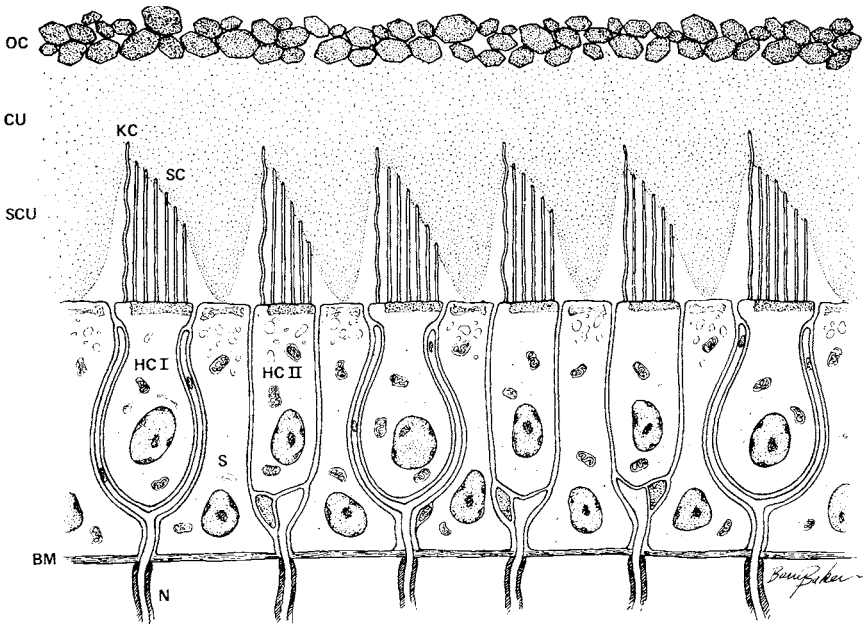


FIGURE 1–16. Schematic drawing of a cross-section of a macula. The gelatinous substance is divided into cupular (CU) and subcupular (SCU) layers. There are two types of hairs: kinocilia, labeled KC (one per hair cell), and stereocilia, labeled SC (many per hair cell). OC = otoconia; HC I = type I hair cell; HC II = type II hair cell; N = nerve fiber; BM = basement membrane; S = supporting cell.

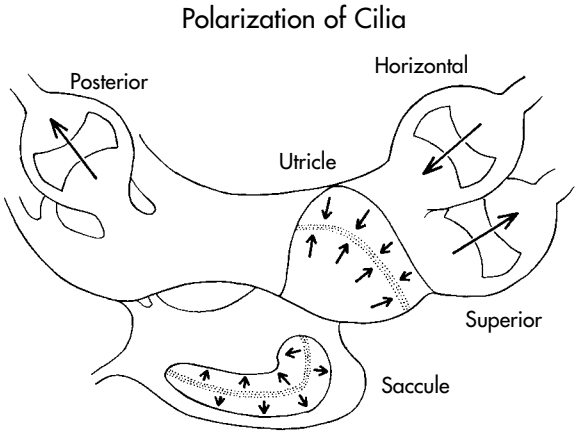


FIGURE 1–17. Diagram summarizing the orientation of hair cells in the vestibular sense organs based on the location of the kinocilium in the ciliary bundle.

complementary input to the vestibular nuclei, the ducts of each labyrinth are coplanar; in other words, the posterior duct of one labyrinth is coplanar with the anterior duct of the contralateral labyrinth. Because of the opposite polarization of hair cells in coplanar ducts, following rotation in a given plane, one coplanar duct will be facilitated and its contralateral partner will be inhibited.[21]

The maculae are flat sense organs that are divided approximately in half by a line called the striola. The striola divides the macula into two halves, in which the hair cell ciliary polarization is in opposite directions (see Figure 1–17). In the macula of the utricle, the hair cells are oriented toward the striola, whereas in the saccule, hair cells are oriented away from the striola. Since the maculae project centrally to the vestibulospinal pathways primarily, but to a lesser extent to the vestibulo-ocular pathways, it is possible that the opposite polarization of regions in the maculae allows excitation and inhibition to antagonistic muscle groups. Type I hair cells seem to predominate near the striola, with type II hair cells furthermost from the striola. The cilia of the hair cells in the macular sense organs are covered by an otolithic membrane, gelatinous in makeup with otoconia composed of calcium carbonate crystals with a specific gravity of 2.71 embedded in the gelatinous blanket. The movement of the otoconial membrane by gravitational forces or inertial forces displaces the hairs of the hair cells, thus bringing about activity in their afferent nerve input.

The innervation of hair cells in the vestibular sense organs is by both afferent and efferent neurons (see Figure 1–16). The afferent neurons can be divided into two types, the large afferent neurons belonging to the large bipolar ganglion cells in Scarpa's ganglia, which innervate type I hair cells with large calyx-like endings,[8,18] and the smaller afferent bipolar neurons in Scarpa's ganglia, which innervate type II hair cells with bouton-type terminals. Vestibular neurons have a high spontaneous activity with a higher range (90 to 100 spikes per second [sps]) in canal afferents than in macular afferents (60 to 70 sps). The large vestibular afferents characteristically show an irregular discharge pattern, whereas the smaller afferents have a regular pattern of discharge.[21] The significance of such different afferent units is not known but implies a functional difference. The efferent neurons are also small in diameter and consist of both myelinated and unmyelinated axons. They penetrate the basement membrane of the neuroepithelium and ramify extensively to form bouton-like terminals filled with many synaptic vesicles on both type II hair cells and the calyces of type I hair cells.

The afferent neurons to the vestibular sense organs are bipolar neurons of Scarpa's ganglion. There are approximately 18,000 to 19,000 ganglion cells in the human vestibular ganglion. The organization of the afferent neurons in the vestibular nerve has been elucidated.[22] The vestibular nerve has superior and inferior divisions, with the superior division innervating the cristae of the lateral and superior semicircular ducts and the utricle as well as a small portion of the saccule (Figure 1–18). The inferior division innervates the macula of the saccule and the posterior duct crista. The bipolar neurons innervating the lateral and the superior duct crista travel in the most rostral (toward the facial nerve) portion of the vestibular ganglion. As they course centrally, they are joined by axons of the bipolar neurons innervating the posterior duct crista to form the rostral portion of the vestibular nerve containing all semicircular duct afferents. The afferent neurons innervating the utricle then bend caudally to join those of the saccule and form the caudal portion of the vestibular nerve as it reaches the brainstem. Between these two groups of neurons is located the parent efferent bundle for both the cochlear and vestibular efferent axons. Consistent with the innervation pattern in the cristae ampullaris, the large afferents travel up the center of the ampullary nerve to innervate type I hair cells at the crests of the crista,[18,22] whereas the smaller fibers

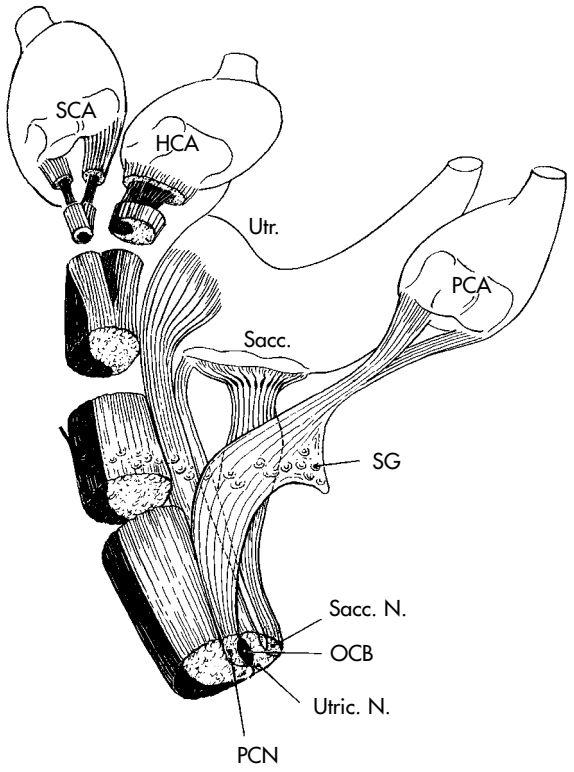


FIGURE 1–18. Drawing of the organization of bipolar neurons of the vestibular nerve. Ganglion cells in Scarpa's ganglion (SG) that innervate the cristae converge into the rostral two-thirds of the nerve trunk as it approaches the brainstem. The ganglion cells that supply the utricular and saccular maculae occupy the caudal third of the vestibular nerve trunk. The cochlear and vestibular efferent axons are located at the interface of these two divisions. The dark portion of the vestibular nerve signifies the location of large neurons that supply type I hair cells in the superior and lateral duct cristae. SCA = superior canal ampulla; HCA = horizontal canal ampulla; PCA = posterior canal ampulla; Utr. = utricle; Sacc. = saccule; Sacc N. = saccular nerve; Utric N. = utricular nerve; PCN = posterior canal nerve; OCB = efferent bundle (cochlear and vestibular). Reproduced with permission from Gacek RR.[22]

innervate the slopes and therefore surround the larger fibers in each ampullary nerve (see Figure 1–18). On the other hand, the distribution of the large and small afferents in the saccular and the utricular nerves is more dispersed since the division between the location of type I and type II hair cells in the maculae is not so precise as in the cristae.

Central Vestibular Pathways

The division of duct afferents in the rostral one half to two-thirds in the vestibular nerve and utricular afferents in the caudal portion of the vestibular nerve is related to their termination in the major vestibular nuclei in the brainstem.[22] The duct afferents after bifurcating send an ascending branch to the superior vestibular nucleus and a descending branch with many collaterals to the rostral portion of the medial vestibular nucleus, with some collaterals to the ventral division of the lateral vestibular nucleus (Figure 1–19). This projection is organized in that the duct innervated by the inferior division terminates most medially in the superior nucleus and most ventrally in the medial nucleus, whereas the two ducts supplied by the superior division of the vestibular nerve terminate dorsolaterally in the superior nucleus and dorsally in the medial nucleus.[22] Short collaterals from

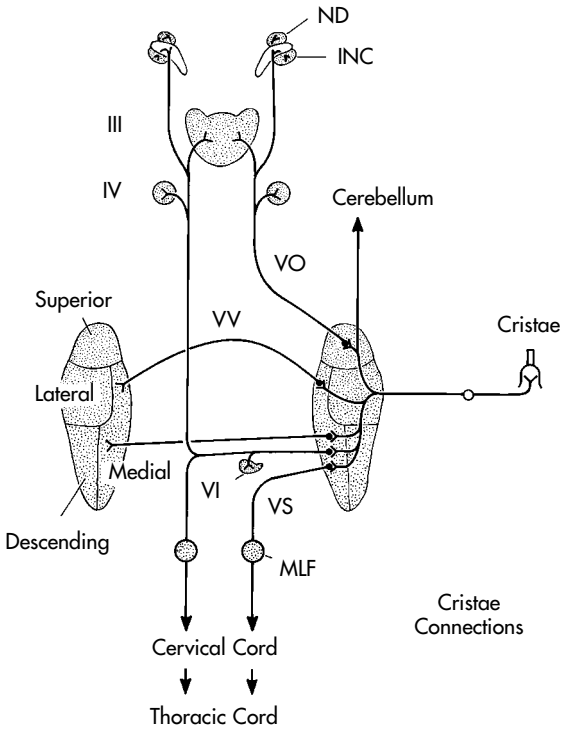


FIGURE 1–19. Diagram summarizing the central projections of afferents from the semicircular ducts. MLF = medial longitudinal fasciculus; VS = vestibulospinal; VV = commissural; VO = vestibulo-ocular; ND = nucleus of Darkeschewitch; INC = interstitial nucleus of Cajal III = oculomotor nucleus; VI = abducens nucleus; IV trochlear nucleus. Reproduced with permission from Gacek RR. Neuroanatomical correlates of vestibular function. Ann Otol Rhinol Laryngol 1980;89:1–5.

the incoming axons of duct afferents also terminate on the intravestibular nerve nucleus, a small nucleus within the vestibular nerve root. The afferents from the utricular and saccular maculae also bifurcate, with their ascending ramus terminating in the lateral and medial nuclei and the descending branch to the descending and group Y vestibular nuclei (Figures 1–20 and 1–21). The projection to the lateral nucleus is on neurons in the ventral division of the lateral vestibular nucleus and on large neurons in the rostral extension of the medial vestibular nucleus. The large neurons in the medial nucleus represent afferents to the abducens nucleus and to the subnucleus in the oculomotor nucleus, which innervates the medial rectus muscle. The latter projection pathway travels over the ascending tract of Deiters.[23] The saccular afferents also send terminations to the group Y nucleus, a minor nucleus that is located between the restiform body and the dorsal acoustic stria and the lateral vestibular nucleus. The group Y nucleus' major projection is to the contralateral group Y and superior nuclei.

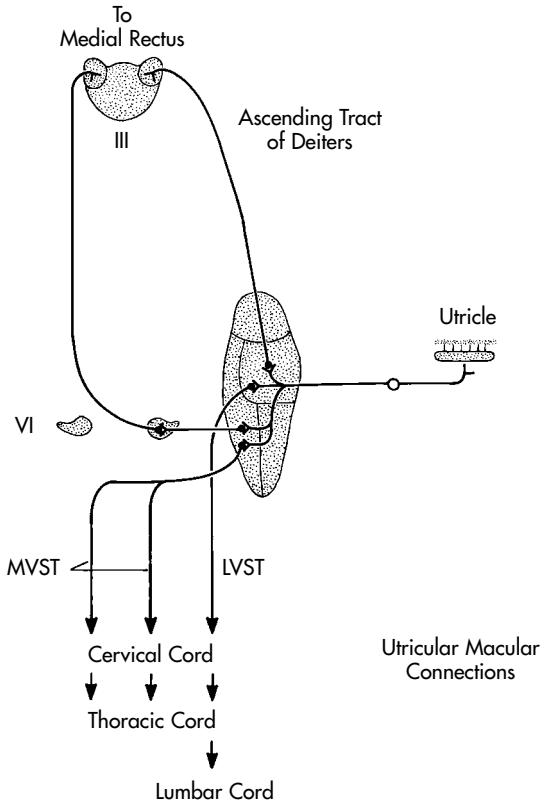


FIGURE 1–20. Central projections of afferents from the utricular macula. LVST = lateral vestibulospinal tract; MVST = medial vestibulospinal tract; III = oculomotor nucleus; VI = abducens nucleus. Reproduced with permission from Gacek RR. Neuroanatomical correlates of vestibular function. Ann Otol Rhinol Laryngol 1980;89: 1–5.

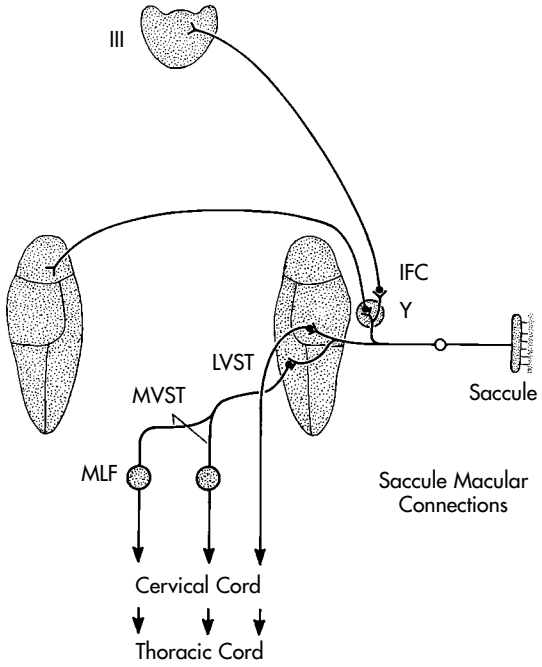


FIGURE 1–21. Central projections of afferents from the saccular macula. IFC = infracerebellar nucleus; Y = group Y nucleus; LVST = lateral vestibulospinal tract; MVST = medial vestibulospinal tract; MLF = medial longitudinal fasciculus; III = oculomotor nucleus. Reproduced with permission from Gacek RR. Neuroanatomical correlates of vestibular function. Ann Otol Rhinol Laryngol 1980; 89:1–5.

The projections of the major nuclei are responsible for the reflex connections of the duct and the macular afferents.[23–25] Second-order pathways activated by canal afferents project from the superior and medial vestibular nuclei and course in a parallel fashion to terminate in the trochlear and oculomotor nuclei with rostral terminations in the nucleus of Darkeschewitch and the interstitial nucleus of Cajal (see Figure 1–19). These pathways travel in the medial longitudinal fasciculus (MLF), with the ipsilateral pathway from the superior nucleus being inhibitory and the contralaterally projecting pathway from the medial nucleus being excitatory. They are responsible for the conjugate eye displacement manifesting as nystagmus from canal stimulation or destruction. Neurons from the medial vestibular nucleus project bilaterally to the abducens nuclei. An

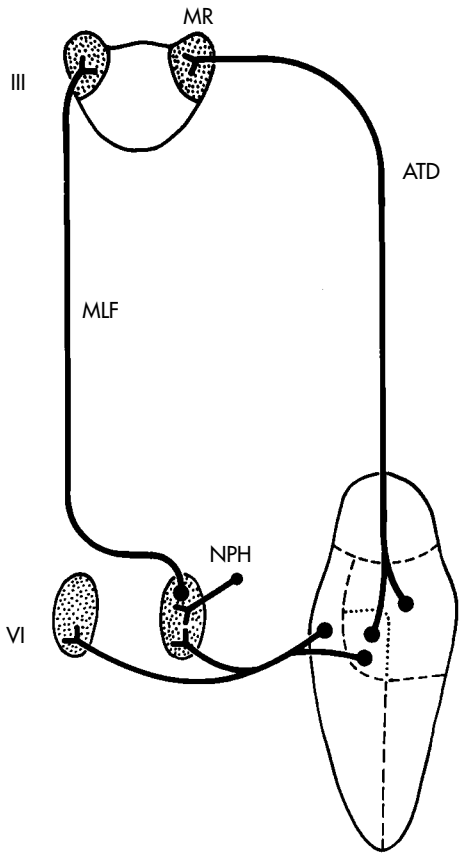


FIGURE 1–22. Central pathways for the horizontal vestibulo-ocular reflex. MR = medial rectus subnucleus; ATD = ascending tract of Deiters; MLF = medial longitudinal fasciculus; NPH = nucleus prepositus hypoglossi; III = oculomotor nucleus; VI = abducens nucleus.

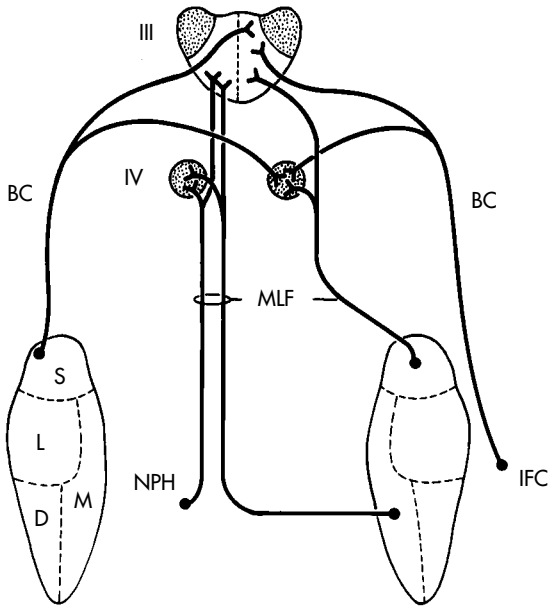


FIGURE 1–23. Central pathways for the vertical-rotatory vestibulo-ocular reflex. BC = brachium conjunctivum; IFC = infracerebellar nucleus; IV = trochlear nucleus; NPH = nucleus prepositus hypoglossi; MLF = medial longitudinal fasciculus; III = oculomotor nucleus; S = superior vestibular nucleus; L = lateral vestibular nucleus; D = descending vestibular nucleus; M = medial vestibular nucleus.

additional pathway lateral to the MLF, the ascending tract of Deiters, projects from neurons in the ventral portion of the lateral vestibular nucleus to the subnucleus of the oculomotor nucleus, which supplies the medial rectus muscle. The vestibulo-ocular reflex pathway for horizontal eye movement is completed with an interneuron located in the abducens nucleus, which carries an excitatory charge to neurons serving the contralateral medial rectus muscle (Figure 1–22). The oculomotor pathways for horizontal and vertical rotatory eye movements are separate, with the latter located in more lateral portions of the brainstem (Figure 1–23), whereas the pathways involved in the horizontal eye movement are located beneath the floor of the fourth ventricle (see Figure 1–23). It is this organization of pathways for horizontal and vertical rotatory eye movements that explains the perversion of nystagmus when the lateral canal is stimulated (calorically) in patients with midline cerebellar or brainstem lesions. That is, instead of seeing a horizontal nystagmus, a vertical nystagmus is seen because the afferent pathways to the abducens nuclei are interrupted, whereas those supplying the remaining extraocular muscles are intact.

The projections from the caudal vestibular nuclei, that is, the medial, descending, and entire lateral vestibular nuclei (both ventral and dorsal divisions), travel down the spinal cord to muscles of the trunk and limbs.[25,26] These projections join the vestibulospinal tract (VST) and make up the major output reflex pathway from the maculae, particularly the utricular macula (see Figures 1–20 and 1–21). The major pathway is the lateral vestibulospinal tract (LVST), which arises from the large multipolar neurons in the lateral vestibular nucleus and to a lesser extent from the descending nucleus. This pathway is somatotopically organized so that the vestibulospinal projections to the cervical and upper thoracic regions arise from neurons in the anteroventral portion of the lateral vestibular nucleus, whereas the most caudal and sacral portions of the spinal cord are innervated by the multipolar neurons in the most dorsal

and caudal portions of the lateral vestibular nucleus. Truncal and limb musculature of the intervening segments are organized in orderly fashion between these extremes. This pathway provides excitatory tone to extensor muscles of the limbs. The medial vestibulospinal tracts (MVSTs) travel down the MLF and are largely inhibitory, although some facilitatory connections are made from the descending vestibular nucleus.[21] The projections through the MLF arise from the caudal portions of the medial and descending vestibular nuclei. Since they are largely inhibitory, they act synergistically with the innervation of neck and upper truncal muscles excited by the LVST. The MVSTs extend down to the upper cord levels but do not reach the thoracic, lumbar, and sacral region levels reached by the LVST.

Other pathways exist in the vestibular nuclei that are of importance. One such pathway is the commissural pathway that interconnects the vestibular nuclei (see Figure 1–19). The major intervestibular or commissural projections are between the superior, medial, and descending vestibular nuclei.[21] A minor projection may exist between the lateral vestibular nuclei. As mentioned earlier, the group Y nucleus also forms a significant commissural projection to the contralateral group Y nucleus as well as the superior nucleus. The commissural pathways are largely inhibitory on second-order neurons activated by canal input. It is possible that they serve to potentiate the differential response arising from stimulation of coplanar canals. The commissural projections are important in the recovery of balance following ablation of one set of vestibular sense organs (labyrinthectomy). It has been demonstrated that the commissural pathways are largely responsible for providing the reactivation of input to the denervated side of the brainstem to approximate that in the intact half.[27]

The interaction of vestibular pathways with the cerebellum is emphasized by reference to certain areas of the cerebellum as the *vestibulocerebellum* (VC). These are the nodulus, uvula, flocculus, and ventral paraflocculus (see Figure 1–19). Other areas of the anterior and posterior lobes of the vermis are also included to a lesser extent. This association is based on the projection of the first- and second-order vestibular afferents to the cerebellar cortex by way of mossy fibers.[21] All canal afferents continue through the superior nucleus after termination there, to converge and terminate on neurons in the VC. Second-order neurons in the superior, medial, lateral, and descending nuclei relay input from the utricular and saccular maculae to the same areas. There is also some relay of canal input to the vestibular nuclei on to the cerebellum. The importance of such direct as well as relayed labyrinthine input to the cerebellum is not known but may have some bearing on different information coming in for modification of cerebellar output. For example, the relayed canal input may be modified by commissural inhibition before entering the cerebellum. Furthermore, other afferent input (spinal) may alter the labyrinthine (macular) input before relay to the cerebellum.

Input from the labyrinth and spinal cord, although important, is not the only source of inputting information to this important part of the vestibular system.[21] The visual system also has significant input to the flocculus and nodulus, where overlap with labyrinth input occurs. Therefore, interaction between vestibular and visual signals in the VC has firm anatomic and physiologic support. These various inputs to the VC probably determine the pattern of output from these regions, which will then impact on neurons in the vestibular nuclei.

In a reciprocal fashion, there are extensive projections to the vestibular nuclei from the cerebellar cortex and nuclei.[21] The anterior and posterior parts of the vermis provide extensive projections to the lateral vestibular nucleus, particularly the dorsal half. The output of the VC is distributed throughout the vestibular nuclei, especially to those areas related to the cerebellar input. These cerebellar projections are entirely inhibitory since Purkinje cells are inhibitory neurons.

However, not all cerebellar output is inhibitory. The fastigial nucleus (FN), which receives labyrinthine input, exerts an excitatory effect on the vestibular nuclei. The caudal part of the FN projects to parts of all of the contralateral vestibular nuclei, whereas the rostral region projects to the ipsilateral nuclei. The FN can be a link between the cerebellar cortex and the vestibular nuclei that provides an excitatory input on vestibular neurons.

Efferent Vestibular Pathway

An efferent vestibular pathway has also been demonstrated in lower and higher mammalian forms.[14,28,29] This pathway originates bilaterally from small neurons located lateral to the abducens nucleus near the

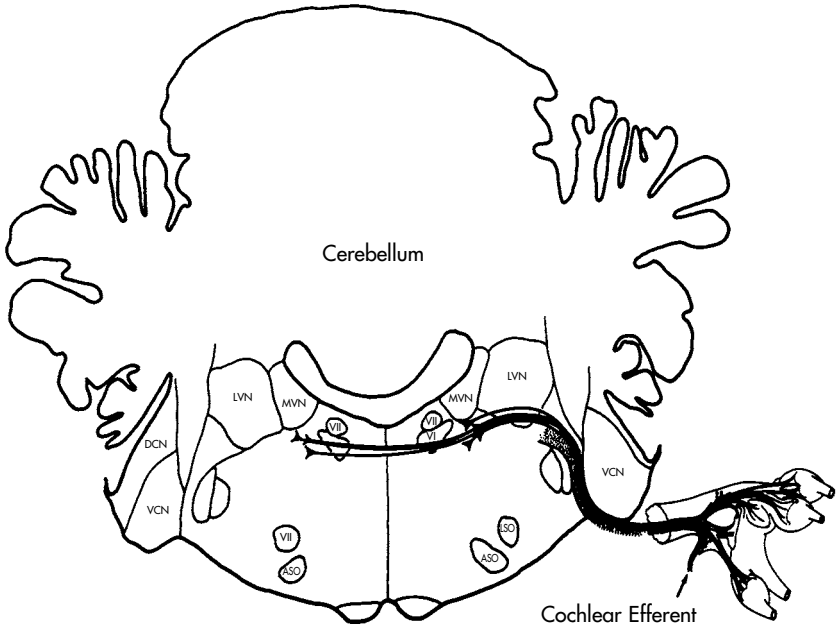


FIGURE 1–24. Diagram of the efferent vestibular pathway (*solid lines*). *Stippled area* = efferent cochlear pathway. LVN = lateral vestibular nucleus; MVN = medial vestibular nucleus; DCN, VCN = dorsal and ventral divisions of the cochlear nucleus; VII = genu of facial nerve; VI = abducens nucleus; ASO = accessory superior olivary nucleus; LSO = lateral superior olivary nucleus. Reproduced with permission from Gacek RR and Lyon M.[30]

medial vestibular nucleus[30] and projects in a symmetric fashion to each vestibular labyrinth (Figure 1–24). The axons, which are fine axons, converge with the olivocochlear bundle as the vestibular root is reached in the brainstem. They travel with the olivocochlear bundle efferents through the vestibular nerve to the point where the vestibular ganglia are reached. At this point, the vestibular efferents diverge from the parent efferent bundle to break up into fascicles and individual axons that then disperse throughout the ampullary and macular nerve branches. They branch and ramify richly along their course within the nerve trunks and after penetration of the basement membrane in the neuroepithelium of the vestibular sense organs. Here they represent fine fibers that form many bouton-like vesiculated endings in termination on type II hair cells and the calyces encapsulating type I hair cells (see Figure 1–16). These fibers are cholinergic, as are the olivocochlear efferents, and can be demonstrated selectively with acetylcholinesterase localization techniques.[15] The number of efferent neurons is relatively small compared with the number of afferent neurons in the vestibular nerve. In the cat, for the 8,000 afferent neurons, there are 200 to 300 efferent neurons that supply through their branching an almost equal number of efferent terminals to the afferent terminals in the sense organs. Although the number of vestibular efferents in the human labyrinth has not been determined, it is presumed that the imbalance in number is similar to that in lower forms. The function of the efferent system to the vestibular sense organs is not known. Most experimental studies indicate an inhibitory effect on afferent vestibular activity.[31] However, an excitatory effect has also been demonstrated in primates.[29] It is possible that the efferent effect is to raise or lower the resting activity level in vestibular afferents, thereby modifying their operating range. Unlike the efferent auditory pathway, the vestibular efferent pathway is represented by small neurons located near the medial vestibular nucleus, where they can be contacted by first-order afferents. Thus, a two-neuron reflex arc exists to type I hair cell afferent terminals and type II hair cell soma.

Fluids of the Inner Ear

As mentioned earlier, there are two fluid compartments of the inner ear, the endolymph and the perilymph. Endolymph is a fluid that has a similar ionic composition to intracellular fluid and fills the membranous auditory and vestibular labyrinth. Endolymph is formed by secretory cells in the stria vascularis and by dark cells near the ampullary ends of the semicircular ducts and the walls of the utricle. The endolymph is thought to be absorbed in the endolymphatic sac. Endolymph composition is characterized by a high potassium level and a low sodium level.[32]

TABLE 1–1. Chemical Composition of Inner Ear Fluids

Mean Value	*Perilymph*	*Endolymph*	*Cap. Serum*	*CSF*
NA^+ (mEq/L)	143	12–16	141	141
K^+ (mEq/L)	5.5–6.25	143.3 (140–160)	5.9	2.9
Protein (mg %)	200 (89–326)	150?	7170	30
Glucose (mg %)	104		104	67
Free cholesterol (mg %)	1.5			0.035
Total cholesterol (mg %)	12		28	
MDH (IU)	95.6–136		63.5	18.3
LDH (IU)	127–155		151	1.9
PO_3^- (mM/L)	0.72		0.95	0.36
CA^{++} (mM/L)	1.16	1.07	2.44	1.12
Lactate (mM/L)	6.78		4.63	3.94

CSF = cerebrospinal fluid; MDH = malate dehydrogenase; LDH = lactate dehydrogenase. Adapted from Smith CA et al.[32]

The membranous labyrinth is suspended within the bony labyrinth by a fine trabecular network in a space that is filled with perilymph. Perilymph, in contradistinction to endolymph, has an ionic composition similar to extracellular fluid, with low potassium and high sodium levels.[32] A comparison of the composition of the two fluids with that of serum and cerebrospinal fluid is found in Table 1–1. This differential chemical makeup in the two fluid compartments is important to the establishment of a standing voltage surrounding the sense organs, which permits the generation of nerve impulse in the hair cell afferent neuron unit. Perilymph is secreted largely as the result of diffusion from the capillary network in the spiral ligament adjacent to the scala tympani with a smaller portion derived from cerebrospinal fluid by way of the cochlear aqueduct.

Blood Vessels of the Ear

Two major vessels have little or nothing to do with the vascular supply of the ear itself but are important in disorders and surgery of the temporal bone. These are the sigmoid sinus and the internal carotid artery. The sigmoid sinus is a continuation of the lateral (transverse) sinus, which is formed by the superior sagittal sinus. It makes an indentation into the posterior fossa plate of the mastoid portion of the temporal bone before taking an abrupt redundant turn on itself (jugular bulb) to exit through the jugular foramen accompanied by cranial nerves IX, X, and XI (see Figure 1–4). This structure is important as it may represent a lethal complication of bacterial mastoiditis (thrombophlebitis) when erosion of its bony covering occurs. The anatomy of the jugular bulb is particularly crucial to the diagnosis and surgical management of neoplasms, which originate in the bulb such as glomus jugulare tumors or schwannomas.

The internal carotid artery enters the temporal bone just anterior to the jugular foramen and travels in a bony canal anterior to the middle ear space, first in a vertical direction and then in a horizontal anteromedial direction medial to the eustachian tube (see Figure 1–4). Its main clinical significance is as an important landmark in surgery of the petrous apex and skull base. Rarely, an anomalous internal carotid artery may appear clinically to be a vascular neoplasm in the middle ear.

The circulatory networks of the external and middle ears and the inner ear are completely separate, with the one being supplied by the carotid system, whereas that of the labyrinth is derived from the vertebrobasilar system.[9] The external ear is supplied by the auriculotemporal branch of the superficial temporal artery and branches of the posterior auricular branches of the external carotid artery. The

middle ear and mastoid are supplied by a different set of arterial branches from the external carotid system. The arterial branches to the middle ear space are the anterior tympanic branch from the internal maxillary artery, which enters through the petrotympanic fissure and travels along the eustachian tube and the semicanal for the tensor tympani. The middle meningeal artery gives off the superior tympanic branch that enters the middle ear through the petrosquamous fissure. The middle meningeal artery also gives off the superificial petrosal artery that travels with the greater superficial petrosal nerve and enters the facial canal at the hiatus. This vessel anastomoses with a branch of the posterior auricular artery, the stylomastoid artery, which enters the facial canal inferiorly through the stylomastoid foramen. A branch of the stylomastoid artery leaves the fallopian canal to travel through the canaliculus with the chorda tympani nerve to enter the middle ear. Finally, the inferior tympanic artery, a branch of the ascending pharyngeal artery, enters the middle ear through the tympanic canaliculus in the hypotympanum with the tympanic branch of the ninth nerve.

The vascular supply to the ossicles is derived from the anterior tympanic artery, the posterior tympanic artery, and branches from the plexus of vessels on the promontory. The most tenuous link in the ossicular chain as regards blood supply is the tip of the long process of the incus, which commonly undergoes necrosis secondary to conditions that compromise its blood supply.[33]

Blood Supply to the Inner Ear The blood supply to the inner ear is derived from the labyrinthine branch of the anterior inferior cerebellar artery off the basilar artery.[9] Occasionally, the internal auditory artery has a direct origin from the basilar artery. This vessel represents an end artery as it receives no known anastomosing arterial vessels. The internal auditory artery as it enters the IAC divides into three branches. The first branch, the anterior vestibular artery, supplies the semicircular ducts, utricle, and saccule. The second branch, the vestibulocochlear artery, supplies the saccule, utricle, posterior duct, and basal turn of the cochlea. Its third and terminal branch is the cochlear artery, which enters the modialus, where it gives off the spiral vessels that form external and internal radiating arterioles. The internal radiating arterioles descend to supply the limbus as well as the basilar membrane and organ of Corti. The external radiating arteriole courses through the interscalar septum to supply the vascular arcades and capillaries of the stria vascularis. This arteriole anastomoses with the venous supply near the spiral prominence, with the venous return being along the floor of the scala tympani, where the veins merge with the collecting venules from the spiral vein to form the posterior spiral vein. This posterior spiral vein, or inferior cochlear vein, then travels along the scala tympani to exit at the cochlear aqueduct through a separate channel and enter the inferior petrosal sinus. The remainder of the labyrinth is drained by a venous system, which parallels to some degree the arterial system. Therefore, an anterior vestibular vein and the posterior vestibular vein drain the posterior duct ampulla as well as the saccule and a portion of the utricle and superior duct ampulla. These merge to join the vein at the cochlear aqueduct. The venous return from the semicircular ducts and the body of the utricle converge to form the vein at the vestibular aqueduct, which travels as the vein of the paravestibular aqueduct. This vein travels along with the endolymphatic duct toward the endolymphatic sac. This venous system drains into the sigmoid sinus.

FACIAL NERVE

The anatomy of the facial nerve and its branches are important in clinical practice, not only because of the cosmetic effects of facial paralysis, whether it be by disease or by surgical injury, but because the various components of facial nerve are important for the diagnosis and localization of lesions in the temporal bone. The facial nerve has sensory function and autonomic afferent and efferent function, together with the motor function for the facial musculature. The neuronal pathway from the motor cortex to the facial nucleus is either two or three neurons in a mono- or disynaptic pathway. The motor control of the facial muscles is located at the inferior end of the presylvian gyrus. The area for facial musculature innervation is located near the area of innervation of laryngeal muscles. Although a modest corticonuclear projection exists, it is not the only higher pathway inputting the facial nucleus. Various extrapyramidal neurons located in the red nucleus as well as the periaqueductal gray nucleus also project to the facial nucleus. The projection to the facial nucleus is largely contralateral, as are most corticobulbar projections, but

there is bilateral innervation to the facial neurons supplying the frontalis muscle.

The facial nerve consists of four major components: the motor component to the facial musculature, a sensory component to receptors in facial muscles as well as to the face, autonomic secretomotor and special sensory pathways. The motor component of the facial nerve is the largest component and arises from the facial nerve nucleus, which is located just caudal to the lateral superior olivary nucleus of the auditory system.[34] Immediately caudal to the facial nerve nucleus is the rostral limit of the nucleus ambiguus, which provides the motor innervation to the laryngeal musculature. It has been demonstrated in the cat with retrograde neuronal tracers that the regional facial muscle groups are supplied by groups of neurons within the facial nerve nucleus in a medial and lateral arrangement.[34] The number of facial motoneurons has been estimated at approximately 10,000 to 20,000. As axons leave the facial nucleus, they travel in a dorsal direction to loop around the abducens nucleus under the floor of the fourth ventricle and then curve in a ventrolateral direction passing between the lateral superior olivary nucleus and the descending trigeminal root to emerge from the brainstem ventral to the eighth nerve. The nerve is joined by the nervus intermedius and travels in the IAC in its anterior and superior compartment. In this portion of the IAC, it has an intimate anatomic relationship with the superior division of the vestibular nerve, and indeed the nervus intermedius may travel within the vestibular nerve for a part of its intracanalicular course. At the distal end of the IAC, the facial nerve diverges in an anterolateral direction to enter the labyrinthine segment of the fallopian canal (see Figure 1–5). At the geniculate ganglion, it makes an acute bend near the floor of the middle cranial fossa to continue within the bony canal in its tympanic segment. The course within the labyrinthine segment of the fallopian canal is associated with an extension of the subarachnoid space. This extension of the subarachnoid space is usually limited by the location of the geniculate ganglion but in a small portion of temporal bones may extend into the tympanic portion of the fallopian canal. The greater superficial petrosal nerve emerges from the geniculate ganglion at the facial hiatus (see Figure 1–9) and courses along the floor of the middle cranial fossa between layers of dura to reach the foramen lacerum, where it joins with the carotid sympathetic nerves to form the vidian nerve in the vidian canal. The tympanic portion of the facial nerve travels superior to the oval window in a posterolateral direction to the level of the horizontal or lateral semicircular canal, where it makes a 90-degree turn in a ventral direction to continue as the mastoid portion of the facial nerve (Figure 1–25). After traveling the length of the vertical canal, the facial nerve emerges from the stylomastoid foramen to enter the parotid gland anterior to the mastoid tip. Before emerging from the stylomastoid foramen, the chorda tympani nerve is given off through a bony canal in a retrograde direction toward the middle ear, where it passes between the neck of the malleus and long process of the incus to enter the epitympanum and leave the middle ear space before joining the lingual nerve. The facial nerve provides innervation to the posterior belly of the digastric muscle just external to the stylomastoid foramen and then emerges into the parotid gland, where it gives rise to four to five branches providing innervation to the facial musculature. Although the motoneurons in the facial nerve are spatially organized within the nucleus, the fibers to the peripheral facial muscle groups are interspersed throughout the course of the facial nerve but gather together again on exiting the temporal bone to provide the four to five distal branches supplying the regional facial musculature.[34]

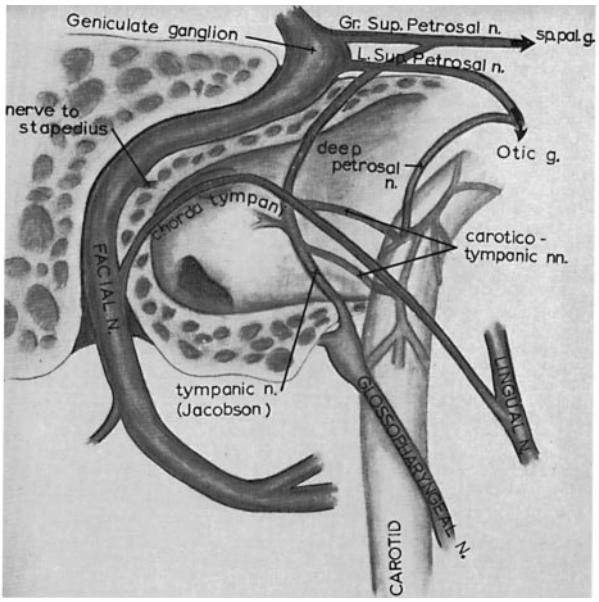


FIGURE 1–25. The facial nerve and its relationships (diagrammatic). Sp.pal.g. = sphenopalatine ganglion.

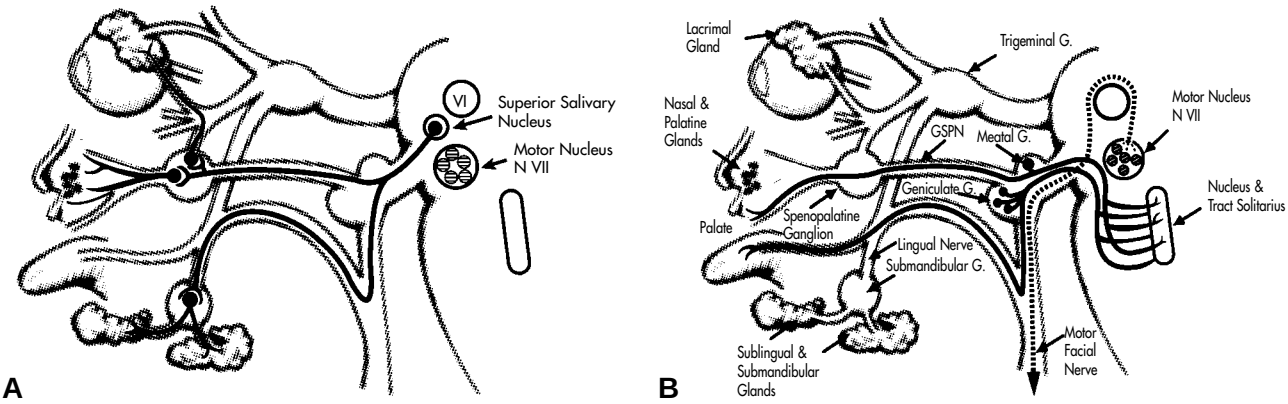


FIGURE 1–26. A, Diagram of the autonomic secretomotor pathways associated with the facial nerve. *B*, Diagram of the special sensory pathways in the facial nerve.

General Sensory Component

An unknown number of sensory afferent neurons are intermixed with the motor axons in the facial nerve trunk. In an animal model (cat), approximately 15 to 20% of the nerve fibers in the medium to large fiber size persist even after total facial nerve transection in the cerebellopontine angle. It is not clear what peripheral receptors these afferents innervate. The possibilities are either spindles within the facial musculature or some sensory receptors in the skin, particularly the dermis.

Autonomic Secretomotor Neurons

Secretomotor or motor preganglionic neurons of the facial nerve are located in the superior salivatory nucleus in the brainstem (Figure 1–26, A). They travel in the nervus intermedius, with some traveling within the vestibular nerve trunk. Those traveling within the vestibular nerve leave in the vestibulofacial anastomosis in the distal part of the IAC to join the nervous intermedius as it travels with the facial nerve trunk in the meatal canal.[9] Some efferent preganglionic neurons pass through the geniculate ganglion to the foramen lacerum and the vidian canal. After leaving the vidian canal, they synapse with postganglionic neurons in the sphenopalatine ganglion that innervate the lacrimal gland and secretory glands of the nose. Other preganglionic neurons travel within the sensory bundle of the facial nerve and leave with the chorda tympani nerve to join the lingual nerve, which carries them to the submandibular ganglion, where they synapse with postganglionic neurons, providing secretomotor function to the submandibular gland. The sensory bundle is located lateral to the motor component throughout the intratemporal course of the facial nerve.[9]

Special Sensory Function

Special sensory function for taste receptors in the anterior two-thirds of the tongue as well as the soft palate and nasopharyngeal mucosa are carried by unipolar ganglion cell masses in the facial nerve (Figure 1–26, B). Taste receptors in the anterior two-thirds of the tongue are innervated by dendrites of the geniculate ganglion and travel by way of the lingual and the facial nerve trunk, where they are contained in the sensory bundle of the facial nerve to ganglion cells in the geniculate ganglion.[35] Axons of these ganglion cells pass into the brainstem in the nervus intermedius to reach the tractus and nucleus solitarius, where they terminate. Taste receptors in the oral cavity, primarily the soft palate and the nasopharynx, are supplied by afferent dendrites, which travel in the vidian nerve and the greater superficial petrosal nerve. They pass through the facial hiatus to join the nervus intermedius, which brings them to their ganglion cells in the meatal ganglion of the facial nerve within the IAC.[35,36] Axons of these neurons also enter the brainstem in the nervus intermedius to terminate in the tractus and nucleus solitarius. The meatal ganglion located within the IAC has special significance in harboring neurotropic viruses (herpes virinae), which gain entry through nerve endings in the oral cavity.[36,37] After long latent periods, the viruses can be reactivated later in life to manifest as viral neuropathies (Bell's palsy, recurrent vertigo).

REFERENCES

1. Bast TH, Anson BJ. The temporal bone and the ear. Springfield (IL): Charles C. Thomas; 1949.
2. Anson BJ, Donaldson JA. Surgical anatomy of the temporal bone and ear. Philadelphia (PA): WB Saunders; 1981.
3. Lim DJ. Tympanic membrane. Electronmicroscopic observations. Part I. Pars tensa. Acta Otolaryngol (Stockh) 1968;66:181–98.
4. Lim DJ. Tympanic membrane. Part II: Pars flaccida. Acta Otolaryngol (Stockh) 1968;66:515–32.
5. Allam AF. Pneumatization of the temporal bone. Ann Otol Rhinol Laryngol 1969;78:49–64.
6. Gacek RR. Diagnosis and management of primary tumors of the petrous apex. Ann Otol Rhinol Laryngol 1975;84 Suppl 18:1–20.
7. Wersall J, Flock A, Lundquist RG. Structural basis for directional sensitivity in cochlear and vestibular sensory receptors. Cold Spring Harbor Symp Quant Biol 1965;30:115–32.
8. Wersall J. Studies on the structure and innervation of the sensory epithelium of the cristae ampullares in the guinea pig. A light and electronmicroscope investigation. Acta Otolaryngol (Stockh) 1956; Suppl 126:1–85.
9. Schuknecht HF. Pathology of the ear. 2nd ed. Philadelphia (PA): Lea and Febiger; 1993.
10. Spoendlin H, Schrott A. The spiral ganglion and the innervation of the human organ of Corti. Acta Otolaryngol (Stockh) 1990;105:403–10.
11. Chow KL. Numerical estimates of the auditory central nervous system of the Rhesus monkey. J Comp Neurol 1951;95:159–75.
12. Rasmussen GL. The olivary peduncle and other fiber projections of the superior olivary complex. J Comp Neurol 1946;84:141–220.
13. Rasmussen GL. Further observations of the efferent cochlear bundle. J Comp Neurol 1953;99:61–74.
14. Warr B, Guinan J, White JS. Organization of the efferent fibers: the lateral and medial olivocochlear systems. In: Altschuler RA, Hoffman GW, Bobbin RP, editors. Neurobiology of hearing: the cochlea. New York: Raven Press; 1986. p. 333–48.
15. Ishii D, Balogh K. Distribution of efferent nerve endings in the organ of Corti: their graphic reconstruction in cochleae by localization of acetylcholinesterase activity. Acta Otolaryngol (Stockh) 1968; 66:282–8.
16. Fex J. Efferent inhibition of the cochlea related to hair cell DC activity: study of post-synaptic activity of the crossed olivocochlear fibers in the cat. J Acoust Soc Am 1967;41:666–75.
17. Kemp DT. Stimulated acoustic emissions from the human auditory system. J Acoust Soc Am 1978;64: 1386–91.
18. Spoendlin H. Ultrastructural studies of the labyrinth in squirrel monkeys. In: Symposium on the role of the vestibular organs in the exploration of space. Washington (DC): National Aeronautics and Space Administration; 1965. p. 7–22. NASA SP-77.
19. Dohlman G. On the mechanism of transformation into nystagmus on stimulation of the semicircular canals. Acta Otolaryngol (Stockh) 1938;26:425–42.
20. Dohlman G. The shape and function of the cupula. J Laryngol Otol 1969;83:43–53.
21. Wilson VJ, Melville Jones G. Mammalian vestibular physiology. New York: Plenum Press; 1979.
22. Gacek RR. The course and central termination of first order neurons supplying vestibular end organs in the cat. Acta Otolaryngol (Stockh) 1969; Suppl 254:1–66.
23. Gacek RR. Anatomical demonstration of the vestibulo-ocular projections in the cat. Laryngoscope 1971;81:1559–95.
24. Highstein SM. The organization of the vestibulo-oculomotor and trochlear reflex pathways in the rabbit. Exp Brain Res 1973;17:285–300.
25. Brodal A, Pompieano O, Walberg F. The vestibular nuclei and their connections, anatomy and functional correlations. The Henderson Trust Lectures. Edinburgh, Scotland: Oliver and Boyd; 1962.
26. Pompieano O, Brodal A. The origin of vestibulospinal fibers in the cat. An experimental-anatomical study with comments on the descending medial longitudinal fasciculus. Arch Ital Biol 1957;95: 166–95.
27. Precht W, Shimazu H, Markham CH. A mechanism of central compensation of vestibular function following hemi-labyrinthectomy. J Neurophysiol 1966; 29:996–1010.
28. Gacek RR. Efferent component of the vestibular nerve. In: Rasmussen GL, Windle WF, editors. Neural mechanisms of the auditory and vestibular systems. Springfield (IL): Charles C. Thomas; 1960. p. 276–84.
29. Goldberg JM, Fernandez C. Efferent vestibular system in the squirrel monkey, abstract. Neurosci Abstr 1977;3:543.
30. Gacek RR, Lyon M. The localization of vestibular efferent neurons in the kitten with horseradish per-

oxidase. Acta Otolaryngol (Stockh) 1974;77: 92–101.
31. Klinke R, Schmidt CL. Efferent influence on the vestibular organ during active movements of the body. Pflugers Arch 1970;318:352–3.
32. Smith CA, Lowry OH, Wu ML. The electrolytes of the labyrinthine fluids. Laryngoscope 1954;64:141–53.
33. Alberti PW. The blood supply to the incudostapedial joint and the lenticular process. Laryngoscope 1963; 73:605–15.
34. Gacek RR, Radpour S. Fiber orientation of the facial nerve: an experimental study in the cat. Laryngoscope 1982;92:547–56.
35. Gacek RR. On the duality of the facial nerve ganglion. Laryngoscope 1998;108:1077–86.
36. Gacek RR. Pathology of facial and vestibular neuronitis. Am J Otolaryngol 1999;20:202–10.
37. Gacek RR, Gacek MR. Meatal ganglionitis: clinical pathologic correlation in idiopathic facial paralysis. Otorhinolaryngol Nova 1999;9:229–38.

CHAPTER 2

Development of the Ear

Doris K-W. Wu, PhD, Daniel I. Choo, MD

The development of the human ear begins around the fourth week of embryonic life. Although the three components of the ear (external, middle, and internal) develop concurrently, their origins are diverse.

EXTERNAL EAR

The auricle develops from a series of mesenchymal condensations referred to as the hillocks of His (auricular hillocks), derived from branchial arches 1 and 2. The groove between the two arches deepens to form the external auditory meatus. The six hillocks are apparent in the 6-week embryo but subsequently fuse to form the gross structure of the pinna by the end of the seventh week. The exact structures generated by each hillock have been proposed as in Figure 2–1. However, the precise contributions of individual hillocks to the auricle have not been demonstrated and remain controversial.

The innervation of the external ear reflects these developmental origins. Trigeminal and facial nerve branches (derived from first and second branchial arch contributions, respectively) provide sensory innervation to the external ear. Cutaneous branches of the cervical plexus also provide sensory innervation to the external ear.

Significantly, the pinna rests in a more ventral and transverse position during early embryogenesis and assumes the typical, more rostral, vertically oriented position with maturation. As a result, improper or arrested development of the ear commonly results in low-set, retroverted auricles, as seen in Treacher Collins syndrome and hemifacial microsomia.

Failure of differentiation or fusion of the auricular hillocks is thought to result in developmental malformations of the auricle ranging from anotia to the spectrum of microtias (Figure 2–2).

During embryogenesis, the external meatus begins as a thickening of ectodermal cells at the dorsal end of the first branchial groove (Figure 2–3).

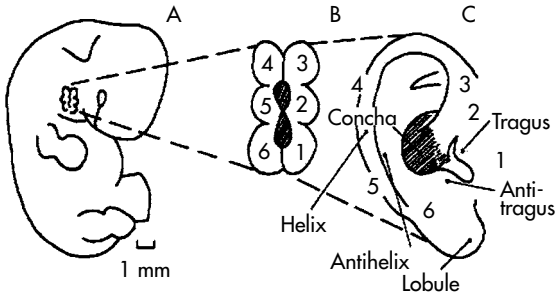


FIGURE 2–1. Development of the external ear. The external ear develops from six mesenchymal condensations known as auricular hillocks, *A* and *B*. Early in gestation, the hillocks rapidly fuse to form the primordial of the auricle. The parts of the auricle generated by the individual auricular hillocks are proposed as labeled in *C*. Courtesy of Drs. Ronan O'Rahilly and Fabiola Müller.

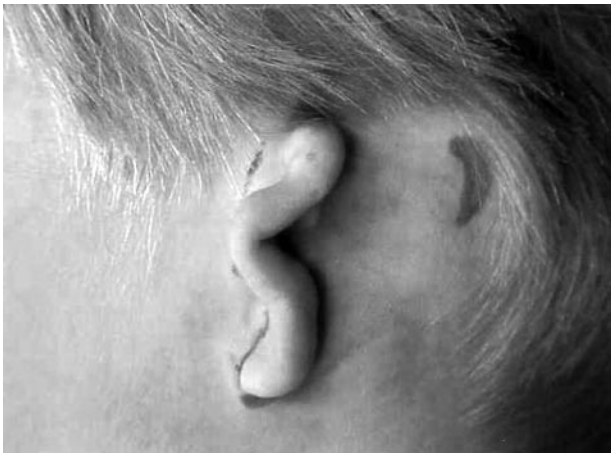


FIGURE 2–2. Grade III microtia showing a discrete earlobe with a cartilaginous and soft tissue rudiment. Grade III anomalies likely result from dysplastic or aplastic development of the auricular hillocks of the first two branchial arches at 5 to 6 weeks of gestation. Microtia is commonly graded from I (mildly dysplastic ears with all components of the auricle present) to III (severely dysplastic ears with most or all components absent).

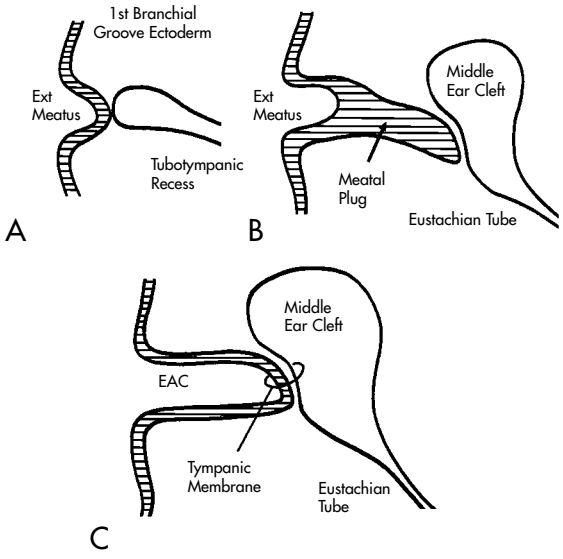


FIGURE 2–3. Drawing showing the development of the external auditory canal and middle ear. Development of the external meatus, ear canal, and middle ear space involves a complex process of ectodermal proliferation, cellular resorption, and coordinated maturation with the tubotympanic recess forming the tympanic membrane and middle ear cleft. The ectodermal cells at the dorsal end of the first branchial groove proliferate and expand medially (*A*, *B*). This solid core of ectodermal cells is referred to as the meatal plug. By the twenty-first week, this meatal plug begins to hollow out, leaving the ear canal lumen. Formation of the ear canal is typically complete by the twenty-eighth week of gestation (*C*). Note that the medial-most ectoderm and endoderm of the tubotympanic recess interface in such a way that these two layers combine to form the tympanic membrane (*C*).

Proliferation of these ectodermal cells generates a meatal "plug" that expands medially. Resorption of cells in the center of this meatal plug leaves a hollow tube-like structure representing the external meatus and ear canal. Failure of this canalization process results in aural atresia. The medial-most ectoderm forms the outer squamous cell epithelial layer of the tympanic membrane. It is noteworthy that external auditory canal development occurs subsequent to external ear morphogenesis. Accordingly, it is possible to see a developmental malformation resulting in isolated aural atresia without any auricular malformations, suggesting abnormal development around the twenty-eighth week. However, it is extremely unusual to see an isolated auricular malformation (indicating abnormal development around weeks 6 to 7) without concomitant middle ear and external auditory canal abnormalities. Therefore, in a child with microtia, imaging studies of the temporal bone are needed to determine the extent of associated middle ear abnormalities.

MIDDLE EAR

The middle ear cavity forms as a lateral extension of the first pharyngeal pouch known as the tubotympanic recess (see Figure 2–3). The proximal end of the extension remains as the eustachian tube. The lateral endoderm of the tubotympanic recess interfaces with the ectoderm of the meatal plug to form the tympanic membrane. The tympanic membrane is a three-layered structure consisting of a medial mucosal epithelium, a fibrous middle layer of mesenchymal origin, and an outermost layer of squamous cell epithelium contiguous with the skin lining the external auditory canal and originating from ectoderm. The malleus and incus appear to develop from a common ossicular mass around the sixth week of development. Both the malleus and the incus are thought to be derived from the neural crest in the first and second arch mesoderm, Meckel's cartilage and Reichert's cartilage, respectively (Figure 2–4). By the sixteenth week, the ossicles attain adult size, and ossification begins at discrete ossification centers (eg, medial surface of the neck of the malleus). Stapes development involves a complex morphogenesis starting as a blastema at 4½ weeks (derived from the neural crest in the second arch mesenchyme). The facial nerve divides the blastema into stapes, interhyale, and laterohyale. By the seventh week, a stapes "ring" forms around the stapedial artery. The interhyale becomes the stapedial muscle and tendon, whereas the laterohyale becomes the posterior wall of the middle ear and also a portion of the fallopian (facial nerve) canal. By the tenth week, the stapes assumes its more typical shape. Notably, the footplate seems to develop in conjunction with the otic capsule, with ossification beginning around the nineteenth week. The mastoid antrum expands late in fetal development and continues to expand significantly after birth.

The two middle ear muscles (the tensor tympani and stapedius muscles) are first and second arch derivatives, respectively. Accordingly, each mus-

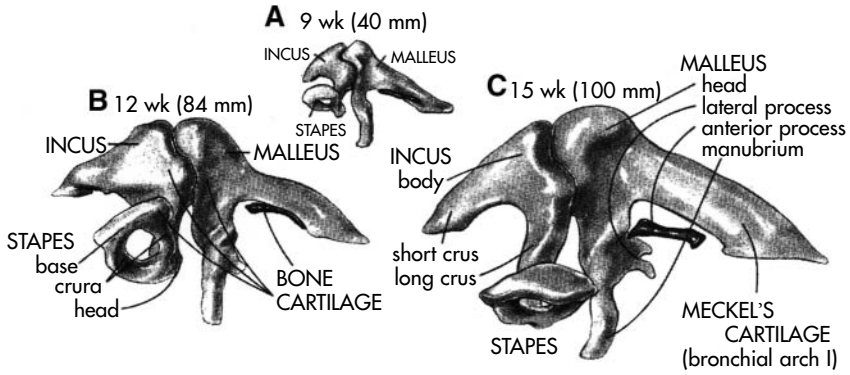


FIGURE 2–4. Ossicular development. *A*, By the ninth week of gestation, the incus and malleus are still fused but begin to differentiate into individual structures. The stapes at this stage has developed beyond the stapes ring stage and more closely resembles the mature stirrup configuration. *B* and *C*, By 12 to 15 weeks, the ossicles are more clearly differentiated and approximate adult size. Shortly following this stage, ossification begins at discrete centers of ossification. Reproduced with permission from Anson BJ, Davies J, Duckert LG. Embryology of the ear. In: Paparella MM, Shumrick DA, editors. Otolaryngology. 3rd ed. Philadelphia (PA): WB Saunders; 1991.

cle obtains its innervation from the respective nerve associated with each branchial arch: trigeminal innervation for the tensor tympani muscle and facial nerve innervation for the stapedius muscle.

INTERNAL EAR

The internal ear consists of two components, a membranous and a bony labyrinth. The membranous labyrinth is derived from the ectoderm, whereas the bony labyrinth/otic capsule is derived from the mesoderm and neural crest surrounding the primordial membranous labyrinth.

Development of the Membranous Labyrinth

The membranous portion of the inner ear orignates from a thickening of the ectoderm adjacent to the hindbrain known as the otic placode (Figure 2–5). In humans, the otic placode is evident at the third week of embryonic development. The otic placode soon invaginates from the surface ectoderm to form an otic cup. By the end of the fourth week, the edges of the otic cup come together and fuse to form the otic vesicle/otocyst. Some cells of the otic epithelium at the otic cup and otocyst stage delaminate from the epithelium and coalesce to form neurons of the statoacoustic (eighth cranial) ganglion just anterior and medial to the otocyst. These neurons subsequently send afferent fibers to innervate sensory organs within the inner ear. Concurrent with the development of the ganglion, the otocyst proper undergoes a complex morphogenesis, and the gross anatomy of the membranous labyrinth is nearly mature by the end of the eighth week of gestation (Figure 2–6).

The first morphologic change in the otocyst is the appearance of a diverticulum that first grows out dorsally and then moves medially and elongates to form the endolymphatic duct and sac. The rest of the otocyst enlarges and lengthens and can be roughly divided into a dorsal vestibular and a ventral saccular and cochlear region. The dorsal region develops into the utricle and three semicircular ducts and their respective ampullae. The ventral region develops into the saccule and the auditory component of the inner ear, the cochlear duct (see Figure 2–6, A).

The three semicircular ducts, superior, posterior, and lateral, develop sequentially during the fifth week, with the superior duct forming first, followed by the posterior and lateral ducts. The superior and posterior ducts form from a vertical outgrowth in the dorsal region of the otocyst. The lateral duct forms from a horizontal outgrowth in the lateral portion of the otocyst. In the central region of each presumptive duct, the opposing epithelia of each outgrowth fuse together and reabsorb, leaving behind a tube-shaped duct (see Figure 2–6, B to D; upper panel). One end of each duct opens into the utricle, and the other end forms a bulge-like struc-

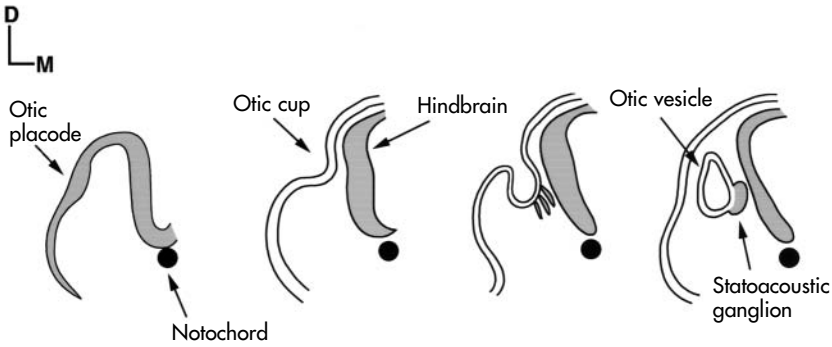


FIGURE 2–5. Diagram of coronal sections illustrating the development of the otocyst. Orientations: D = dorsal; M = medial. Reproduced with permission from Moore KL. The developing human. Philadelphia (PA): WB Saunders; 1977.

ture known as the ampulla that houses one of the vestibular sensory organs, the crista ampullaris. The superior and lateral ampullae form at the anterior end of their respective ducts, whereas the posterior ampulla forms at the posterior end of the posterior duct (see Figures 2–6, E, and 2–7).

The primordium of the cochlear duct evaginates from the ventral portion of the otocyst starting at the fifth week. This evagination extends anteromedially and gradually begins to coil. By the eighth week, the cochlea has formed 1½ turns (see Figure 2–6, E) and at the tenth week, two turns; at the twenty-fifth week, the cochlear duct has acquired the mature 2½ turns. The utricle and saccule start to form during the sixth week. As the cochlear duct extends, the opening between the saccule and cochlea becomes constricted and forms the narrow cochleosaccular duct. In the mature inner ear, two additional ducts are evident medially: the utricular and saccular ducts that connect the endolymphatic duct and sac to the utricle and saccule, respectively (see Figure 2–7).

Development of the Sensory Organs

Six major sensory organs are present in the human inner ear. Only the organ of Corti is responsible for auditory function. The remaining five sensory organs are vestibular in function and include three cristae ampullaris and the maculae of the utricle and saccule. In addition, a small vestibular organ, the macula neglecta, is thought to be present in only 1% of humans. The origin and lineage relationships among these sensory organs during development are not clear. Based on morphologic studies in the chicken, it

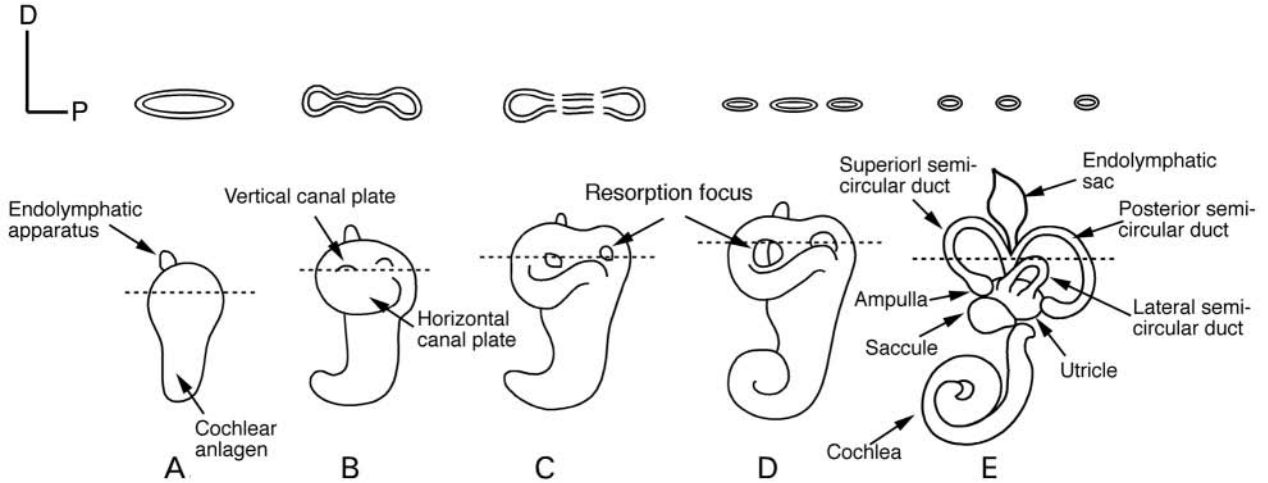


FIGURE 2–6. Diagram showing the development of the membranous labyrinth of the internal ear. *A* to *E*, Lateral views of the left inner ear from the fifth to eighth weeks. The *dotted lines* represent the level of section for the diagrams shown in the upper panel, illustrating the development of the superior and posterior semicircular ducts. Orientations: D = dorsal; P = posterior. Reproduced with permission from Moore KL. The developing human. Philadelphia (PA): WB Saunders; 1977.

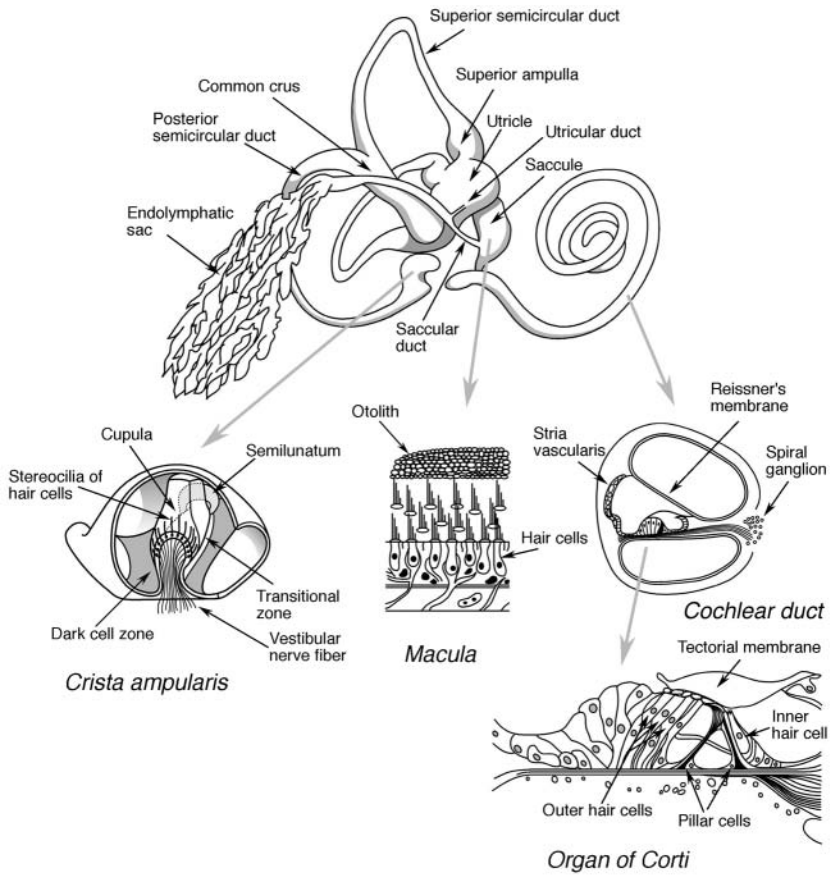


FIGURE 2–7. Diagram illustrating the anatomy of a mature membranous labyrinth and its sensory organs. Reproduced with permission from Schuknecht HF.[3]

has been proposed that all sensory organs in the inner ear originate in the ventromedial wall of the otic cup, and each sensory organ is derived from this common region as morphogenesis continues.[1] Thus far, there are no early fate-mapping studies in chicken or other animal models that lend direct support to this hypothesis. However, early gene expression studies of the inner ear at stages before the onset of histologic differentiation suggest that the sensory organs might be defined differently molecularly. For example, in mice, BMP-4 (bone morphogenetic protein 4) and *Msx1* (muscle segment homeobox gene) are expressed only in the three presumptive cristae and not other sensory organs.[2] Whether these molecular differences represent differences in the origin and/or mechanisms of specification for these sensory organs remains to be verified.

Once the position and type of a sensory organ have been specified, morphologic differentiation follows, and the sensory epithelium can be identified by an increase in the thickness of the otic epithelium into a pseudostratified layer that later differentiates into sensory hair cells and supporting cells. The Notch signaling pathway is used in a variety of tissues for generating cell type diversity and regulating differentiation during development.[4,5] Genes in the Notch signaling pathway have also been implicated in the determination of hair cells and supporting cells in the inner ear. Mutation or deletion of genes in this pathway resulted in the absence or aberrant number of sensory hair cells in zebra fish and mice.[6–8]

Cristae Ampullaris The sensory epithelium for the crista elevates into a ridge-like fold where hair cells and supporting cells develop. Cells surrounding the sensory epithelium are secretory in function and are thought to generate the gelatinous cupula in which the stereocilia of the sensory hair cells are embedded. The mound-like elevation of the crista is evident by the eighth week, and the sensory structure is mature by the twenty-third week (see Figure 2–7).

Maculae The maculae develop in a similar fashion as the cristae except the sensory epithelium is flat, and it is covered by an otolithic membrane that con-

tains superficial calcareous deposits, the otoconia. The maculae appear to be fully differentiated between 14 and 16 weeks (see Figure 2–7).

Organ of Corti The organ of Corti develops from the posterior wall of the cochlear duct. As the cochlear duct increases in length, the cross-sectional shape of the duct changes from round to oval to triangular (see Figures 2–7 and 2–8). The posterior wall develops into the sensory tissue, the organ of Corti; the anterior wall forms part of Reissner's membrane; and the lateral wall forms the stria vascularis. The organ of Corti starts to differentiate at the basal region of the cochlear duct around the seventh week, whereas the epithelium in the apical turn is undifferentiated and pseudostratified at this age. By the twenty-fifth week, the organ of Corti is fully differentiated.

Development of the Statoacoustic Ganglion

Based on studies in animal models, neurons of the statoacoustic ganglion are thought to originate from the otic epithelium, whereas glial cells in the ganglion are derived from the neural crest.[9] Morphologic studies suggest that cells in the anteroventral lateral region of the otic cup/otocyst delaminate from the epithelium, migrate away from the otocyst, and undergo proliferation before aggregating to form the ganglion.[10] In parallel with the growth of the otocyst, the statoacoustic ganglion forms a pars superior and a pars inferior at about the time when the membranous labyrinth is divided into a dorsal vestibular and a ventral cochlear region. The pars superior of the ganglion provides the peripheral neural connections to the superior and lateral cristae and the macula of the utricle. The pars inferior gives rise to a distinct vestibular portion that innervates the macula of the saccule and posterior crista and the spiral ganglion that innervates the organ of Corti.

Helix loop helix transcription factors, such as Neurogenin1, have been implicated in the formation of the statoacoustic ganglion. In mice with a deletion in *Neurogenin1*, the statoacoustic ganglion fails to form.[11] Growth factors such as fibroblast growth factor 3 are also important for the formation of the vestibular ganglion in the mouse.[12] Once the neurons of the ganglion reach their final position and finish proliferation, they express high-affinity neurotrophin receptors and become dependent on neurotrophins secreted by the presumptive sensory organs in the membranous labyrinth. Mice with targeted deletion (knockout) of genes encoding brain-derived nerve growth factor or neurotrophin 3 or their respective receptors had a loss of ganglionic neurons and innervations to the sensory organs.[13] However, this trophic dependency of ganglionic neurons on their target tissues is not reciprocal. In the absence of afferent innervation, the differentiation of sensory hair cells appears unaffected, at least until birth.[14,15]

Development of the Bony Labyrinth

Concurrent with the development of the otocyst, the mesenchymal cells surrounding the otic vesicle differentiate into a cartilaginous otic capsule by the eighth week. As the membranous labyrinth enlarges, vacuoles appear in areas surrounding the otic epithelium as a result of programmed cell death. These vacuoles soon coalesce to form the perilymphatic space that is filled with perilymph. In the cochlea, the perilymphatic space develops in two divisions; the scala tympani forms before the scala vestibuli (see Figure 2–8).

Ossification of the cartilaginous capsule to form the bony labyrinth does not occur until the membra-

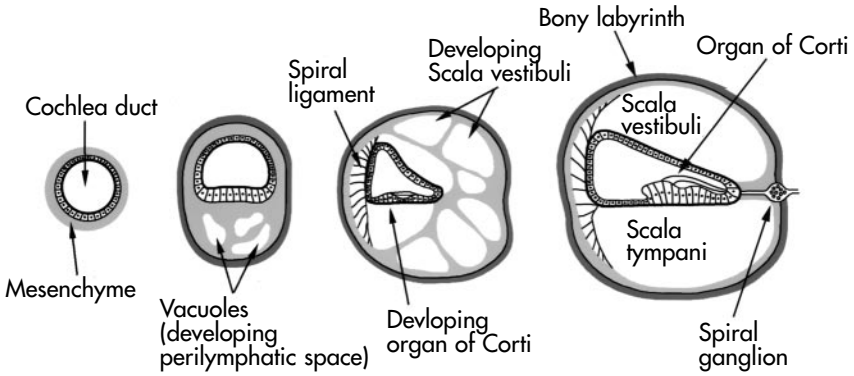


FIGURE 2–8. Diagram showing development of the bony labyrinth. Cross-sections through the cochlea showing the development of the organ of Corti, bony labyrinth, and perilymphatic spaces from the eighth to the twelfth weeks. Reproduced with permission from Moore KL. The developing human. Philadelphia (PA): WB Saunders; 1977.

nous labyrinth has acquired its adult size. Bone formation starts at the fifteenth week and ends by the twenty-first week, with a total of 14 ossification centers. The first ossification center appears at the base of the cochlea, and ossification advances more rapidly in the cochlea than in the canal region, where membranous growth continues to about the twenty-first week. By the twenty-third week, all ossification centers have fused with no "suture lines," thus limiting further growth in the membranous or bony labyrinth.

HUMAN EAR MALFORMATIONS AND MODELS OF MOLECULAR DEVELOPMENT

Many genes causing human deafness have been identified recently. In some cases, the functional deficit is the result of gross malformations of the external, middle, or inner ear or a combination of more than one of these components (Table 2–1). Examples of some of these syndromes are described below. For syndromes that are associated with the lack of proper differentiation and functioning of sensory hair cells, see Chapter 13 on hereditary hearing impairment, as well as a review by Steel and Bussoli.[16]

Branchio-Oto-Renal and Branchio-Oto Syndromes

Branchio-oto-renal (BOR) syndrome is an autosomal dominant disorder characterized by branchial arch anomalies (branchial cleft cysts and sinuses), ear malformations (affecting the external, middle, and inner ear regions), and various renal anomalies ranging from undetected in the branchio-oto (BO) syndrome to bilateral aplasia.[17] Hearing loss is the

TABLE 2–1. Genes Involved in the Development of the Human Ear

Syndrome/Disease	*Gene*	*Type of Product*	*Structures Affected*	*Reference*
Branchio-oto, branchio-oto-renal	*EYA1*	Transcription coactivator	Atresia and stenosis of external auditory canal, malformed auricle Hypoplasia or absence of three ossicles Absent or abnormal semicircular canals and cochlea	16, 20
Crouzon's disease	*FGFR2* *FGFR3*	Growth factor receptor	Atresia of external auditory canal Malformation of ossicles	21, 22 23
Craniosynostosis	*FGFR3*	Growth factor receptor	Sensorineural deafness	24
DFN3/Gusher	*POU3F4/ BRN4*	Pou-domain transcription factor	Conductive hearing loss Abnormally wide connection between internal acoustic canal and the inner ear Stapes fixation	25
Pendred's DFNB4	*PENDRIN*	Anion transporter	Sensorineural deafness Widened vestibular aqueduct Shortened cochlea	26, 27
Proximal symphalangism (SYM1), multiple synostoses (SYNS1)	*NOGGIN*	Secreted factor, antagonist of bone morphogenetic proteins	Fixation of stapes to petrous part of temporal bone	28
Townes-Brocks	*SALL1*	C2H2 zinc-fingered transcription factor	Malformed auricle, sensorineural hearing loss	29, 30
Treacher Collins (mandibulofacial dyostosis)	*TCOF1*	Nucleolar trafficking protein	Atresia of external auditory canal, malformed auricle Middle and inner ear anomalies	31

most common presenting symptom and occurs in 88% of patients with BOR syndrome.[18,19]

The most common external ear anomalies are preauricular pits and skins tags, but microtia and aural atresia are also noted in patients with BOR syndrome. The middle ear cleft is commonly hypoplastic in BOR syndrome, as can be the middle ear ossicles. In the inner ear, aplasia or hypoplasia of the cochlea has been reported as a BOR phenotype (Figure 2–9). The vestibular portions of the inner ear similarly show absent or malformed semicircular canals. Depending on the ear phenotype, hearing loss can be sensorineural, conductive, or mixed, as seen in approximately one half of the patients with BOR syndrome.[17]

The gene responsible for causing BOR and BO syndromes has been identified to be the human homologue of the *Drosophila eyes absent* (*Eya*) gene.[16,19] The spectrum of malformations seen in BOR syndrome suggests a defect in the *EYA1* molecular pathway occurring sometime between the fourth and the tenth week of gestation. The types of DNA mutations identified in patients with BOR syndrome suggest that the developmental malformations result from reduced gene dosage.[19]

Mice with a targeted deletion of *Eya1* also displayed similar severe external, middle, and inner ear *defects, suggesting that the function(s)* of this protein in ear development is evolutionary conserved.[32] In mice, *Eya1* is specifically expressed in the early otocyst along the ventromedial wall and the adjacent statoacoustic ganglion.[33] At later embryonic time points, *Eya1* is expressed along the neuroepithelium that eventually gives rise to key cochlear structures, including the inner and outer hair cells. In addition, *Eya1* is also expressed in the external and middle ear primordia. These results suggest that the encoded protein of *Eya1* plays a direct role in the formation of all three components of the ear in mice as well as in humans.

Studies in *Drosophila* eye imaginal disks show that the *Eya* gene product is a transcriptional coactivator that acts as part of a complex with other transcription factors during eye morphogenesis.[34] Its mode of action in ear development remains to be established.

Pendred's Syndrome

Pendred's syndrome is an autosomal recessive disorder classically characterized by deafness and goiter. The presence of goiter is usually variable in severity; however, most patients have a positive perchlorate discharge test indicative of a defective organification of iodine in the thyroid gland. The onset of deafness is typically congenital, profound, and sensorineural in nature. Cochlear malformation with missing apical turns, known as Mondini's phenotype, and widened vestibular aqueducts are often described in these patients.[35,36]

The gene responsible for causing Pendred's syndrome was identified using a positional cloning approach and was named *PENDRIN*.[25] Functional studies in *Xenopus* oocytes show that the encoded protein functions as a chloride/iodide transporter.[37] The clinical significance of this gene broadened when mutations in *PENDRIN* (*PDS*) were also shown to cause autosomal, recessive deafness locus 4 (DFNB4), a recessive nonsyndromic form of deafness.[27] Expression and gene-targeting studies in mice have provided insights into the etiology of this disorder. In mice, *Pds* is activated during embryonic development. Its expression is restricted to nonsensory regions of the inner ear including the endolymphatic duct and sac, nonsensory regions of the utricle and saccule, and the external sulcus region in the cochlea.[38] *Pendrin* knockout mice are deaf and display variable vestibular defects.[39] However, the cause of inner ear dysfunction in these mice most likely is not attributable to structural malformations as implied by the human disorder. The membranous labyrinth of the *Pds* knockout mice swells during late embryonic development indicative of an ionic imbalance presumably owing to a lack of the transporter function of *Pds*. The sensory hair cells develop and appear normal at birth but degenerate later in life, most likely as a consequence of disrupted fluid homeostasis. It is likely that the etiology of human Pendred's syndrome is similar to that of the *Pds* knockout mouse model, where the deficit lies in ionic imbalance of the endolymph rather than structural malformations during development. The widened vestibular aqueduct associated with Pendred's syndrome supports this hypothesis. Mutant mice containing mutations similar to those found in patients with Pendred's syndrome will help to clarify this issue.

Symphalangism and Multiple Synostoses Syndromes

Symphalangism syndrome (SYM1), an autosomal dominant disorder with aberrant bone fusion at

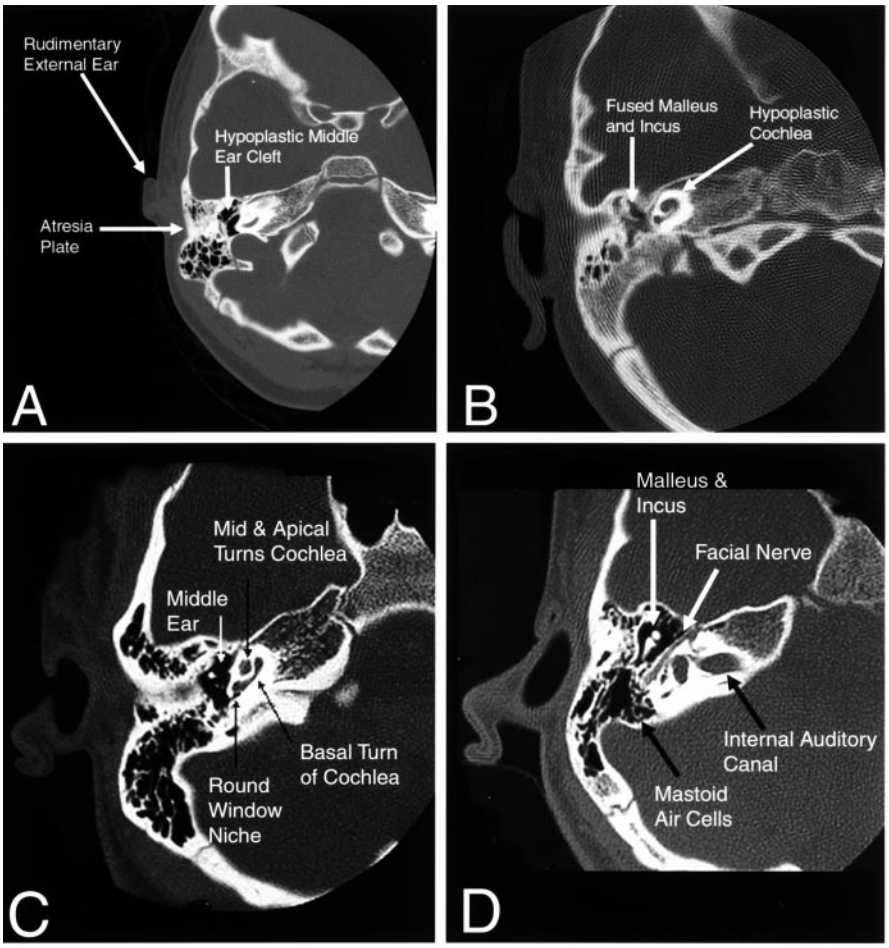


FIGURE 2–9. The external, middle, and inner ears are all potentially affected in branchio-oto-renal (BOR) syndrome. At the external ear level, auricular malformations vary from mild hypoplasia to complete aplasia. *A*, A rudimentary ear lobe is apparent even on this axial computed tomographic (CT) scan of the temporal bones in a child with BOR syndrome. Note also the failure of development of an external auditory canal (aural atresia) and the associated small middle ear cavity. *B*, Another patient with BOR syndrome demonstrates hypoplastic development of the cochlea with a discrete basal turn but poorly differentiated middle and apical turns appearing as a bulbous mass on high-resolution CT imaging. *C* and *D* demonstrate normative CT imaging data with middle and inner ear structures labeled as shown.

interphalangeal joints, is associated with conductive hearing loss. The ear anomaly is attributable to the fixation of the stapes to the petrous part of the temporal bone. Multiple synostoses syndrome (SYNS1) is also associated with joint fusion and conductive hearing loss. Both of these disorders are caused by mutations in *NOGGIN*, which encodes a secreted polypeptide that inactivates members of the transforming growth factor β family, in particular bone morphogenetic proteins (BMP-2 and -4).[28,34,40] *Noggin* knockout mice die at birth with multiple defects, including abnormal joint formation.[41,42] During embryogenesis, *Noggin* is usually expressed in regions closely associated with BMP-producing regions. *NOGGIN* is thought to function by modulating the available levels of BMPs. The abnormal bone fusion at interphalangeal joints and the stapes associated with symphalangism in humans is most likely caused by an alteration of BMP levels in those regions during development.

X-linked, Deafness Locus 3

Another conductive hearing loss disorder with fixation of the stapes is the X-linked DFN3, which is caused by mutations in a member of the POU domain transcription factors, *POU3F4/BRN4*.[25] In mice, *Pou3f4* is important for the development of the ear. Transcription of this gene is activated in mice at the otocyst stage in the mesenchymal cells surrounding the otocyst.[43] *Pou3f4* knockout mice show a variety of deficits, including abnormal fibrocytes surrounding the cochlea, misshapen stapes, and shortened cochlea.[44,45] The deficits caused by mutations of this gene in humans are usually conductive in nature and can be corrected by surgery.

Treacher Collins Syndrome

Treacher Collins syndrome is an autosomal dominant disorder of craniofacial development associated

with bilateral conductive hearing loss, and occasionally malformations of the inner ear have also been reported.[3] This syndrome is caused by mutations in the *TCOF1* gene that encodes for a nucleolar protein.[31,46] *TCOF1* is widely expressed in fetal and adult tissues and is thought to function as a trafficking protein between the nucleolus and cytoplasm.

Crouzon's Disease and Craniosynostosis

Crouzon's disease is characterized by craniofacial abnormalities and hearing loss. The hearing loss is usually conductive owing to anomalies of the middle ear and atresia of the external auditory canal. Mutations in both fibroblast growth factor receptor (*FGFR*) 2 and 3 have been shown to cause Crouzon's disease.[21–23] A unique mutation in *FGFR3* has also been shown to cause craniosynostosis that is associated with deafness and abnormalities on radiographs of hands and feet.[24] In mice, knockout of *Fgfr3* affected the development of the cochlea in which the pillar cells were missing and the tunnel of Corti failed to form.[47]

Townes-Brocks Syndrome

Townes-Brocks syndrome is an autosomal dominant syndrome associated with anal, renal, limb, and ear anomalies. External ear malformations and sensorineural hearing loss are frequently found in this syndrome. Malformations of the malleus and incus have also been reported.[29] This syndrome is caused by mutations in the *SALL1* gene, a putative C2H2 zinc-fingered transcription factor gene.[30]

RETINOIDS AND EAR DEVELOPMENT

Since the early 1950s, it was evident to clinicians that retinoids (vitamin A, retinoic acid [RA], etc) had potent teratogenic effects when administered early in pregnancy.[48] Either excess or inadequate maternal retinoid intake could result in a spectrum of congenital anomalies affecting the central nervous system, eye, and heart or result in craniofacial defects such as midface hypoplasia and cleft lip and palate as well as ear and limb malformations. Retinoic acid is one of the better-studied retinoids whose role in molecular development of the ear has been extensively studied in animal models. Several lines of evidence suggest an essential role for RA in ear development. First, reporter gene assays have demonstrated the endogenous production of RA by the organ of Corti.[49] In addition, several of the key RA receptors, both at the cytoplasmic and nuclear levels, have been demonstrated in the developing ear region.[50] Gain of function studies have revealed in both chicken and mouse that exogenous RA, administered at low levels during early embryogenesis, induces a spectrum of inner ear malformations that are similar to the variety of ear malformations observed with human retinoid embryopathy.[51,52] Intriguingly, delivery of exogenous RA to cultures of the organ of Corti has also been shown to cause development of supernumerary hair cells.[49] Most recently, loss of function studies have probed the role of individual and multiple RA receptors in ear development.[53–55] Generation of null mutant mice carrying targeted deletions of nuclear retinoic acid receptor and retinoid X receptor isoforms and cytoplasmic retinoic acid binding proteins isoforms has demonstrated the complexity of the molecular signaling system existing for RA during development.[56–58] In several examples, loss of individual RA receptor isoforms fails to yield a significant or informative phenotype, possibly owing to redundancy in the molecular pathways.[59] Accordingly, investigators have been forced to recognize the importance of normal RA signaling during development and the compensatory mechanisms in place so that perturbation of normal signaling does not occur, despite interruption of any single signal transduction pathway. Also interesting to note is the interaction of RA receptors with other members of the steroid hormone superfamily of nuclear receptors. For example, data suggest that thyroid hormone nuclear receptors may heterodimerize with activated RA receptors and regulate transcription.[60,61] Such mechanisms would provide a plausible explanation for the widespread developmental effects of retinoids that seemingly affect various organ systems.

Specific to the developing inner ear, RA has been shown to have significant effects on cell growth and differentiation as early as the otocyst stage. Using in vitro[62] and in vivo[51] animal models of inner ear development, exogenous RA has been shown to down-regulate cellular proliferation in the otocyst in association with altered levels of *c-fos* expression. Furthermore, RA has been shown to down-regulate the expression of critical genes such as *Bmp4* that in the chicken and mouse appear to be involved in sensory organ generation. However, a

key consideration in trying to dissect the molecular effects of RA rests in the likelihood that RA impacts multiple signaling pathways; therefore, isolating individual effects and their sequelae will be extremely difficult. However, the extensive work with RA receptor isoform-specific knockouts combined with work on the various (endogenous and synthetic) retinoid metabolites may provide the means with which to start dissecting this complex system and thereby determine the role of RA in ear development.

REFERENCES

1. Knowlton VY. Correlation of the development of membranous and bony labyrinths, acoustics ganglia, nerves, and brain centers of the chick embryos. J Morphol 1967;121:179–208.
2. Morsli H, Choo D, Ryan A, et al. Development of the mouse inner ear and origin of its sensory organs. J Neurosci 1998;18:3327–35.
3. Schuknecht HF. Developmental defects. In: Pathology of the ear. 2nd ed. Philadelphia (PA): Lea & Febiger; 1993. p. 115–89.
4. Artavanis-Tsakonas S, Rand M, Lake R. Notch signaling: cell fate control and signal integration in development. Science 1999;284:770–6.
5. Bray S. Notch signaling in Drosophila: three ways to use a pathway. Semin Cell Dev Biol 1998;9:591–7.
6. Haddon C, Jiang Y, Smithers L, Lewis J. Delta-notch signalling and the patterning of sensory cell differentiation in the zebrafish ear: evidence from the mind bomb mutant. Development 1998;125: 4637–44.
7. Lanford PJ, Lan Y, Jiang R, et al. Notch signalling pathway mediates hair cell development in mammalian cochlea. Nat Genet 1999;21:289–92.
8. Bermingham NA, Hassan BA, Price SD, et al. Math1: an essential gene for the generation of inner ear hair cells. Science 1999;284:1837–41.
9. D'Amico-Martel A, Noden D. Contribution of placode and neural crest cells to avian cranial peripheral ganglia. Am J Anat 1983;163:351–72.
10. Carney PR, Couve E. Cell polarity changes and migration during early development of the avian peripheral auditory system. Anat Rec 1989;225:156–64.
11. Ma Q, Anderson DJ, Fritzsch B. Neurogenin 1 null mutant ears develop fewer, morphologically normal hair cells in smaller sensory epithelia devoid of innervation. I J Assoc Res Otokiryn 2000;1:129–43.
12. Mansour SL, Goddard JM, Capecchi MR. Mice homozygous for a targeted disruption of the proto-oncogene int-2 have developmental defects in the tail and inner ear. Development 1993;117:13–28.
13. Fritzsch B, Pirvola U, Ylikoski J. Making and breaking the innervation of the ear: neurotrophic support during ear development and its clinical implications. Cell Tissue Res 1999;295:369–82.
14. Fritzsch B, Farinas I, Reichardt LF. Lack of NT-3 causes losses of both classes of spiral ganglion neurons in the cochlea in a region specific fashion. J Neurosci 1997;17:6213–25.
15. Silos-Santiago I, Fagan AM, Garber M, et al. Severe sensory deficits but normal CNS development in newborn mice lacking trkB and trkC tyrosine protein kinase receptors. Eur J Neurosci 1997;9: 2045–56.
16. Steel KP, Bussoli TJ. Deafness genes: expressions of surprise. Trends Genet 1999;15:207–11.
17. Vincent C, Kalatzis V, Abdelhak S, et al. BOR and BO syndromes are allelic defects of EYA1. Eur J Hum Genet 1997;5:245–6.
18. Fraser FC, Sproule JR, Halal F. Frequency of the branchio-oto-renal (BOR) syndrome in children with profound hearing loss. Am J Med Genet 1980;7: 341–9.
19. Millman B, Gibson WS, Foster WP. Branchio-oto-renal syndrome. Arch Otolaryngol Head Neck Surg 1995;121:922–5.
20. Abdelhak S, Kalatzis V, Heilig R, et al. A human homologue of the Drosophila eyes absent gene underlies branchio-oto-renal (BOR) syndrome and identifies a novel gene family. Nat Genet 1997;15:157–64.
21. Wilkes D, Rutland P, Pulleyn LJ, et al. A recurrent mutation, ala391glu, in the transmembrane region of FGFR3 causes Crouzon syndrome and acanthosis nigricans. J Med Genet 1996;33:744–8.
22. Reardon W, Winter RM, Rutland P, et al. Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. Nat Genet 1994;8:98–103.
23. Meyers GA, Orlow SJ, Munro IR, et al. Fibroblast growth factor receptor 3 (FGFR3) transmembrane mutation in Crouzon syndrome with acanthosis nigricans. Nat Genet 1995;11:462–4.
24. Muenke M, Gripp KW, McDonald-McGinn DM, et al. A unique point mutation in the fibroblast growth factor receptor 3 gene (FGFR3) defines a new craniosynostosis syndrome. Am J Hum Genet 1997;60:555–64.
25. de Kok YJ, van der Maarel SM, Bitner-Glindzicz M, et al. Association between X-linked mixed deafness

and mutations in the POU domain gene POU3F4. Science 1995;267:685–8.
26. Everett LA, Glaser B, Beck JC, et al. Pendred syndrome is caused by mutations in a putative sulphate transporter gene (PDS). Nat Genet 1997;17:411–22.
27. Li XC, Everett LA, Lalwani AK, et al. A mutation in PDS causes non-syndromic recessive deafness. Nat Genet 1998;18:215–7.
28. Gong Y, Krakow D, Marcelino J, et al. Heterozygous mutations in the gene encoding noggin affect human joint morphogenesis. Nat Genet 1999;21:302–4.
29. Ferraz FG, Nunes L, Ferraz ME, et al. Townes-Brocks syndrome. Report of a case and review of the literature. Ann Genet 1989;32:120–3.
30. Kohlhase J, Wischermann A, Reichenbach H, et al. Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. Nat Genet 1998;18:81–3.
31. The Treacher Collins Collaborative Group. Positional cloning of a gene involved in the pathogenesis of Treacher Collins syndrome. Nat Genet 1996;12: 130–6.
32. Xu PX, Adams J, Peters H, et al. Eya1-deficient mice lack ears and kidneys and show abnormal apoptosis of organ primordia. Nat Genet 1999;23:113–7.
33. Kalatzis V, Sahly I, El-Amraoui A, Petit C. Eya1 expression in the developing ear and kidney: towards the understanding of the pathogenesis of branchio-otorenal (BOR) syndrome. Dev Dyn 1998;213:486–99.
34. Pignoni F, Hu B, Zavitz KH, et al. The eye-specification proteins So and Eya form a complex and regulate multiple steps in Drosophila eye development. Cell 1997;91:881–91.
35. Phelps PD, Coffey RA, Trembath RC, et al. Radiological malformations of the ear in Pendred syndrome. Clin Radiol 1998;53:268-273.
36. Johnsen T, Jorgensen MB, Johnsen S. Mondini cochlea in Pendred's syndrome. A histological study. Acta Otolaryngol (Stockh) 1986;102:239–47.
37. Scott DA, Wang R, Kreman TM, et al. The Pendred syndrome gene encodes a chloride-iodide transport protein. Nat Genet 1999;21:440–3.
38. Everett LA, Morsli H, Wu DK, et al. Expression pattern of the mouse ortholog of the Pendred's syndrome gene (Pds) suggests a key role for pendrin in the inner ear. Proc Natl Acad Sci U S A 1999;96: 9727–32.
39. Everett LA, Belyanseva IA, Noben-Trauth K, et al. Targeted disruption of mouse Pds provides insight about the inner-ear defects encountered in Pendred syndrome. Hum Mol Genet 2001;10:153–61.
40. Holley SA, Neul JL, Attisano L, et al. The Xenopus dorsalizing factor noggin ventralizes Drosophila embryos by preventing DPP from activating its receptor. Cell 1996;86:607–17.
41. Brunet LJ, McMahon JA, McMahon AP, Harland RM. Noggin, cartilage morphogenesis, and joint formation in the mammalian skeleton. Science 1998; 280:1455–7.
42. McMahon JA, Takada S, Zimmerman LB, et al. Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. Genes Dev 1998;12:1438–52.
43. Phippard D, Heydemann A, Lechner M, et al. Changes in the subcellular localization of the Brn4 gene product precede mesenchymal remodeling of the otic capsule. Hear Res 1998;120:77–85.
44. Phippard D, Lu L, Lee D, et al. Targeted mutagenesis of the POU-domain gene Brn4/Pou3f4 causes developmental defects in the inner ear. J Neurosci 1999; 19:5980–9.
45. Minowa O, Ikeda K, Sugitani Y, et al. Altered cochlear fibrocytes in a mouse model of DFN3 nonsyndromic deafness. Science 1999;285:1408–11.
46. Isaac C, Marsh KL, Pazekas WA, et al. Characterization of the nucleolar gene product, treacle, in treacher collins syndrome. Mol Biol Cell 2000;11: 3061–71.
47. Colvin JS, Bohne BA, Harding GW, et al. Skeletal overgrowth and deafness in mice lacking fibroblast growth factor receptor 3. Nat Genet 1996;12: 390–7.
48. Wilson J, Roth C, Warkany J. An analysis of the syndrome of malformations induced by maternal vitamin A deficiency. Effects of restoration of vitamin A at various times during gestation. Am J Anat 1953; 92:189–217.
49. Kelley MW, Xu XM, Wagner MA, et al. The developing organ of Corti contains retinoic acid and forms supernumerary hair cells in response to exogenous retinoic acid in culture. Development 1993;119: 1041–53.
50. Smith SM, Eichele G. Temporal and regional differences in the expression pattern of distinct retinoic acid receptor–beta transcripts in the chick embryo. Development 1991;111:245–52.
51. Choo D, Sanne JL, Wu DK. The differential sensitivities of inner ear structures to retinoic acid during development. Dev Biol 1998;204:136–50.
52. Frenz DA, Liu W, Galinovic-Schwartz V, Van De Water TR. Retinoic acid-induced embryopathy of the mouse inner ear. Teratology 1996;53:292–303.

53. Lohnes D, Mark M, Mendelsohn C, et al. Function of the retinoic acid receptors (RARs) during development (I). Craniofacial and skeletal abnormalities in RAR double mutants. Development 1994;120: 2723–48.
54. Lohnes D, Mark M, Medelsohn C, et al. Developmental roles of the retinoic acid receptors. J Steroid Biochem Mol Biol 1995;53:475–86.
55. Mark M, Ghyselinck NB, Wendling O, et al. A genetic dissection of the retinoid signalling pathway in the mouse. Proc Nutr Soc 1999;58:609–13.
56. Morriss-Kay GM, Ward SJ. Retinoids and mammalian development. Int Rev Cytol 1999;188: 73–131.
57. Collins MD, Mao GE. Teratology of retinoids. Annu Rev Pharmacol Toxicol 1999;39:399–430.
58. Ross SA, McCaffrey PJ, Drager UC, DeLuca LM. Retinoids in embryonal development. Physiol Rev 2000;80:1021–54.
59. Mendelsohn C, Mark M, Dolle P, et al. Retinoic acid receptor beta 2 (RAR beta 2) null mutant mice appear normal. Dev Biol 1994;166:246–58.
60. Brown NS, Smart A, Sharma V, et al. Thyroid hormone resistance and increased metabolic rate in the RXR-gamma-deficient mouse. J Clin Invest 2000;106:73–9.
61. Lee SK, Lee B, Lee JW. Mutations in retinoid X receptor that impair heterodimerization with specific nuclear hormone receptor. J Biol Chem 2000; 275:33522–6.
62. Represa J, Sanchez A, Miner C, et al. Retinoic acid modulation of the early development of the inner ear is associated with the control of c-fos expression. Development 1990;110:1081–90.

SUGGESTED READING

Anson JA, Davies J, Duckert LG. Embryology. In: Paparella MM, Shumrick DA, editors. Otolaryngology. Vol 1. Philadelphia (PA): WB Saunders; 1991. p. 3–22.

Donaldson JA, Miller JM. Anatomy of the ear. In: Paparella MM, Shumrick DA, editors. Otolaryngology. Vol 1. Philadelphia (PA): WB Saunders; 1973. p. 75–110.

Donaldson JA, Lambert PM, Duckert LG, Rubel EW. The ear: developmental anatomy. In: Donaldson JA, Lambert PM, Duckert LG, Rubel EW, editors. Surgical anatomy of the temporal bone. New York: Raven Press; 1992. p. 19–108.

Krsti'c RV. Human microscopic anatomy. New York: Springer-Verlag; 1997.

Moore KL. The developing human. Philadelphia (PA): WB Saunders; 1977.

Schuknecht HF. Anatomy. In: Schuknecht HF. Pathology of the ear. Philadelphia (PA): Lea & Febiger; 1993. p. 31–75.

Sulik KK. Embryology of the ear. In: Gorlin RJ, Toriello HV, Cohen MM, editors. Hereditary hearing loss and its syndromes. New York: Oxford University Press; 1995. p. 22–42.

CHAPTER 3

Molecular Biology of Hearing and Balance

Allen F. Ryan, PhD, David J. Lim, MD

Molecular biology is a branch of modern biology in which biologic processes are studied using physical, biochemical, cellular, and genetic principles at the molecular level. The techniques of molecular biology use deoxyribonucleic acid (DNA) sequences that encode proteins or that regulate protein expression.

This emerging science has contributed to major advances in our understanding of cellular function, in health as well as in disease. It also has resulted in a deeper understanding of the pathogenesis of many diseases and the revolutionary development of new approaches in diagnosis, prevention, and treatment of diseases, including gene therapy. This chapter reviews the rapidly developing molecular methodologies that have made contributions to advance our understanding of the cellular functions and dysfunctions of the auditory and vestibular systems, in molecular terms. It reviews recent advances achieved with these techniques in both the basic and clinical arenas.

MOLECULAR METHODOLOGY

The rapid expansion of molecular research has been the result of the introduction of several powerful methodologies for isolating gene sequences and characterizing their expression. At the center of these methodologies are gene-cloning techniques, which rely on the replication of DNA sequences in bacteriophages and plasmids, generating large amounts of DNA that can be screened, characterized, and sequenced. Another crucial method is reverse transcription of messenger RNA (mRNA), in which viral enzymes capable of copying RNA into DNA are used to generate complementary DNA (cDNA) copies of mRNAs isolated from tissue. Reverse transcription and cloning have allowed the rapid discovery of new genes, usually based on homology with known DNA sequences or on assay of functional characteristics when cDNAs are cloned in expression vectors. This has led to the isolation of many families, some quite extensive, of genes encoding related proteins. Other widely used methods in molecular biology include the polymerase chain reaction (PCR) for amplification of DNA sequences present in low copy numbers[1] and in situ mRNA hybridization for the localization of gene expression to individual cells.[2] (Terms that are frequently used in molecular biology are listed in Table 3–1, and a number of important concepts and methods are illustrated in Figures 3–1 through 3–5).

Gene Expression Systems

The expression of cloned genes in biologic systems ranging from bacteria and *Xenopus* oocytes[3] (Figure 3–6) to eukaryotic cell lines and intact mammals has become a major tool in biologic research. The expression of functional proteins in cells has added to our knowledge of numerous molecules such as neuronal receptors, ion channels, and transcription factors. Targeted mutations in the translated sequences of genes have made fundamental contributions to the molecular determinants of protein function, as in studies that identify the functional domains of potassium channels and provide evidence to support a "ball-and-chain" model of their action[4,5] and amino acid sequences governing interaction between subunits of ion transport adenosine triphosphatases (ATPases).[6] The use of different promoter elements and mutations within promoters has dramatically increased our understanding of the control of gene expression.[7]

Transgenic Animals

The influence of genes in the organism has been studied by placing the gene sequences of interest under control of tissue-specific promoter sequences that target expression to given tissue and inserting

TABLE 3–1. Terms Used in Molecular Methods

Alternative splicing	Different ways of assembling exons to produce different mature mRNAs
Base pair (bp)	A combination of A with T or G with C formed by hydrogen bonding between two antiparallel strands of a DNA double helix
cDNA clone	A vector containing a cDNA molecule from another organism
cDNA	A single strand of DNA complementary to an mRNA and synthesized from it by reverse transcription
DNA library	A set of cloned genomic DNA fragments from an organism or cDNAs from a cell type or tissue
Exon	Any segment of an interrupted gene that is represented in a mature mRNA. Segments of DNA separating exons are called introns.
Gene family	A group of genes related by sequence similarity
Genomic clone	A selected host cell with a vector containing a fragment of genomic DNA from a different organism
Genomic DNA	All DNA sequences of an organism
Host cell	A cell (usually a bacterium) in which a vector can be propagated
Hybridization	Localization of mRNA or DNA in a cell or tissue section by hydrogen bonding of complementary sequences found in a probe that can be detected
Library	A complete set of genomic clones from an organism or of cDNA clones from one cell type
Linkage analysis	A means of identifying the chromosomal location of genes based on the frequency of recombination between the gene of interest and known loci
Oligonucleotide	A short DNA sequence, usually less than 100 base pairs
PCR	The polymerase chain reaction. Exponential amplification of a DNA sequence using paired oligonucleotide primers, DNA polymerase to extend the primers by synthesizing new DNA strands, and thermal cycling to denature the duplex DNA and allow a new round of primer binding.
Phage	Bacteriophage, a bacterial virus, often engineered to carry inserted DNA, as in DNA libraries
Plasmid	A small, circular, extrachromosomal DNA molecule capable of replicating independently in a host cell
Polymorphism	Two or more alleles of the same gene (ie, Rh+ or Rh– individuals) in a breeding population
Promoter	A region of DNA involved in binding RNA polymerase to initiate transcription
Recombination	A reciprocal exchange of DNA between sister chromatid during meiosis
Restriction enzyme	A bacterial enzyme that recognizes and cleaves a particular short DNA sequence. Often used to cut and join two DNA sequences, as in cloning.
Reverse transcription	DNA synthesis from an RNA template, mediated by a retroviral enzyme called reverse transcriptase
RNA splicing	Removal of introns from an RNA sequence to produce a mature mRNA
Transfection	A method for introducing exogenous DNA into eukaryotic cells
Transgene	An artificial gene constructed from a promoter active in the tissue or cell to be studied, a coding sequence of interest, and a termination sequence
Transgenic	Animal (or plant) that has stably incorporated one or more genes from another cell or organism
Transcript	The RNA product produced by template-dependent RNA synthesis from a DNA sequence
Transcription	RNA synthesis on a DNA template, mediated by an enzyme called RNA polymerase
Transcription factor	Term loosely applied to any protein required to initiate or regulate transcription in eukaryotic or prokaryotic cells
Translation	Synthesis of protein on an mRNA template
Transposon	DNA elements that can transpose themselves (move) from one position in a DNA molecule to another

cDNA = complementary deoxyribonucleic acid; mRNA = messenger ribonucleic acid.

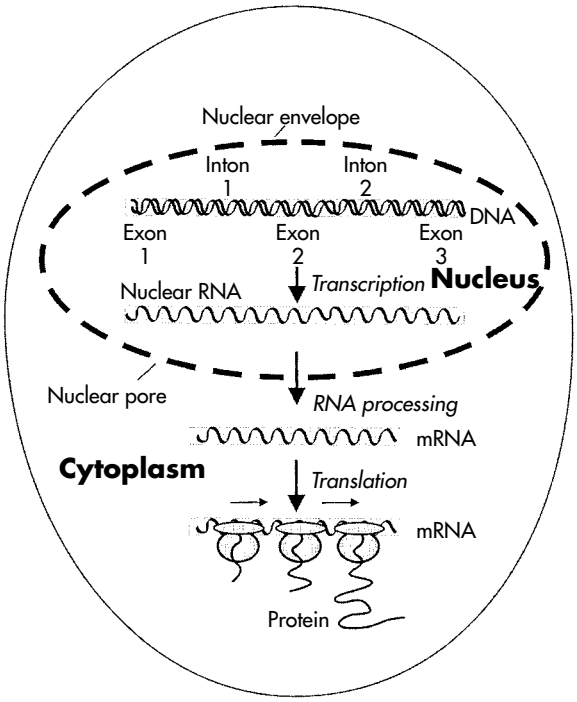


FIGURE 3–1. Summary of steps by which proteins are synthesized by deoxyribonucleic acid (DNA) in a eukaryotic cell. A nuclear envelope with pores segregates the DNA from cytoplasm. Adapted from Gilbert SF, 1.[8] mRNA = messenger ribonucleic acid.

these transgenes into the DNA of mice to create a transgenic animal (Figure 3–7). The resultant transgenic animals then display the effects of transgene expression and can be used to document the effects of genes and/or promoters.[9] A serendipitous by-product of this technique has been insertional mutations. In a small percentage of transgenic animals, insertion of the transgene disrupts an endogenous gene and produces a morphologic or behavioral change completely unrelated to the transgene. In these animals, the mutation site is marked by the inserted DNA sequence,[10,11] simplifying the process of localizing the mutated gene and its normal counterpart. Insertional mutagenesis has led to the identification of several new genes, including genes controlling limb development[12] and motor control.[13] However, transgenic insertions are frequently extremely complex, often involving deletion of hundreds of kilobases at or near the site of transgene insertion. For this reason, the identification of genes mutated by insertion in transgenic animals has sometimes proven to be difficult, if not impossible.

Several insertional mutations in transgenic animals that affect the inner ear have been identified. For example, Crenshaw et al found that 1 of 14 lines of mice with a v-src transgene showed behavior consistent with an inner ear defect.[14] The inner ear of these animals showed collapse and degeneration of the pars superior of the inner ear at about 1 month of age. This was preceded by a period of severe endolymphatic hydrops, suggesting that the gene disrupted in mutation affects inner ear fluid balance. The gene has been localized to a 6 cM (centimorgan) region on chromosome 1, near the interleukin-1 (IL-1) receptor.[15]

Gene Knockouts (Gene Targeting)

A method by which a gene knockout can be created is illustrated in Figure 3–8. Single-gene knockouts can be lethal or have extensive effects on development and adult function. For example, knockouts of key factors in embryonic development are typically lethal prior to birth, and deletion of the genes encoding the neurotrophins or their receptors can produce serious deficits in sensory neurons.[16,17] Alternatively, knockouts can have little influence since the process of development appears to have a high degree of redundancy. An example is the

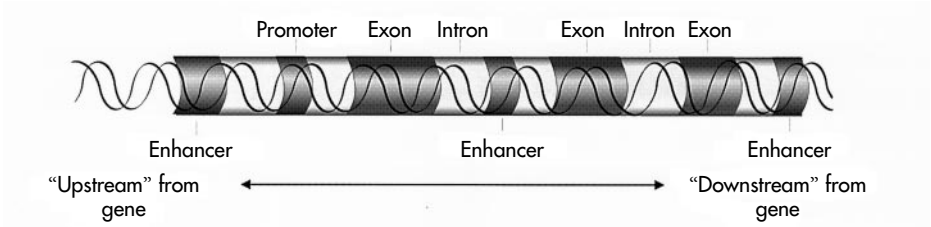


FIGURE 3–2. Basic structure of a developmentally regulated gene is illustrated. The promoters of most genes encoding proteins are found at the 5′ (upstream) end of the gene. Enhancers are often even farther upstream, but they can also occur within an intron of the gene or at the 3′ end. Proteins that bind to promoters and enhancers interact to regulate transcription of the gene. Adapted from Gilbert SF.[8]

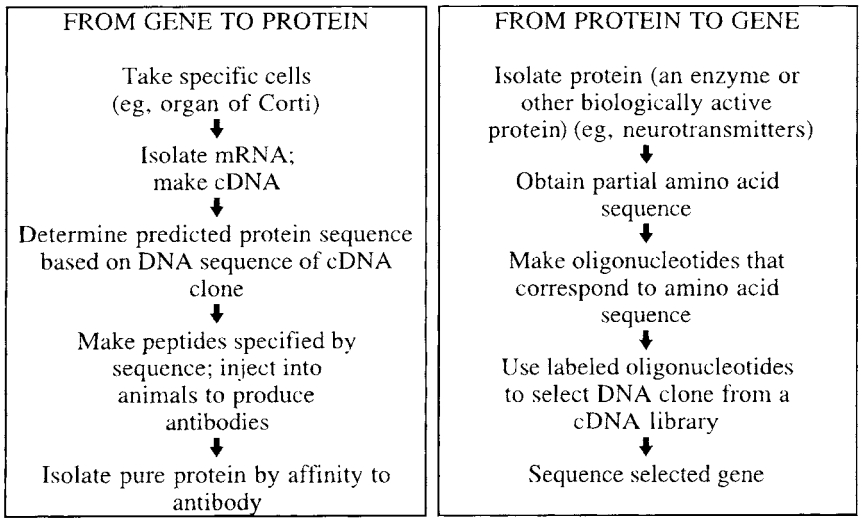


FIGURE 3–3. It is now possible to identify a messenger ribonucleic acid (mRNA) of interest (such as obtained from the inner ear or even the organ of Corti) and to use it to isolate the protein it encodes without knowing the function of that protein. Conversely, it is also possible to sequence part of a protein that has a specific function and then to synthesize an oligonucleotide that can be used to identify and isolate the gene that encodes the complete protein. Adapted from Gilbert SF.[8] cDNA = complementary deoxyribonucleic acid.

regulation of differentiation in muscle, in which several factors act to drive a cell toward the muscle phenotype.[18]

Molecular Diversity (Gene Families)

Gene cloning has revealed the existence of many families of genes that mediate important aspects of cell function, including the genes that encode neuronal receptors such as acetylcholine receptors.[19,20] The first of these receptor genes was identified by injecting RNA molecules transcribed in vitro from cDNA clones isolated from brain cDNA libraries into *Xenopus* oocytes.

When exposed to the appropriate transmitter, the oocytes showed intracellular potential changes consistent with the presence of specific receptors on their surfaces. Subsequent low-stringency screening of cDNA libraries for related sequences yielded families of receptor genes.[21]

Other gene families that have been isolated recently include voltage-gated potassium channels,[22] Ca ATPases,[23] G proteins, and numerous growth factors and their receptors.[24] Several gene families consist of subunits that can be assembled in a variety of combinations, leading to increased diversity of functional expression from a limited set of genes. For example, the nicotinic acetylcholine receptor consists of α and β subunits. The existence of genes encoding several different α and β subunits allows a multitude of different combinations, each with different functional properties.[19] The non–(*N*-methyl-D-aspartate) glutamate receptors have been hypothesized to amalgamate as heteropentamers of subunits, with each subunit isoform capable of complexing with several other isoforms.[25,26] This provides the potential for a very large number of structurally and functionally different glutamate receptors, explaining in part the high level of pharmacologic diversity observed at synapses using the

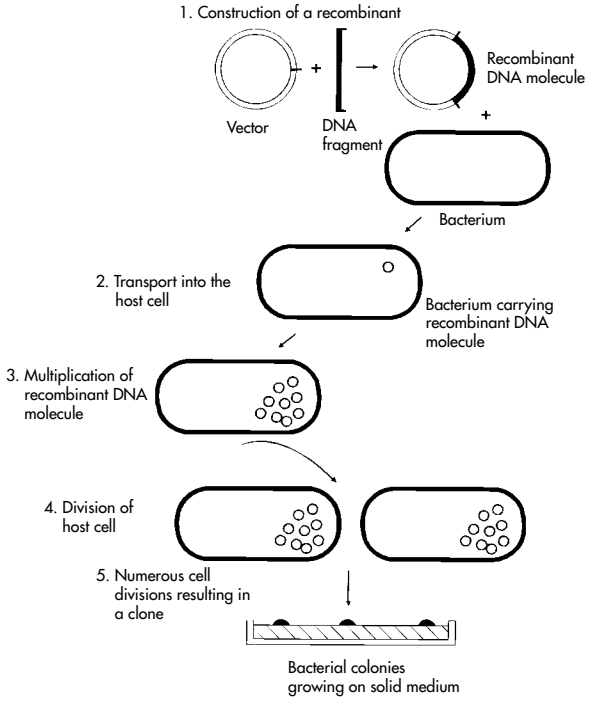


FIGURE 3–4. The basic steps in gene cloning are schematically illustrated. Adapted from Brown TA.[27]

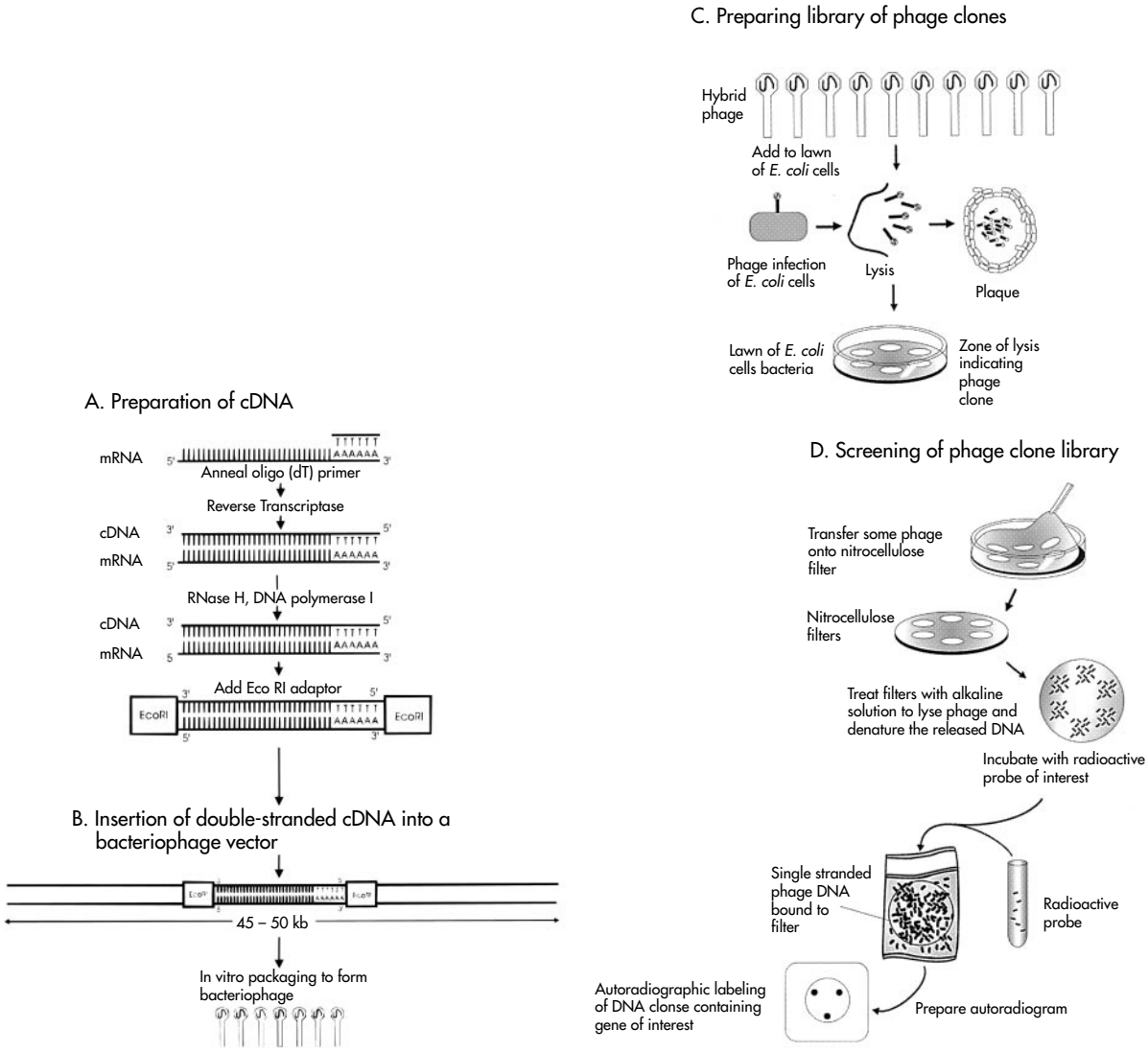


FIGURE 3–5. Schematic illustration of protocol to make complementary deoxyribonucleic acid (cDNA) libraries. Step A: Messenger ribonucleic acid (mRNA) is isolated from a tissue of interest and reverse transcribed into cDNA. This cDNA represents all of the genes being expressed in the tissue. It is made double-stranded in a nick translation reaction, in which ribonuclease H nicks the mRNA strand, creating a primer for second-strand DNA synthesis by DNA polymerase I. Sequences (adapters) that allow ligation into a cloning vector are added to double-stranded cDNA. Step B: The cDNA can then be inserted into specially modified vectors, in this case bacteriophages. Step C: Phages containing the recombinant DNA will infect *Escherichia coli* and reproduce, making many copies and eventually lysing the bacteria forming plaques. Step D: The plaques are transferred to nitrocellulose paper and treated with alkali to lyse the phages and denature the DNA in place. These filters are then incubated in a radioactive probe. This can be a short DNA sequence (oligonucleotide), a cDNA, or other DNA fragment. In the case of differential cDNA library screening, the same phage library is screened with radioactive cDNA probes reverse transcribed from mRNA isolated from two different tissues. This allows the identification of a relatively abundant mRNA that would be found in one type of tissue but not in the other. Adapted from Gilbert SF.[8]

same neurotransmitter. In addition, mRNA transcribed from a single gene can often be assembled into mRNAs encoding functionally different forms of the protein product from different combinations of exons by the process of alternative splicing, leading to even greater variation.[23] Increased understanding of the molecular basis for functional diversity of proteins may be one of the most significant contributions of molecular biology.

Molecular Genetic Methods

Molecular methods have also transformed the field of genetics. Gene localization by kindred analysis classically depended on the linkage of a phenotypic trait associated with a genetic disorder to a relatively small number of phenotypic traits controlled by genes whose locations were already known. Recently, several developments have led to a dramatic increase in our ability to localize genes. The number of genetic markers at known locations in the genome increased exponentially after the introduction of molecular methodology, from a few dozen in the early 1980s to nearly 10,000 in 1991,[28] and the number continues to increase. This increase was the result of several critical technical developments. The isolation of large numbers of restriction enzymes to cut DNA at specific locations has led to the discov-

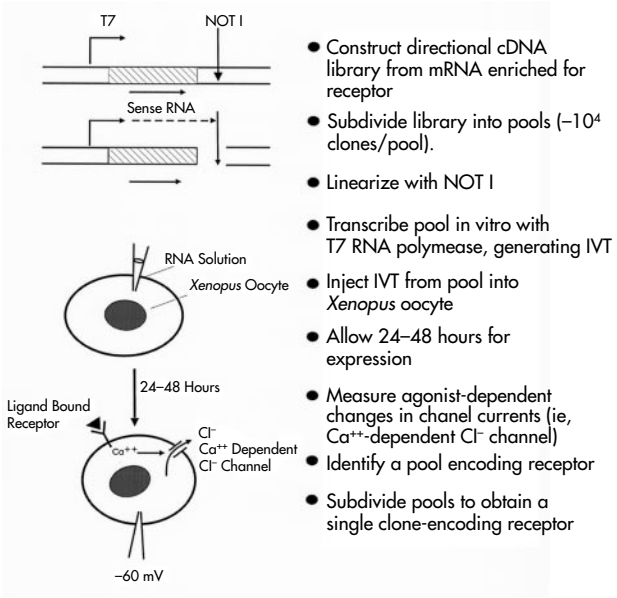


FIGURE 3–6. Expression cloning using *Xenopus* oocytes. The steps required to generate in vitro transcripts from pools of directional complementary deoxyribonucleic acid (cDNA) library clones are shown. Included are in vitro transcription, injection into *Xenopus* oocytes, and voltage clamp assay for a ligand-evoked opening of the calcium-dependent chloride channel. *Shaded boxes* indicate the receptor-coding region. IVT = in vitro transcript. Adapted from Battey JF et al.[3]

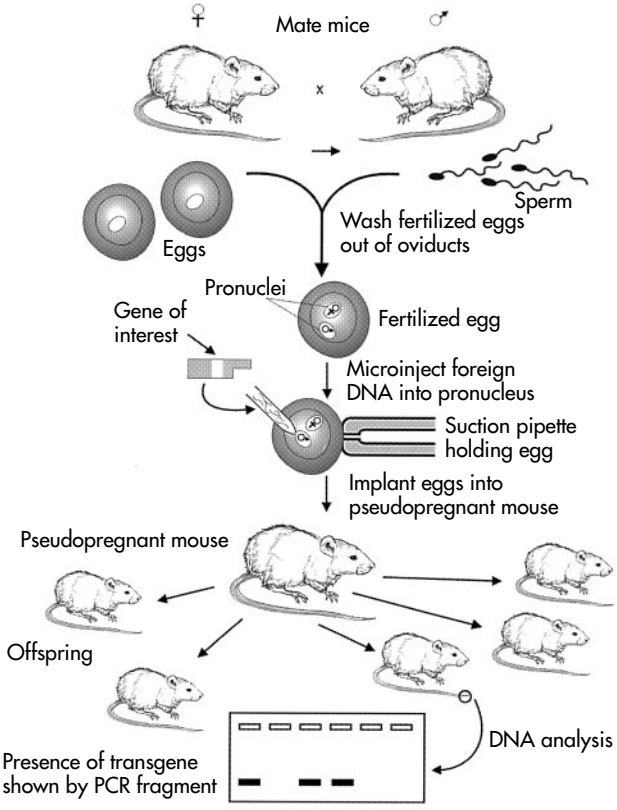


FIGURE 3–7. Schematic illustration of production of transgenic mice using the microinjection technique. A transgene is constructed from a promoter that directs expression in the tissue of interest, the gene to be studied, and a termination sequence. Fertilized eggs are collected from mated females, and the transgene is injected into one of the two pronuclei. The injected eggs are transferred to foster mothers, which are female mice made pseudopregnant by mating with vasectomized males. Three weeks after birth, the offspring are checked for the presence of the transgene by Southern blot analysis of DNA extracted from a small piece of the tail. Screening can be performed rapidly using the polymerase chain reaction (PCR) if suitable primers are available. Three of the offspring carry the transgene in the example given. Adapted from Watson JD et al.[29]

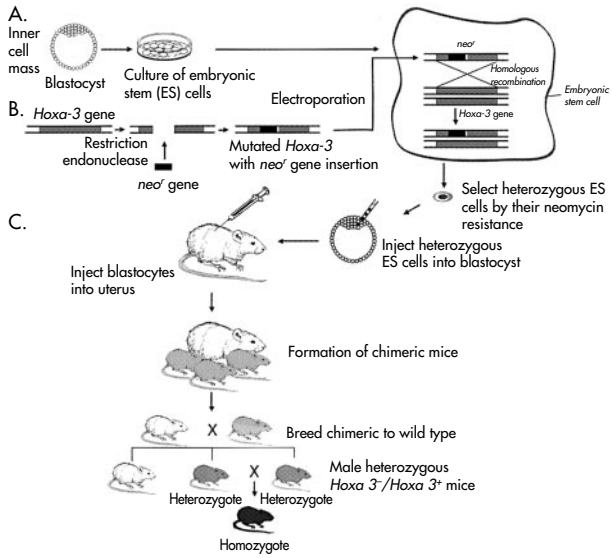


FIGURE 3–8. Schematic illustration of technique for gene targeting using the *Hoxa-3* gene as an example. (A) Embryonic stem (ES) cells are cultured from inner cell mass. (B) The cloned *Hoxa-3* genes are cut with a restriction enzyme, and the neomycin resistance gene is inserted into the region that encodes the DNA-binding site of the protein. The mutant *Hoxa-3* genes are introduced into ES cells, in which homologous recombination exchanges a wild-type gene for the mutant copy. These cells are selected with neomycin. (C) Heterozygous ES cells are inserted into the inner cell mass of a wild-type embryo, and the blastocyte is returned to the uterus. The resulting mouse is a chimera composed of heterozygous *Hoxa-3* tissues and wild-type *Hoxa-3* tissues. Mating of the chimeric mice to wild-type mice produces heterozygous *Hoxa-3* offspring if the ES cells contributed to the germ line of the chimeras. These heterozygous mice can be bred together, and some of their progeny should be homozygous mutants of the *Hoxa-3* gene. Adapted from Gilbert SF.[8]

ery of many restriction fragments at defined chromosomal locations whose length differs between individuals (a phenomenon called polymorphism; the more polymorphic a marker, the more useful it is for gene localization). These differences can readily be identified by Southern blotting (Figure 3–9).

Another family of markers has been developed by exploiting the tendency of DNA to form repeating sequences called minisatellites. Repeats of short DNA sequences, or of C-A or G-T pairs, are ubiquitous in the genome, are often highly polymorphic, and can be identified by PCR. In animals, interspecies crosses between the laboratory mouse *Mus musculus* and closely related mouse species such as *Mus spretus* have been used to generate markers at many loci.

Taken together, these developments have greatly increased both the ease of linkage analysis and the number of defined chromosomal loci against which to test a given phenotype for linkage. This has allowed more precise localization of the genes determining traits of interest, increasing the chance that the genes can be identified at the molecular level by fine mapping, chromosome walking, testing of candidate genes within a region, and/or exon trapping. The proliferation of gene markers has contributed to the location and, in many cases, identification of genes involved in diseases, such as retinoblastoma[30] and Duchenne's muscular dystrophy.[31]

MOLECULAR BIOLOGY OF THE INNER EAR

For the past 20 years, molecular biology has been increasingly applied to research problems in virtually every area of biology and medicine. The expansion of molecular research has been fueled by the introduction of many powerful methodologies for isolating and characterizing gene sequences, determining their expression patterns, inducing expression in cells or organisms, and producing specific mutations. Owing in part to the small size and difficult access of the ear, the research community studying the auditory system was relatively slow to use the techniques of molecular biology, with initial publications using these methods appearing in the early 1990s. Since this time, the use of molecular biology as a tool to study hearing and deafness has increased dramatically. Many auditory laboratories in the United States and worldwide have added molecular methods to their research programs, and several molecular laboratories have begun to study the auditory system. This has led to the identification and characterization of numerous genes involved in hearing and deafness.

Progress beyond the gene identification stage has been more limited. A survey of molecular studies in other fields suggests that our field lags behind other medical disciplines in applying these powerful

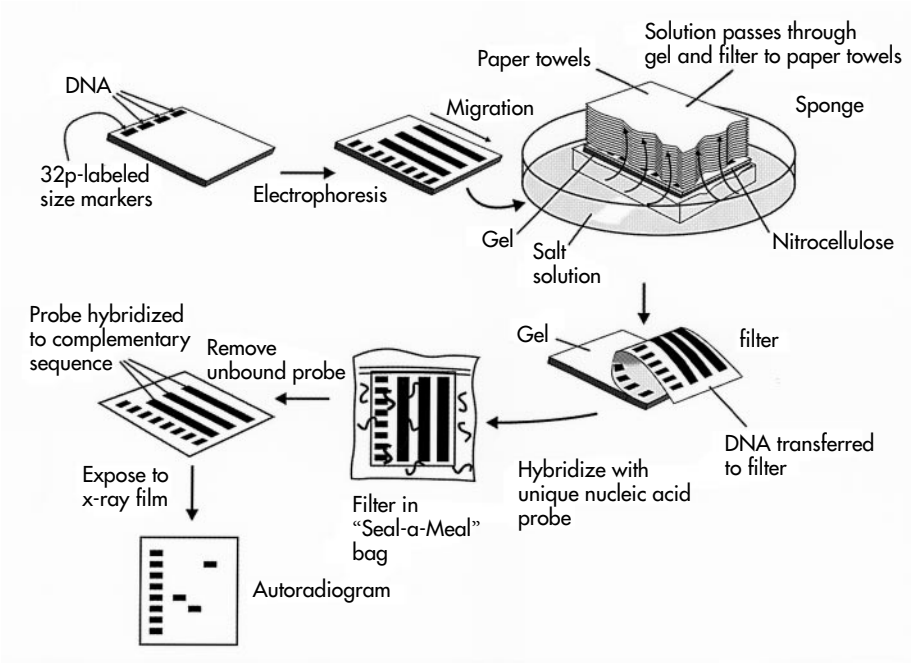


FIGURE 3–9. Schematic illustration of Southern blotting. Deoxyribonucleic acid (DNA) digested with restriction enzymes is loaded onto a gel and electrophoretically separated by molecular size. After the fragments of DNA are separated, the DNA is denatured into single strands. The gel is then placed on a support on top of filter paper saturated with buffer. A nitrocellulose or nylon filter is placed on top of the gel, and absorbent towels are placed on top of the filter. The transfer buffer makes its way through the gel, nitrocellulose paper, and towels by capillary action, taking the DNA with it. The single-stranded DNA sticks to the filter. The filter is then removed, the DNA is immobilized on to the filter, and the filter is hybridized with a radioactively labeled probe. The position of the DNA fragment complementary to the probe appears as a band on x-ray film, which reflects the position of the DNA fragment in the gel. Adapted from Watson JD et al.[29]

methods. The limited amount of tissue that is available from the middle ear, labyrinth, and many central auditory structures, coupled with the diversity of cell types present, has contributed to the disparity. This has slowed the generation of cDNA libraries, especially those for specific cell types. It has also made the application of other methods for cataloguing gene expression, such as gene array technology, difficult. A second impediment to progress is a relative paucity of in vitro models of many aspects of auditory development and function that are appropriate for molecular manipulation. Thus, molecular methods often need to be applied to the intact organism as opposed to culture systems or cell lines.

Molecular biology has been minimally applied in the clinic. Molecular diagnostic techniques are becoming available for genotyping individuals with potential genetic disorders affecting hearing or for identifying disease pathogens as in otitis media (OM). However, molecular therapy has yet to be applied to the ear. However, many otologists recognize the potential of molecular medicine to improve diagnosis and treatment in the future.

As noted above, molecular biologic studies of the ear have increased greatly in the past several years, to the point where it is no longer possible to provide a comprehensive review in the space available. However, some recent highlights are presented below.

Deafness Genes

Molecular Basis of Deafness A notable exception to the generally slower rate of molecular progress in

our field is the molecular genetics of deafness, perhaps because this field is less dependent on extracting molecular information from the auditory tissues themselves. We now know the identity of more than 16 genes of the 60+ loci known to be involved in nonsyndromic deafness in humans, and many new syndromic deafness genes have also been identified.[32] For example, Alport's syndrome has been associated with a mutation in the gene encoding the basement membrane collagen (*COL4A5*).[33,34] Waardenburg's syndrome types I and III have been identified with mutations in the transcription factor gene *PAX3*.[35] Myosin VIIA (MyoIIIA) has been identified as the gene responsible for Usher's syndrome type IB.[36] Linkage of neurofibromatosis 2 (NF2) to chromosome 22q[37,38] led to fine-scale mapping[39] and then to identification of a mutated gene, which encodes a new member of the family of 4.1 cytoskeletal associated proteins known as merlin.[40,41] Neurofibromatosis 2 is an autosomal dominant inherited disease characterized by bilateral vestibular schwannomas and other central nervous system tumors and is often associated with hearing impairment caused by cochlear nerve compression. Similarly, testing of candidate genes near the linkage site for X-linked deafness with stapes fixation (*DFN3* = nonsyndromic deafness X-linked gene third to be discovered) led to the identification of the mutated gene as *Brn-4,* which encodes a transcription factor with a POU domain (*POU3F4*).[42] A mutation in the related gene, *Brn-3.1*, causes nonsyndromic deafness in *DFNA15* (nonsyndromic deafness autosomal dominant gene fifteenth to be discovered). These transcription factors are members of a gene family that is expressed in the cochlea during development.[43] The indentification of mutations in the connexin 26 gene, encoding a gap junction protein,[44] as the basis for DFNA1 is particularly significant. Mutations in this gene are particularly common, and it is estimated that up to 50% of all nonsyndromic hereditary deafness in the United States is caused by mutations in this gene.[45] A defect in mitochondrial DNA that increases susceptibility to aminoglycoside ototoxicity and that, in combination with an autosomal mutation, leads to maternally inherited nonsyndromic deafness has also been characterized.[46–48]

The gene responsible for Usher's syndrome type 1C has been identified and reveals a unique pathologic mutation.[49] Also of note is the finding in DFNA10 of mutations in *EYA4*,[50] a member of a family of genes in which defects in another member, *EYA1*, cause a separate hereditary hearing impairment disorder, branchio-oto-renal syndrome.[51,52]

In addition, around 50 new hearing- and/or balance-defective mouse mutants have been created by two large European ethylnitrosourea (ENU) mutagenesis programs,[53,54] with more becoming available from other, similar programs in the United States. A number of the genes involved in deafness in these new mouse mutants and other older mouse mutations have been identified. An example is the novel unconventional myosin, myosin 15, which is mutated in the Shaker 2 mouse. This example is particularly notable for the use of bacterial artificial chromosome (BAC) transgenic mice to identify the mutated gene.[55] With the completion of draft sequences of the human genome and progress toward completion of mouse genome sequencing, the transition from linkage to gene identification has and will become increasingly more rapid. Many gene knockouts also have effects on the development or function of the ear, which has allowed us to build up rapidly our knowledge of the molecular basis of auditory function and development.[56] Examples include genes encoding PMCA2,[57] Brn-4,[58] and otogelin.[59]

There is an important trend emerging toward using the information obtained from genetic screening to study the functional biology of the inner ear, for example, the work of Karen Steel and others showing changes in susceptibility to ototoxicity in myosin VIIA–deficient animals.[60] Another example is pendrin, which is giving insights into cochlear ion transport processes.[61] Microarray analysis is also beginning to be used to investigate changes in patterns of gene expression resulting from single-gene mutations, giving clues to the pathways affected by the mutations.

Molecular genetics is an arena in which there is extensive interaction between clinicians and basic scientists. In particular, several genes involved in inherited deafness have been first cloned in basic science laboratories or identified with the aid of mouse models.[49,62,63]

Linkage and identification of mutations causing deafness and vestibular disorders have immediate clinical significance since they allow genetic testing and counseling and may eventually provide a basis for gene therapy. They are also proving to be an important means of identifying genes whose expres-

sion is important for normal cochlear development and function.

Inner Ear Development

Many genes involved in auditory ontogeny have been identified by screening for gene expression in developing auditory tissues. Polymerase chain reaction and in situ hybridization have provided a wealth of information regarding which genes are expressed and in which tissues. Whereas expression of a gene in a cell or tissue suggests a functional role, mutation of the gene leading to a phenotype provides much more convincing evidence that the molecule functions within the context of the intact tissue and organism. Many genes involved in syndromic and nonsyndromic deafness appear to be critical developmental regulators, such as the unconventional myosins that are involved in stereociliary morphogenesis.[64] Similarly, despite the risks of redundancy, the targeted deletion of an increasing number of genes has illuminated their role in auditory development and/or function. This includes a variety of developmental genes that influence organogenesis of the labyrinth or the development of inner ear cell types.[56] Many such genes involved in inner ear morphogenesis are known to play roles in the morphogenesis of other developing structures.[65–67]

Several genes have been shown using molecular techniques to affect the basic morphogenesis of the labyrinth, beginning at the stage of the otocyst. For example, retinoids[68,69] affect ear development. The retinoic acid receptor (RAR) family of nuclear receptors for retinoic acid (RA) is one of two groups of DNA binding proteins that are activated by interaction with this ligand. They serve as ligand-inducible regulators of transcription by activating RA responsive promoters. There are three RAR gene isotypes, RARa, RARb, and RARg, each with their own isoforms, and they have widespread expression patterns in the developing inner ear. Retinoic acid receptor a expression is ubiquitous in cochlear sensory and nonsensory structures. Retinoic acid receptor b is expressed mainly in structures of mesenchyme origin and in vestibular sensory epithelia. Retinoic acid receptor g is located in the cochlear and vestibular compartments in the epithelium of the inner ear.[70] Because of functional redundancy, deletion of individual receptor isoforms has relatively little effect on the development of the auditory system. However, deletion of both a and g isotypes results in absence of the stapes, an abnormal incus, a small and incomplete cartilaginous otic capsule, and absence of the organ of Corti and spiral ganglion.[71]

Bone morphogenetic proteins (BMPs) are members of the transforming growth factor-alpha (TGF-α) superfamily. Oh and Wu investigated BMP (BMP-4, -5, -7) expression during development of the chick otocyst.[72] In situ hybridization analysis showed that BMP-4 mRNA was present in the future ampullae of the three semicircular ducts on embryonic day 2.5 to 3. Bone morphogenetic protein 5 mRNA, however, was only found transiently in the future ampulla of the posterior semicircular duct. Bone morphogenetic protein 7 mRNA, on the other hand, was initially expressed in most parts of the otocyst and became restricted to specific regions by embryonic day 3. These investigators suggested that BMPs may play an important role in the differentiation of inner ear sensory organs. Inactivation of BMP-4 results in embryonic lethality[73]; however, mice heterozygous for BMP-4 deletion exhibit abnormal vestibular behavior and malformations of the lateral semicircular canal,[74] suggesting a role in labyrinthine patterning.

In the sensory epithelia of the inner ear, several genes involved in hair cell fate selection and differentiation have recently been identified with this approach. For example, deletion of the *Math1* gene in mice results in failure of cells in the organ of Corti to adopt the hair cell phenotype.[75] In addition, exit of committed hair cells from the cell cycle appears to be dependent on expression of *p27Kip1* since deletion of this gene leads to formation of supernumerary hair cells.[76,77] The related factor *p19Ink4d* may play a similar role (Chen P and Segil N, unpublished observations). Hair cell differentiation requires expression of the POU domain transcription factor Brn-3.1. In mice null for *Brn-3.1*, hair cells never develop cuticular plates or stereocilia, and many eventually die.[78,79] The differentiation of hair cells from supporting cells appears to be regulated in addition by Notch/Delta signaling since deletion of the gene encoding the Delta family protein Jagged 2 results in increased numbers of hair cells, perhaps via down-regulation of *Notch* in supporting cells.[80]

Understanding the lineage relationships between hair and supporting cells and the molecules involved in a cell's decision to become either a hair

or a supporting cell provides insights into mechanisms of hair cell regeneration. Delta appears to be involved in hair cell regeneration in the avian basilar papilla. The underlying supporting cells begin expressing Delta upon hair cell loss.[81] In perhaps the most intriguing recent observation, Zheng and Gao have shown that transfection of the developing organ of Corti with the Math1 coding sequence results in the adoption of the hair cell phenotype by cells far outside the normal location of hair cells.[82] Math1 thus appears to be both necessary and sufficient for promoting hair cell fate selection and differentiation.

Neurotrophic growth factors such as basic fibroblast growth factor (bFGF), TGF-α, and neurotrophins, such as neurotrophin 3 (NT-3) and brain-derived neurotrophic factor (BDNF), are necessary for survival and neurite extension in auditory and vestibular neurons during both early and late stages of ear development.[83] Messenger RNAs encoding both BDNF and NT-3 are expressed in the inner ear during development.[84,85] Cochlear and vestibular neurons both respond to BDNF and NT-3 in vitro.[16] Inhibition of BDNF and NT-3 expression using antisense oligonucleotides prevents neurite extension in vitro.[86] Finally, using mutant mice with knockout of the *BDNF* gene or *NT-3* gene or both genes, Ernfors et al demonstrated that BDNF is critical for survival of the vestibular ganglion and maintenance of both afferent and efferent innervation.[87] These data are consistent with observations in mice lacking a functional *NT-3* gene reported by Farinas et al.[17] In the cochlea, *BDNF* gene–deleted mice showed loss of presumptive type II ganglion cells and afferent innervation to outer hair cells, whereas *NT-3* gene–deleted mice showed loss of presumptive type I ganglion cells and the majority of afferent innervation. The double mutants lose all vestibular and cochlear ganglion cells.[16]

Molecular Substrates of Auditory Function

Hair Cell Transduction The major function of the sensory hair cells in the inner ear is to transduce mechanical energy into neural information. The sensory cells of both the cochlea and the vestibular sensory organs are mechanoreceptors equipped with a geometrically arranged sensory stereociliary bundle on each sensory cell surface. Deflection of the ciliary bundle by mechanical force toward the kinocilium (or basal body in the case of the cochlea) is excitatory and toward the opposite direction is inhibitory for neural discharges. Deflection of the stereociliary bundle toward the excitatory direction is now known to open transduction channels located on the upper part of the stereocilia.[88–90] The opening of the transduction channels is believed to be mediated by filamentous structures connecting the tip of the lower stereocilium to the neighboring taller stereociliary surface known as tip links.[91,92] Most of the molecules responsible for the hair cell transduction process have yet to be identified. However, some progress has been made in the transduction channel[93,94] and the adaptation motor in vestibular sensory hairs.[95] The precise location of the transduction channels is not yet established. It is believed to be associated with tip links because a calcium chelator eliminates both transduction current and tip links.[91] Hackney et al used an antibody to the amiloride-sensitive sodium channel to label the side of the stereocilia just below the tip-link attachment site.[93] A recent advance is the identification of a mechanically sensitive ion channel of the TRP superfamily that is required for mechanotransduction in the bristle organ of *Drosophila*.[96] A mammalian homologue would be an excellent candidate for the transduction channel of hair cells.

To retain sustained sensitivity during transient displacement, the vestibular sensory cell must adapt to sustained stimuli. Adaptation of vestibular sensory cell is thought to be mediated by a "slipping" of tip-link attachment and by an active "tensioning."[97] Readjusting tension of the tip link is necessary for the adaptation of the vestibular sensory cells.[95,98] This tensioning is suggested to be mediated by 120 kDa myosin motors, which are presumably located in the insertional plaque and run on the surface of the stereociliary actin filaments as guides.[94,95,98]

The discoveries of the acoustic emission[99] and the electromotor activity of the dissociated outer hair cells in mammals[100,101] led to the concept of active hearing. According to this concept, amplification of auditory sensitivity and sharp tuning of the basilar membrane frequency response are the result of motor activity by the outer hair cell. There is strong evidence that this motor activity is driven by multiple molecular motors, which are an integral part of the cell membrane.[102,103]

An exciting development is the identification of prestin as a candidate for the outer hair cell motor protein. This molecule, related to the anion transporter pendrid, was one of several identified in a differential screen of cDNAs performed by subtractive PCR between cDNAs derived from outer versus inner hair cells. A full-length cDNA produced membrane motility when transfected into an unrelated cell type, identifying this molecule as a strong candidate for the outer hair cell membrane motor protein.[104] Supporting the evolutionary relationship between prestin and an anion transporter is the observations that the intracellular anions chloride and bicarbonate are required for prestin's voltage sensitivity.[105,106]

Neurotransmitters Afferent and efferent (medial and lateral) systems and the innervation patterns of the auditory organ are relatively well established. Although remarkable progress has been made in recent years, our knowledge concerning inner ear neurochemistry is yet to be completed. Afferent neurotransmission between the cochlear hair cells and spiral ganglion neurons is mediated by an as yet unidentified neurotransmitter. Based on molecular analysis evidence, the strongest candidate for this transmitter is glutamate or a related amino acid.[107] The expression of mRNAs encoding glutamate receptors has been detected in spiral ganglion neurons. These include members of the AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionate),[108] NMDA, and kainate[108,109] glutamate receptor families. These observations strongly support the hypothesis that an excitatory amino acid like glutamate is one of the cochlear afferent neurotransmitters. The primary efferent neurotransmitter between the brainstem and cochlea is thought to be acetylcholine, with other transmitters, including γ-aminobutyric acid (GABA), opioid and other peptides, and possibly adenosine triphosphate (ATP), also being involved. Members of several neuronal receptor gene families have been shown to be expressed in the cochlea. Expression of genes encoding nicotinic,[110] muscarinic,[111] and metabotropic[112] acetylcholine receptors; GABA receptors[113]; and ATP receptors[114] has been documented in the cochlea, whereas preproenkephalin mRNA has been found in cochlear and vestibular efferent cell bodies in the brainstem.[115] In particular, the α9 member of the nicotinic acetylcholine receptor has been shown to be strongly expressed in cochlear hair cells and to match exactly the pharmacology of the outer hair cell response to this transmitter.[116] In general, the diversity of receptor and transmitter expression associated with the efferent system matches the complex pharmacology of the inner ear.

Inner Ear Fluid Regulation Differences in the ionic composition between endolymph and perilymph have been shown to play a critical role in inner ear function. It is, therefore, not surprising that the expression of genes encoding a number of ion transport enzymes[117,118] and ion channels[119] has been found in the inner ear. Of particular interest are the Na, K-ATPase genes since this enzyme appears to be the most important determinant of the composition of perilymph and endolymph. In the stria vascularis, the α_1 and β_2 isoforms are the only ones expressed. [117] This combination is not found in isolation in any other tissue in the body. However, the β_2 subunit is associated with transporting sodium against a high electrochemical gradient. In order for the stria vascularis to transport sodium against the highest gradient in the body, perhaps this unique composition of the enzyme is required. Several additional proteins that appear to be involved in circulation of K^+ ions from the organ of Corti to the stria vascularis have been identified via natural or induced mutations.[120] These include connexin 26 expressed in cells of the organ of Corti,[44] a Na-K-Cl cotransporter expressed in basal cells of the stria vascularis,[121] and an ISK potassium channel expressed by marginal cells.[122]

Inner Ear Proteins

Genes important for inner ear function have often been highlighted by mutational analysis. An additional paradigm for identifying functionally important proteins is to look for genes whose pattern of expression is limited to the inner ear and to test if these genes are important for inner ear function, creating animal models with nonfunctional mutant versions of these genes. To date, very few genes showing a pattern of expression limited to the inner ear have been reported. The organ of Corti–specific proteins OCP I and OCP II have been identified,[123] and the OCP II gene has been cloned.[124] Tectorial membrane proteins (tectorins) with little homology to known proteins have also been cloned.[125]

An inner ear–specific novel structural protein in the sunfish saccule has been cloned and characterized.[126] A cDNA, obtained by differential screening of a saccular cDNA library, that encodes an inner ear–specific collagen molecule was identified. The predicted amino acid sequence of this protein showed 40% identity and 56% overall homology with collagen types VIII and X.[127] In situ hybridization with an antisense saccular collagen cRNA showed that transcripts encoding this protein are localized only to the edge of the saccular epithelium, indicating that these specialized epithelial (supporting) cells may secrete this protein, which is believed to be a component of the otolithic membrane.[126]

Much of the molecular data generated to date has involved identification of novel gene sequences for structural proteins and has not been well integrated into the functions of the inner ear. Future studies will be needed to test the importance of these genes and others for normal inner ear function.

Gene Expression in Auditory Tissues

Several laboratories have generated cDNA libraries from specific tissues or cell types of the inner ear, such as outer hair cells,[127] and the National Institute on Deafness and Other Communication Disorders (NIDCD) is currently sponsoring the generation of cDNA libraries from auditory tissues such as the otocyst. Differential screening techniques, such as subtraction, differential display, and suppressive subtractive hybridization (SSH), are currently being employed by a number of laboratories in an attempt to identify genes specifically expressed in different inner ear cell types.[104] Efforts to use gene arrays to provide broad profiles of gene expression in auditory tissues are under way, including efforts sponsored by the NIDCD. A hair cell–specific promoter has been isolated and used to direct expression of green fluorescent protein (GFP) limited to hair cells.[128] Using homologous recombination, Zuo et al introduced GFP into the α9 acetylcholine receptor locus in a BAC, which was used to create a transgenic mouse that showed hair cell–specific expression.[129] Analysis of these and other tissue-specific promoters will help to define the nuclear signals that direct expression to auditory cell types and allow the development of tissue-specific viral vectors for the auditory system. It will also permit the production of other proteins, or antisense RNA for suppression of targeted proteins, as well as cell-specific gene deletions using CRE-LOX recombinant and similar methodologies, in auditory cells.

GENE THERAPY IN ANIMAL MODELS

The use of gene therapy for disorders of the ear remains a prospect for the future. However, research on this topic gives promise to the idea that at least some inner ear conditions might be treated using molecular biologic methodologies.

For example, mutations that delete a critical gene can, in some instances, be corrected by supplying the missing gene. Yoo et al studied the *shiverer* mouse, which has a neurologic disorder caused by a mutation in the gene encoding myelin basic protein and characterized by defects in the myelination of nerve tracts.[130] These mice show deficits in both the amplitude and latency of auditory brainstem responses (ABRs) owing to abnormal conduction in auditory nerve fibers. When a transgene encoding normal myelin basic protein was integrated into the genome of *shiverer* mice, the ABR deficits were partially corrected. A similar approach, combined with in vitro fertilization, might someday be used to prevent certain devastating genetic disorders in humans. Transgenic mice carrying a bacterial antibiotic resistance gene were found to be resistant to ototoxicity.[131] Molecular studies of otologic disease and gene therapy research are both areas in which there is considerable interaction between basic scientists and clinicians.

A more practical form of therapy is one that can be used in the adult organism. Several laboratories are studying methods for gene delivery into the tissues of the inner ear, either for investigative purposes or as the basis for eventual gene therapy. Several viral vectors carrying reporter genes under the control of constitutive promoters have been shown to mediate the transduction of inner ear cells, both in vivo[132,133] and in vitro.[134] These studies have generally found that transduction of hair cells is less effective than that of other cell types, such as Schwann cells or cells lining the perilymphatic spaces. Vectors have also been used to deliver growth factors, resulting in the rescue of damaged inner ear ganglion neurons[135] and hair cells.[136] Ryan et al explored different molecular biologic methods to deliver a factor into the inner ear.[137] They injected

into the adult mouse inner ear fibroblasts that had been genetically engineered to secrete high levels of acidic FGF. To suppress proliferation, the fibroblasts were irradiated, and the researchers found that the irradiated cells were well tolerated in the cochlea, with few side effects, and continued to produce the growth factor for several weeks. Because the ganglion cells are bathed with fluid freely communicating with perilymph, growth factors secreted by the fibroblast in the spiral ligament will be available to the ganglion cells. Staecker et al also found that inner ear injections of fibroblasts engineered to produce BDNF promoted survival of spiral ganglion neurons after hair cell destruction in guinea pigs.[138] However, the use of nonirradiated cells in this study led to destructive overproliferation. Although preliminary, these studies demonstrate theoretical possibilities of molecular genetically based therapy for diseases in the human inner ear.

MOLECULAR PATHOGENESIS OF OTITIS MEDIA

In the middle ear, molecular methods are being applied in diagnosis, research, and experimental treatments for disease. Polymerase chain reaction has proven to be extremely useful in the identification of microbes in OM,[139] and quantitative PCR has been used to measure cytokine production.[140] Genetics has been shown to be an important factor in OM susceptibility.[141] Mechanisms of mucosal proliferation and immunoregulation have been probed using a variety of molecular techniques, including in situ hybridization[142] and the implantation of genetically modified cells.[137] Extraction of RNA from human archival temporal bones may also prove to be useful for the analysis of such samples for the expression of genes involved in the inflammatory process.[143]

Purification and sequencing of gene products are important in understanding the pathogenesis of OM. Cloning bacterial genes has allowed investigators to produce a number of mutant or genetically engineered bacteria, in which a gene is either deleted or altered. These mutants or genetically engineered bacteria can be used to probe the molecular mechanisms of bacteria–host interaction, virulence factors, and immunogenicity. The genetically engineered bacteria can also be used to mass-produce recombinant outer membrane proteins (OMPs) for vaccine candidates. Although not yet fully explored, transgenic and gene knockout animals are being developed for the study of host defense mechanisms against infection.

Bacterial Adherence

To be successful pathogens on a mucous membrane, bacteria must be able to colonize, evade host defense, and express toxins that damage host cells and tissues. To accomplish these goals, the bacteria must be able to adapt to the new environment by altering their phenotypic expression through transcriptional regulation. Although the exact mechanisms by which bacteria attach to the mucosal surface are poorly understood, it is generally believed that receptor ligand–mediated bacterial binding plays a major role, even though nonreceptor ligand binding may also play a role in the adherence. The host receptor for *Streptococcus pneumoniae* is now known to be Glc Nac1-3Gal,[144] whereas the host receptor for nontypable *Haemophilus influenzae* (NTHi) is not yet fully characterized. However, several of the bacterial ligands of NTHi have been described. The suggested ligands include 22 kDa pilin,[145] 27.5 kDa LKP1 (a single serotype) pilin,[146] 36.4 kDa fimbrin,[147,148] and 120 kDa high-molecular-weight OMPs (HMW-1, HMW-2),[149] which may mediate bacterial binding to the host epithelial cell surface. Pilin is a subunit of the rigid, tubular surface appendage pilus and mediates hemagglutination. Fimbrin is a subunit of the surface appendage fimbria, which is a nontubular filament, thinner than pilus and nonhemagglutinating.[146,148]

The gene encoding LKP1 pilus (27.5 kDa) from an NTHi strain obtained from a middle ear effusion has been cloned and expressed in *Escherichia coli,* and the recombinant pili were capable of mediating both the binding of bacteria to the buccal mucosa and hemagglutination.[146] Coleman et al cloned and sequenced the gene encoding the 22 kDa hemagglutinating pilin from another strain of NTHi, found the DNA sequence to be 77% identical to that of the *H. influenzae* type b pilin gene, and showed it to be 68% homologous via the derived amino acid sequence.[145] Because only less than 5% of middle ear isolates and 35% nasopharyngeal isolates expressed pili,[150] the precise role of the pili in the pathogenesis of OM is unclear.

Although it has not yet been established that fimbrial expression is mandatory for the establishment of infection, there is good evidence indicating

that 100% of NTHi isolates from the middle ear of patients with chronic OM with effusion are fimbriated.[151] Sirakova et al sequenced and cloned a 36.4 kDa fimbrial protein from another strain of NTHi recovered from middle ear effusion.[148] The translated amino acid sequence of fimbrin was found to be homologous with various members of the outer membrane protein A (OMP A) family of proteins of other gram-negative bacteria[152] and with type b *H. influenzae* OMP P5, with which it showed 92% identity. Disruption of fimbrin gene resulted in a mutant lacking fimbrial expression and loss of immunoreactivity to the antisera directed against isolated fimbrial proteins.[148] Importantly, this isogenic mutant showed reduced adherence to human oropharyngeal cells in vitro and significantly reduced induction of OM through intranasal inoculation of NTHi in the chinchilla animal model. Thus, fimbrin not only causes adhesin-mediating bacterial adherence but is also a virulence factor for NTHi.[148]

Although the exact mechanism or role of the fimbria in the pathogenesis of OM is not clear, one possibility is that when fimbrin is expressed, it may help to establish an intimate association between the bacteria and host cell.[151] It is possible that through such an interaction, bacterial toxins can be delivered to the host cell, causing cell injury and mediating inflammatory cell responses.

Virulence Factors

Endotoxin plays an important role in the pathogenesis of OM.[153] Evidence indicates that phase variation of the endotoxin lipo-oligosaccharide (LOS) is a virulence factor in type b *H. influenzae*.[154] Phase variation is a reversible switch between two stable genotypes. Phase variation provides bacteria with an advantage in evading host immune response and becoming successful pathogens. The genes responsible for phase variation have been identified, and three *lic* genes (*LIC1, LIC2,* and *LIC3*) are known to be responsible for variable translation of the *lic* loci, which enables a number of different LOS structures to be produced from a limited set of genes.[154,155] Although the loss of phase variation did not affect bacterial colonization, it has affected the ability to invade the bloodstream from the respiratory epithelium when the genes are inactivated.[156] It is possible that similar phase variation of LOS may also exist in NTHi.

Pneumococcal proteins pneumolysin and autolysin contribute significantly to the virulence of pneumococci.[157] *Streptococcus pneumoniae* type 3 mutants deficient in the production of either pneumolysin or autolysin were constructed by transformation of DNA from derivatives of a rough strain, in which the respective genes had been interrupted by insertion-duplication mutagenesis using internal fragments of the cloned genes in the vector.[157] Both the pneumolysin-negative and the autolysin-negative strains had significantly reduced virulence in mice, as judged by survival time after intraperitoneal challenge. When pneumolysin and autolysin productions are restored by back-transformation of the mutants with an intact copy of the respective cloned gene, the mean survival time was indistinguishable from that of mice challenged with the wild-type strain. Thus, pneumolysin and autolysin are important virulence factors in type 3 pneumococci.[157] Mice were challenged with a genetically modified mutant strain of pneumococcus, which was unable to express active pneumolysin. Preimmunization of such mice with autolysin failed to provide any significant protection against the challenge.[158] The authors suggested that the most important contribution made by autolysin to the virulence of *S. pneumoniae* may be its role in mediating the release of pneumolysin from the pneumococcal cytoplasm during infection.

MOLECULAR BIOLOGIC APPROACHES IN THE DEVELOPMENT OF VACCINES AGAINST OTITIS MEDIA

The intention of this section is not meant to be a review on current OM vaccines; rather, it is intended to focus on the molecular biologic aspects of vaccine development. Excellent reviews on the OM vaccine development are available elsewhere.[159,160]

Nontypable *Haemophilus influenzae* Vaccines

Several lines of evidence suggest that a number of OMPs of NTHi are potential candidates for vaccines against this microorganism. The genes encoding several of these OMPs have been cloned, sequenced, and expressed.[161–163] This nucleotide sequence information, together with mapping of bactericidal epi-

tope molecules, enabled investigators to define the structure and function of these OMPs and identify domains that are conserved across the strains.[164,165] Such information is critical in selecting antigens as vaccine candidates. Among those identified as potential vaccine candidates, NTHi OMP P6 (peptidoglycan-associated lipoprotein), which is common to both nontypable and type b *Haemophilus*, is the most promising because P6 is a target of bactericidal antibodies in convalescent sera. A truncated P6 gene without the signal peptide has been constructed by Green and his associates,[166] and the recombinant protein encoded by this construct is capable of eliciting bactericidal antibody in vitro. Using a monoclonal antibody, Murphy and Kirkham identified a region of P6 encoding an epitope recognized by bactericidal antibody.[167] Another OMP, P26, also shows conservation across strains of NTHi and a recombinant preprotein-induced protection in animals.[168]

Another *Haemophilus* OMP known as P6 cross-reactive protein (PCP) is also considered to be an excellent vaccine candidate because polyclonal antisera against PCP are bactericidal. The PCP is antigenically conserved and present in both type b and NTHi. Deich et al recently cloned and sequenced this protein.[169]

Other *Haemophilus* OMPs, believed to mediate bacterial adherence and also considered as potential vaccine candidates, include pilin and fimbrin. Antibody response to such proteins by immunization can block the bacterial adherence mediated by these proteins, thus reducing colonization and infection.[148,150] Another vaccine candidate is a group of high-molecular-weight surface-exposed proteins of NTHi, which are related to the filamentous hemagglutinin protein of *Bordetella pertussis*. The high-molecular-weight proteins are known to be critical adhesion molecules but also are major targets of human serum antibody.[170] To further characterize these proteins, Barenkamp and Leninger cloned and sequenced genes encoding two related high-molecular-weight proteins (120 kDa and 125 kDa).[171]

Streptococcus pneumoniae Vaccines

Many new vaccines including multivalent polysaccharide vaccines and polyvalent pneumococcal conjugate vaccines are under development and being field-tested, and they will not be covered here.

Pneumococcal proteins autolysin and pneumolysin contribute significantly to the virulence of this microorganism. Autolysin and a defined toxoid derivative of pneumolysin were tested for efficacy in a mouse model as antigens protecting against challenge with virulent, wild-type *S. pneumoniae,* and the result showed that each antigen alone provided significant protection.[158] When mice were challenged with a genetically modified mutant strain of pneumococcus unable to express active pneumolysin, preimmunization of mice with autolysin failed to provide any significant protection against the challenge.[158] When mice were immunized with a genetically engineered toxoid version of pneumolysin derived from serotype 2 pneumococcus and challenged with 12 strains of pneumococci covering capsular types 1 to 6, 7F, 8, and 18c, pneumolysin toxoid conferred protection against all nine pneumococcal serotypes.[172] Thus, these investigators suggested that pneumolysin toxoids warrant consideration for inclusion in pneumococcal vaccines. Molecular analysis of pneumococcal surface protein A also reveals several relatively conserved domains,[173] suggesting that this molecule may also be a vaccine candidate.

DNA Vaccines against Respiratory Viruses

Acute OM is frequently preceded by a viral upper respiratory infection. Recent data using PCR demonstrated that a high number of middle ear effusions contained evidence of upper respiratory viruses,[174] and upper respiratory viral infection is causally related to the bacterial OM. It has been shown that attenuated influenza viral vaccine conferred protection against pneumococcal OM in the chinchilla.[175] Thus, preventing upper respiratory viral infection will reduce the incidence of acute OM. In recent years, clinical trials have been initiated with live attenuated vaccines for respiratory syncytial virus, influenza, and parainfluenza viruses and adenovirus.[159] Such trials can provide an excellent opportunity to test the above hypothesis. In a clinical trial in Finland, attenuated influenza virus was administered to children under 4 years old. During a 6-week epidemic, 83% of the children were protected from confirmed influenza A–associated acute OM and 36% from acute OM in general.[176] However, efficacies of viral vaccines against disease in susceptible populations are modest at best owing to rapid

changes of viral surface antigens and heterogeneity of surface antigens among different strains, as well as incomplete knowledge of the characteristics of protective immunity of different upper respiratory viruses.[159]

Most of the current viral and bacterial vaccines are largely using surface antigens. However, such an approach poses a serious problem because a microbe's surface structures frequently change at a rapid rate, rendering vaccines useless. For example, evolution of the genes for surface proteins of influenza virus occurs so fast that a new vaccine must be made every year to be effective. However, proteins in the interior of viruses are more highly conserved (stable) than those of the surface. The discovery that naked DNA from a virus injected into the body could function as a vaccine may provide a new opportunity to develop an effective viral vaccine.[177] Futhermore, investigators described the results of experiments on mice using DNA from influenza virus.[178] The opponents of viral DNA vaccines caution a theoretical possibility of activation of an oncogene, which may induce cancer in the host.

MOLECULAR MECHANISMS OF ANTIBIOTIC RESISTANCE AMONG OTITIS MEDIA PATHOGENS

Antibiotic Resistance in *Haemophilus influenzae*

The genetic basis of ampicillin and chloramphenicol resistance in *H. influenzae* has been well characterized, and the principal mechanism of the resistance is largely enzymatic.[179] The gene that encodes the β-lactamase enzyme is located within a large gene segment, transposon A.[180] Mendelman et al reported a small number of strains that demonstrate ampicillin resistance in the absence of β-lactamase production, which is believed to be mediated by alterations in one or more penicillin-binding proteins (PBPs).[181] The PBPs 3a and 3b showed a decrease in affinity for β-lactams in antibiotic-resistant strains.[182] The gene for the altered PBPs has been cloned.[183] The genetic basis for resistance of *H. influenzae* type b to chloramphenicol has been elucidated, and it is believed to be caused largely by the enzyme acetyltransferase. The genetic sequence that codes for this enzyme is also located within a transposon.[184] Molecular cloning and mechanisms of trimethoprim resistance in *H. influenzae* were elucidated by cloning the gene for trimethoprim resistance into a cosmid vector and transducing recombinant plasmids into *E. coli*.[185] The results indicated that the acquisition of trimethoprim resistance involved a chromosomally mediated rearrangement or change of nucleotide sequences. The mechanism of trimethoprim resistance is overproduction of dihydrofolate reductase.[185]

Antibiotic Resistance in *Streptococcus pneumoniae*

Penicillin resistance to *S. pneumoniae* is believed to be mediated by changes in the production of altered PBPs.[186] The PBPs la, 2b, and 2x have been cloned and sequenced.[186–188] No β-lactamase-mediated penicillin resistance was reported in pneumococci. Dowson and his associates investigated PBP 2b from several sensitive and resistant pneumococci and demonstrated a major change in the carboxyl-terminal sequence.[187] It is postulated that this change is responsible for the development of penicillin resistance.

MOLECULAR EPIDEMIOLOGY AND DIAGNOSIS

Molecular biologic techniques have been used to diagnose bacterial and viral diseases and fingerprint specific microorganisms in epidemiologic surveys.[189] Using a sensitive total genomic DNA restriction fingerprinting method, different isolates of *H. influenzae* have been identified.[190] Restriction endonuclease analysis of bacterial chromosomal DNA was used to compare NTHi obtained from middle ear effusions with that from nasopharynges of patients with OM. The restriction digest profiles of strains isolated simultaneously from the middle ear effusion and nasopharynx of an individual child were identical, whereas the restriction profiles of strains isolated from different children were different from one another.[191] In this study, the isolates from recurring episodes of NTHi in six children were different from those that caused the initial episode.

Like in the NTHi, another gram-negative bacteria, *Branhamella catarrhalis*, was subject to restriction fragment mapping for epidemiologic study for this newly emerging pathogen for OM. Dickinson et

al were able to fingerprint and compare *B. catarrhalis* isolates obtained from middle ear effusions and those from nasopharynges of the same patients using restriction fragment mapping following digestion of genomic DNA with restriction endonuclease such as *Pst*1 and *Cla*1.[192] This microorganism was considered as a normal flora of the nasopharynges of healthy children as well as adults. However, it has become increasingly associated with acute purulent OM and OM with effusion in recent years. Therefore, it is important to learn about strains that are pathogens versus nonpathogens, and the restriction fragment mapping technique allows investigation of the epidemiology of this microorganism reliably.

Polymerase chain reaction assay based on the amplification of pneumolysin gene fragments in sera was developed to diagnose acute pneumococcal pneumonia.[193] Analysis of DNA restriction patterns of genomic DNA provided a sensitive measure of genetic similarity between strains and a convenient method for use in epidemiologic studies.

Response of the Middle Ear Mucosa to Otitis Media

The middle ear mucosa exhibits dramatic responses to OM. It is unique among mucosae in that it has the capacity for extreme hyperplasia when stimulated by inflammatory events. The simple squamous cell epithelium that lines much of the normal middle ear can rapidly proliferate and differentiate into a respiratory mucosa with a well-organized and vascularized stroma and a pseudocolumnar epithelium with ciliated and goblet cells.

Growth Factors and Signal Transduction Extracellular stimuli induce both physiologic and pathologic cellular responses via activation of signal transduction pathways extending from receptors on the cell surface to the transcription factors in the nucleus. Therefore, investigating these signal transduction mechanisms not only provides new insights into molecular mechanisms of physiologic responses, development, and pathogenesis but may also open up novel therapeutic targets for the treatment of ear diseases such as OM.

Progress has been made recently in the studies of signal transduction mechanisms involved in the pathogenesis of OM. Oehl and Ryan used reverse transcriptase PCR to examine the expression of vascular endothelial growth factors (VEGFs) and their receptors (VEGF-Rs), important receptor tyrosine kinases involved in normal and pathologic angiogenesis, and found that VEGF, VEGF-C, VEGF-D, and VEGFR-2 (flk-1) were up-regulated in the mucosa of rat middle ears inoculated with bacteria.[194] DeFoire-Hyrmer and Bakaletz provided evidence to link phosphorylation of tyrosine with NTHi adherence and actin nucleation.[195] Sudhoff et al documented the presence of angiogenic growth factors and receptors in cholesteatoma.[196] Using immunocytochemisry, Huang et al found epidermal growth factor, FGF, and platelet-derived growth factor in cholesteatoma keratinocytes,[197] whereas Ishibashi et al detected keratinocyte growth factor and receptor mRNA.[198] Huang and colleagues found evidence that ras, c-jun, and p53 may be involved in signaling from these receptors.[197,199] Schilling et al observed high levels of tenascin in proliferating cholesteatoma keratinocytes, suggesting a role for integrin signaling.[200] In addition, the role of the signaling molecule phospholipase C (PLC)-γ-1 was explored by Park et al.[201] Their studies suggested a possible involvement of PLC-γ-1 in cholesteatoma pathogenesis.

Expression of interleukins has been documented by several investigators. In general, genes encoding proinflammatory cytokines are expressed very rapidly, within 6 hours of middle ear inoculation with bacteria, and decline by 24 hours.[140,202] Anti-inflammatory and immunoregulatory cytokines are expressed over a much longer time course, with some lasting for weeks.[203,204]

John and Nam reported that the concentration of nitric oxide (NO), an important intracellular second messenger, is higher in middle ear effusion.[205] In experimental chronic OM, endothelial nitric oxide synthase (eNOS) mRNA is increased and inducible nitric oxide synthase (iNOS) gene expression is dramatically up-regulated.[142] In addition to the role of NO in the pathogenesis of OM, Juhn et al provided evidence indicating that NO may be involved in the cochlear transduction process.[206]

The role of mitogen-activated protein (MAP) kinases in the pathogenesis of OM has been explored. Li et al investigated the signal transduction mechanisms involved in *H. influenzae*–induced mucin up-regulation and found that *H. influenzae* up-regulates mucin gene transcription via activation of a Rac-dependent MEKK-SEK-p38 MAP kinase

pathway, suggesting that *H. influenzae* up-regulates mucin transcription via activation of an intracellular signaling pathway different from *Pseudomonas aeruginosa*.[207,208] In addition to the involvement of p38, Xu et al reported that oxygen radical species (ROS) may be involved in activation of p38 MAP kinase, which, in turn, led to the activation of mucin transcription.[209] Chun et al also showed that MAP kinase ERK (extracellular signal regulated) is involved in IL-1-induced down-regulation of surfactant protein B in the middle ear.[210]

Transcription factor nuclear factor kappa B (NF-κB) has been shown to play an important role in inflammatory responses. One possible involvement of NF-κB is to regulate the transcription of proinflammatory cytokines. It has been shown that activation of NF-κB is involved in IL-8 gene expression in human adenoidal fibroblasts.

Mucin Gene Expression in the Middle Ear and Up-regulation of Mucin in Otitis Media Mucins are high-molecular-weight glycoproteins that constitute the major component of mucus in the middle ear, trachea, and digestive and reproductive tracts. They protect epithelia and trap particulates, including bacteria, for mucociliary clearance. In diseases such as OM, chronic bronchitis, and cystic fibrosis, mucin production is up-regulated, which, in turn, contributes to the pathogenesis of diseases. In OM with effusion, overproduced mucin is believed to play an important role in causing conductive hearing loss as well as defective mucociliary clearance and recurrent infection.

Progress has been made in the study of mucin gene expression in the middle ear and eustachian tube. Chun et al examined human mucin gene expression and found that mucin *MUC2, MUC5AC*, and *MUC5B* genes are expressed in human middle ear epithelial cells.[210] A similar result was found in human primary cultures of middle ear epithelial cells by Moon et al.[211] In addition to mucin gene expression in the middle ear, Lin et al also reported the expression of glycoconjugates in human eustachian tubes.[212]

More attention has been given to determining mucin expression in OM middle ear mucosa and effusions. Hutton et al isolated and purified high-molecular-weight mucins from effusions of children with OM with effusion and provided direct evidence for expression of different mucin gene products.[213] In consistence with this study, Jung et al reported that human *MUC5AC, MUC5B, MUC7*, and *MUC8* are all expressed in the middle ear mucosa of patients with chronic OM.[214] In the study reported by Severn et al, human *MUC1* and *MUC2*, in addition to *MUC5AC* and *MUC5B*, were also found to be expressed in middle ear mucosa.[215]

Efforts have also been made toward understanding the molecular mechanisms controlling mucin overproduction. Wang et al investigated the role of *H. influenzae* in mucin up-regulation and found that *H. influenzae* up-regulates mucin *MUC5AC* gene transcription via activation of a p38 MAP kinase pathway, suggesting that *H. influenzae* up-regulates mucin transcription via activation of intracellular signaling pathway different from *P. aeruginosa*.[207,216] Interestingly, in addition to the involvement of p38 in mucin up-regulation, Shuto et al showed that activation of p38 also plays an important role in inflammatory responses.[217] The role of proinflammatory cytokines in mucin regulation has also been studied by Lin et al.[212] Their studies showed that tumor necrosis factor-alpha (TNF-α) up-regulates mucin *MUC2* gene expression in the middle ear of rats. All of these studies suggest the complexity of regulatory mechanisms underlying mucin up-regulation in OM.

Epithelial Antimicrobial Peptides and Proteins in the Pathogenesis of Otitis Media It is now believed that in addition to the components of the adaptive immune system, the homeostasis of the nasopharyngeal tract, eustachian tube, and middle ear is also maintained by epithelial antimicrobial proteins and peptides that function to defend these tissues against microbial invasion. The larger antimicrobial proteins are often lytic enzymes, nutrient-binding proteins, or proteins containing sites that target specific microbial macromolecules. The smaller antimicrobial peptides (defined as peptides containing fewer than 100 amino acids) act, at least in part, by disrupting the structure or function of microbial cell membranes. Included in the molecules of the innate immune system are the defensins, lysozyme, lactoferrin, and members of the collectin family, including the surfactant proteins A and D.[218]

The defensins are cationic (polar) molecules with spatially separated hydrophobic and charged regions. In vitro, the defensins (at micromolar con-

centrations) have a broad spectrum of antimicrobial activity against bacteria, fungi, and even some enveloped viruses.[219] In humans and other vertebrates, the defensins can be divided into the α- and β-defensin subfamilies. The α-defensins are produced by neutrophils and intestinal Paneth's cells.[220] The β-defensins, on the other hand, are mainly produced by epithelial cells of the skin, kidneys, and tracheabronchial lining of nearly all vertebrates.[221] The β-defensins are released on microbial invasion and are located at the host–environment interface, such as mucosal surfaces and skin. Two types of human β-defensins, β-defensins 1 (HBD-1) and 2 (HBD-2), have been identified. Human β-defensin 2 is produced by epithelial cells and exhibits potent antimicrobial activity against gram-negative bacteria and *Candida.* But it is not as effective against gram-positive *Staphylococcus aureus.*[222,223] Human β-defensin represents the first human defensin that is produced following stimulation of epithelial cells by contact with microorganisms or cytokines such as TNF-α and IL-1α. Human β-defensin 2 functions as a NF-κB target gene in the intestinal epithelium, but HBD-1 is not affected by IL-1α and other proinflammatory stimuli.[221] The presence of HBD-1 has been reported in the pars tensa and pars flaccida of the tympanic membrane and in the meatal skin.[224] Recently, Lim et al demonstrated that both HBD-1 and HBD-2 are expressed in human middle ear epithelial cells and that both molecules have bactericidal/bacteriostatic activity against NTHi.[218]

Lysozyme is an important component of innate immunity against pathogens at mucosal surfaces. Human chronic middle effusions contain high levels of this molecule, which is produced by secretory cells of the middle ear mucosal epithelia, as well as by polymorphonuclear leukocytes and macrophages.[218,226] Lim and coworkers have shown that human milk–derived lysozyme, in combination with other innate immune molecules, exhibits antibacterial activity against NTHi.[218]

Lactoferrin is an iron-binding glycoprotein found in the milk and exocrine secretions of mammals that is released from neutrophilic granules during inflammation.[227,228] Lactoferrin has also been detected in middle ear effusion and has been localized to the serous cells of eustachian tube glands, as well as the cuboidal epithelium of the transition zone of the middle ear.[229,230]

The collectins are proteins that are related structurally and functionally to the first component of the classic complement pathway C1q and are known to play significant roles in innate immunity through opsonization and complement activation.[231,232] The collectin family of proteins includes the mannose binding protein, conglutinin, CL-43, and surfactant proteins A and D. Surfactant proteins A and D are expressed in the middle ear and eustachian tube and are likely to be involved in the protection of the tubotympanum against OM pathogens.[218]

Gene Transfer to the Middle Ear Mucosa Ryan et al successfully implanted cells that had been stably transformed to secrete FGF-1 into the subepithelial compartment.[137] Mondain et al demonstrated that an adenovirus vector can be used to transduce cells of the adult guinea pig middle ear mucosa.[233] Similarly, Dazert et al demonstrated adenovirus transduction of mucosal cells in the neonatal rat.[133] These studies demonstrate the feasibility of gene therapy in the middle ear cavity for the treatment of OM.

FUTURE DIRECTIONS

The explosive pace of development of molecular techniques does not allow us to do justice in this chapter to all aspects of molecular biology that are relevant to otology. We have just begun to scratch the surface of many important areas of biology of the ear with these new tools. The availability of such powerful techniques is already beginning to impact our understanding of cellular functions and malfunctions on a molecular level. Identification of genes important for the formation of the ear and the maintenance and normal function of the sensory and nonsensory cells of the inner ear are now possible, and discovery of many new genes of importance will be accelerated. Many more transgenic animals and gene knockout animals of interest will be made available, and these animals will propel our knowledge concerning specific gene functions important for the ear.

Although only limited attempts have been made, many innovations for gene or gene-based therapy will allow scientists to develop new approaches (eg, antisense oligonucleotide therapy) to prevent or cure human diseases, which currently cannot be ameliorated.

For OM, molecular biologic techniques already have made a major impact on our understanding of bacterial pathogenesis at the molecular level. We are beginning to understand the molecular mechanisms involved in bacteria–host interaction. A number of OM-causing bacterial surface antigens important for the protective host immune response have now been identified, and their genes are cloned. Molecular methods will allow us to incorporate this new knowledge in developing strategies to prevent and treat OM. The molecular biology of host responses, including the innate and acquired immune responses, cytokine and β-defensin regulation, inflammatory cell mobilization, B and T cell activation, tissue proliferation, and mucin regulation, constitutes a large body of knowledge that is relevant for understanding the pathogenesis and progression of OM. Rapid progress is being made in these areas, including identification of genes regulating various aspects of host responses, which will impact the way we manage OM in the future.

In summary, molecular biologic techniques are no longer an esoteric scientific curiosity; rather, they are beginning to be used in diagnosis, new vaccine development, and even gene therapy. Many more molecular biology–based innovative approaches will become available in the future, which will ultimately change our understanding and treatment of these diseases.

REFERENCES

1. Saiki RK, Gelrand DH, Stoffel S, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 1988;239:487–91.
2. Simmons DM, Arriza JL, Swanson LW. A complete protocol for in situ hybridization of messenger RNAs in brain and other tissues with radiolabeled single-stranded RNA probes. J Histotechnol 1989;12: 169–81.
3. Battey JF, Fathi S, Wada E, et al. Molecular genetic approaches to the analysis of neuropeptide receptors. Cell Physiol Biochem 1993;3:3–230.
4. Hoshi T, Zagotta WN, Aldrich RW. Biophysical and molecular mechanisms of Shaker potassium channel inactivation. Science 1990;250:533–8.
5. Zagotta WN, Hoshi T, Aldrich RW. Restoration of inactivation of mutants of Shaker potassium channels by a peptide derived from ShB. Science 1990; 250:568–71.
6. Fambrough DM, Lemas MV, Hamrick M, et al. Analysis of subunit assembly of Na, K-ATPase. Am J Physiol 1994;6:579–89.
7. Siddique M, Goswami S, Qasba P, et al. Regulation of muscle tissue specific transcription: analysis of chicken cardiac myosin light chain 2 gene promoter. Proceedings of the 1st World Congress of Cellular Molecular Biology 1991. p. 345.
8. Gilbert SF. Genes and development: introduction and techniques; development biology. 4th ed. Sunderland (MA): Sinauer Associates, Inc; 1994. p. 34–76.
9. Rosenfeld M, Crenshaw E, Lira S, Swanson L. Transgenic mice: applications to the study of the nervous system. Ann Rev Neurosci 1988;17:353–72.
10. Gridley T, Soriano P, Jaenish R. Insertional mutagenesis in mice. Trends Genet 1987;3:162–6.
11. Naori H, Kumura M, Otami H, et al. Transgenic mouse model of hemifacial microsomia: cloning and characterization of insertional mutation region on chromosome 10. Genomics 1994;23:515–9.
12. Xiang X, Benson KF, Chada K. Mini-mouse: disruption of the pygmy locus in a transgenic insertional mutant. Science 1990;247:967–9.
13. Gordon J, Uehlinger J, Dayani N, et al. Analysis of the hotfoot (ho) locus by creation of an insertional mutation in a transgenic mouse. Dev Biol 1990;137: 349–58.
14. Crenshaw EB, Ryan AF, Dillon SR, et al. Wocko, a neurological mutant generated by insertional mutagenesis in a transgenic mouse pedigree. J Neurosci 1991;11:1524–30.
15. Friedman RA, Adir Y, Crenshaw EB, et al. A transgenic insertional inner ear mutation on mouse chromosome 1. Laryngoscope 2000;110:489–96.
16. Ernfors P, Lee K-F, Jaenisch R. Mice lacking brain-derived neurotrophic factor develop with sensory deficits. Nature 1994;368:147–50.
17. Farinas I, Jones KR, Backus C, et al. Severe sensory and sympathetic deficits in mice lacking neurotrophin-3. Nature 1994;369:658–61.
18. Gu W, Schneider JW, Condorelli G, et al. Interaction of myogenic factors and the retinoblastoma protein mediates muscle cell commitment and differentiation. Cell 1993;72:309–24.
19. Boulter J, Connoly J, Deneris E, et al. Functional expression of two neural nicotinic acetylcholine receptors from cDNA clones identifies a gene family. Proc Natl Acad Sci U S A 1987;84:7763–7.

20. Wada K, Ballivet M, Boulter J, et al. Functional expression of a new pharmacological subtype of brain nicotinic acetylcholine receptor. Science 1988; 240:330–4.
21. Boulter J, Hollmann M, O'Shea-Greenfield A, et al. Molecular cloning and functional expression of glutamate receptor subunit genes. Science 1990;249:1033–7.
22. Young RM, Shull GE, Lingrel JB. Multiple mRNAs from rat kidney and brain encode a single Na+, K+-ATPase beta subunit protein. J Biol Chem 1987;262: 4905–10.
23. Hammes A, Oberdorf S, Strehler EE, et al. Differentiation-specific isoform mRNA expression of the calmodulin-dependent plasma membrane Ca(2+)-ATPase. FASEB J 1994;8:428–35.
24. Johnson DE, Williams LT. Structural and functional diversity in the FGF receptor multigene family. Adv Cancer Res 1993;60:1–41.
25. Heinemann S, Bettler B, Boulter J, et al. The glutamate receptor gene family: evidence that the GluRl, GIuR2, GluR3 subunits form multiple glutamate receptors. In: Sidia Research Foundation transmitter amino acid receptors: structures, transduction and models for drug development. London: Royal Society; 1990.
26. Wenthold RJ. Structure and distribution of glutamate receptors: subunit-specific antibody studies. Neurosci Facts 1992;3:2–3.
27. Brown TA. Gene cloning: an introduction. 2nd ed. Norwell (MA): Chapman & Hall; 1990.
28. Farrer LA. Gene localization by linkage analysis. Otolaryngol Clin North Am 1992;25:907–22.
29. Watson JD, Gilman M, Witkowski J, Zoller M. Recombinant DNA. 2nd ed. New York: Scientific American Books; 1992.
30. Lee WH, Bookstein R, Hong F, et al. Human retinoblastoma gene: cloning, identification and sequencing. Science 1987;235:1394–9.
31. Koenig M, Hoffman EP, Bertelson CJ, et al. Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell 1987;50:509–17.
32. Tseng CJ, Lalwani AK. Cracking the auditory genetic code: part II. Syndromic hereditary hearing impairment. Am J Otol 2000;21:437–51.
33. Barker DF, Hostikka SL, Zhou J, et al. Identification of mutations in the COL4A5 collagen gene in Alport syndrome. Science 1990;248S:1224–7.
34. Tryggvason K, Zhou J, Hostikka SL, Shows TB. Molecular genetics of Alport syndrome. Kidney Int 1993; 43:38–44.
35. Tassabehji M, Read AP, Newton VE, et al. Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. Nature 1992;355:635–6.
36. Weil D, Blanchard S, Kaplan J, et al. Defective myosin VIIA gene responsible for Usher syndrome type 1B. Nature 1995;374:60–1.
37. Seizinager BR, Martuza RL, Gusella JF. Loss of genes on chromosome 22 in tumorigenesis of human acoustic neuroma. Nature 1986;322:644–7.
38. Rouleau GA, Wertelecki W, Haines JA, et al. Genetic linkage of bilateral acoustic neurofibromatosis to a DNA marker on chromosome 22. Nature 1987; 329:246–8.
39. Frazer KA, Boehnke M, Budarf ML, et al. A radiation hybrid map of the region on human chromosome 22 containing the neurofibromatosis type 2 location. Genomics 1992;34:574–84.
40. Trofatter JA, MacCollin MM, Rutter JL, et al. A novel moesin-, radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor. Cell 1993;72: 791–800.
41. McCollin M, Ramesh V, Jacoby LB, et al. Mutational analysis of patients with neurofibromatosis 2. Am J Hum Genet 1994;55:314–20.
42. de Kok YJM, van der Maarel SM, Bitner-Glindzicz M, et al. Association between X-linked mixed deafness and mutations in the POU domain gene POU3F4. Science 1995;267:685–8.
43. Ryan AF, Simmons DM, Crenshaw EB. Gene expression in normal and abnormal inner ears. Ann N Y Acad Sci 1991;630:129–32.
44. Carrasquillo MM, Zlotogora J, Barges S, Chakravarti A. Two different connexin 26 mutations in an inbred kindred segregating non-syndromic recessive deafness: implications for genetic studies in isolated populations. Hum Mol Genet 1997;6:2163–72.
45. Denoyelle F, Weil D, Maw MA, et al. Prelingual deafness: high prevalence of a 30delG mutation in the connexin 26 gene. Hum Mol Genet 1997;6:2173–7.
46. Jaber L, Shohat M, Bu X, et al. Sensorineural deafness inherited as a tissue specific mitochondrial disorder. J Med Genet 1992;29:86–90.
47. Prezant TR, Shohat M, Jaber L, et al. Biochemical characterization of a pedigree with mitochondrially inherited deafness. Am J Med Genet 1992;44:465–72.

48. Fischel-Ghodsian N, Prezant TR, Bu X, Oztas S. Mitochondrial ribosomal RNA gene mutation in a patient with sporadic aminoglycoside ototoxicity. Am J Otolaryngol 1993;74:399–403.
49. Verpy E, Leibovici M, Zwaenepoel I, et al. A defect in harmonin, a PDZ domain-containing protein expressed in the inner ear sensory hair cells, underlies Usher syndrome type 1C. Nat Genet 2000;26: 6–7.
50. Wayne S, Robertson NG, DeClau F, et al. Mutations in the transciptional activator Eya4 cause late-onset deafness at the DFNA10 locus. Hum Mol Genet 2001;10:195–200.
51. Abdelhak S, Kalatzis V, Heilig R, et al. A human homologue of the *Drosophila* eyes absent gene underlies branchio-oto renal (BOR) syndrome and identifies a novel gene family. Nat Genet 1997;15:157–64.
52. Xu PX, Adams J, Peters H, et al. Eya1-deficient mice lack ears and kidneys and show abnormal apoptosis of organ primordia. Nat Genet 1999;23:113–7.
53. Hardisty RE, Mburu P, Brown SD. ENU mutagenesis and the search for deafness genes. Br J Audiol 1999;33:279–83.
54. Nolan PM, Peters J, Strivens M, et al. A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. Nat Genet 2000;25:440–3.
55. Probst FJ, Fridell RA, Raphael Y, et al. Correction of deafness in shaker-2 mice by an unconventional myosin in a BAC transgene. Science 1998;280: 1444–7.
56. Mullen LM, Ryan AF. Transgenic mice: genome manipulation and induced mutations. In: Willott J, editor. Handbook of mouse auditory research: from molecular biology to behavior. Boca Raton (FL): CRC Press; 2001. p. 457–74.
57. Kozel PJ, Friedmann RA, Erway L, et al. Balance and hearing deficits in mice with a null mutation in the gene encoding plasma membrane Ca2+-ATPase isoform 2. J Biol Chem 1998;273:18693–6.
58. Minowa O, Ikeda K, Sugitani Y, et al. Altered cochlear fibrocytes in a mouse model of DFN3 nonsyndromic deafness. Science 1999;285:1408–11.
59. Simmler MC, Cohen-Salmon M, El-Amraoui A, et al. Targeted disruption of otog results in deafness and severe imbalance. Nat Genet 2000;24:139–43.
60. Richardson G, Forge A, Kros C, et al. A missense mutation in myosin VIIA prevents aminoglycoside accumulation in early postnatal cochlear hair cells. Ann N Y Acad Sci 1999;884:110–24.
61. Everett LA, Morsli H, Wu DK, Green ED. Expression pattern of the mouse ortholog of the Pendred's syndrome gene (Pds) suggests a key role for pendrin in the inner ear. Proc Natl Acad Sci U S A 1999;96:9727–32.
62. Vahava O, Morell R, Lynch ED, et al. Mutation in transcription factor POU4F3 associated with inherited progressive hearing loss in humans. Science 1998;279:1950–4.
63. Verhoeven K, Van Laer L, Kirschhofer K, et al. Mutations in the human alpha-tectorin gene cause autosomal dominant non-syndromic hearing impairment. Nat Genet 1998;19:60–2.
64. Friedman TB, Sellers JR, Avraham KB. Unconventional myosins and the genetics of hearing loss. Am J Med Genet 1999;89:147–57.
65. Acampora D, Merlo GR, Paleari L, et al. Craniofacial, vestibular and bone defects in mice lacking the Distal-less-related gene Dlx5. Development 1999; 126:3795–809.
66. Morsli H, Tuorto H, Wu DK, et al. Otx1 and Otx2 activities are required for the normal development of the mouse inner ear. Development 1999;26:2335–43.
67. Salminen M, Meyer BI, Bober E, Gruss P. Netrin 1 is required for semicircular canal formation in the mouse inner ear. Development 2000;127:13–22.
68. Granstrom G. Retinoid-induced ear malformations. Otolaryngol Head Neck Surg 1990;103:702–9.
69. Jarvis BL, Johnston MC, Sulik KK. Congenital malformations of the external, middle, and inner ear produced by isotretinoin exposure in mouse embryos. Otolaryngol Head Neck Surg 1990; 102:391–401.
70. Romand R, Sapin V, Dollé P. Spatial distributions of retinoic acid receptor gene transcripts in the prenatal mouse inner ear. J Comp Neurol 1998;393:298–308.
71. Lohnes D, Mark M, Mendelsohn C, et al. Function of the retinoic acid receptors (RARs) during development (I). Craniofacial and skeletal abnormalities in RAR double mutants. Development 1994;120: 2723–48.
72. Oh SH, Wu D. Expression of bone morphogenetic protein 4, 5 and 7 in the developing chick inner ear [abstract 422]. Assoc Res Otolaryngol Mid-Winter Meeting; 1995. p. 6.
73. Winnier G, Blessing M, Labosky PA, Hogan BL. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes Dev 1995;9:2105–16.
74. Teng X, Ahn K, Bove M, et al. Malformations of the lateral semicircular canal occur in heterozygous

Bmp4 knockout mice [abstract]. In: Assoc Res Otolaryngol Mid-Winter Meeting; 2000 Feb 20–24. p. 51.
75. Bermingham NA, Hassan BA, Price SD, et al. Math1: an essential gene for the generation of inner ear hair cells. Science 1999;284:1837–41.
76. Chen P, Segil N. p27(Kip1) links cell proliferation to morphogenesis in the developing organ of Corti. Development 1999;126:1581–90.
77. Lowenheim H, Furness DN, Kil J, et al. Gene disruption of p27(Kip1) allows cell proliferation in the postnatal and adult organ of Corti. Proc Natl Acad Sci U S A 1999;96:4084–8.
78. Erkman L, McEvilly RJ, Luo L, et al. Role of transcription factors Brn-3.1 and Brn-3.2 in auditory and visual system development. Nature 1996;381:603–6.
79. Xiang M, Gao WQ, Hasson T, Shin JJ. Requirement for Brn-3c in maturation and survival, but not in fate determination of inner ear hair cells. Development 1998;125:3935–46.
80. Lanford PJ, Lan Y, Jiang R. Notch signalling pathway mediates hair cell development in mammalian cochlea. Nat Genet 1999;21:289–92.
81. Stone JS, Rubel EW. Delta1 expression during avian hair cell regeneration. Development 1999;126:961–73.
82. Zheng JL, Gao WQ. Overexpression of Math1 induces robust production of extra hair cells in postnatal rat inner ear cultures. Nat Neurosci 2000;3:580–6.
83. Lefebvre PP, Van De Water TR, Weber T, et al. Growth factor interactions in cultures of dissociated adult acoustic ganglia: neuronotrophic effects. Brain Res 1991;567:306–12.
84. Pirvola U, Ylikoski J, Palgi J, et al. Brain-derived neurotrophic factor and neurotrophin 3 mRNAs in the peripheral target fields of developing inner ear ganglia. Proc Natl Acad Sci U S A 1992;89:9915–9.
85. Ylikoski J, Privola U, Moshnyakow M, et al. Expression patterns of neurotrophin and their receptor mRNAs in the rat inner ear. Hear Res 1993;65:69–78.
86. Staecker H, Lefebvre P, Liu W, et al. Brain derived neurotrophic factor and its influence on the otocyst [abstract]. In: Assoc Res Otolaryngol Mid-Winter Meeting; 1993 Feb 7–11.
87. Ernfors P, Van De Water T, Loring J, Jaenisch R. Complementary roles of BDNF and NT-3 in vestibular and auditory development. Neuron 1995; 14:1153–64.
88. Corey DP, Hudspeth AJ. Ionic basis of the receptor potential in a vertebrate hair cell. Nature 1979;281: 675–7.
89. Corey DP, Hudspeth AJ. Kinetics of the receptor current in bullfrog saccular hair cells. J Neurosci 1983; 3:962–76.
90. Hudspeth AJ. Extracellular current flow and the site of transduction by vertebrate hair cells. J Neurosci 1982;2:1–10.
91. Pickles JO, Comis SD, Osborne MP. Cross-links between stereocilia in the guinea pig organ of Corti, and their possible relation to sensory transduction. Hear Res 1984;15:103–12.
92. Assad JA, Shepherd CMC, Corey DP. Tip-link integrity and mechanical transduction vertebrate hair cells. Neuron 1991;7:985–94.
93. Hackney CM, Furness DN, Benos DJ, et al. Putative immunolocalization of the mechanoelectrical transduction channels in mammalian cochlear hair cells. Proc R Soc Lond B Bio Sci 1992;8:5–2.
94. Hudspeth AJ, Gillespie PG. Pulling springs to tune transduction: adaption by hair cells. Neuron 1994; 12:1–9.
95. Gillespie PG, Wagner MC, Hudspeth AJ. Identification of a 120 kd hair-bundle myosin located near stereociliary tips. Neuron 1993;11:581–94.
96. Walker RG, Willingham AT, Zuker CS. A *Drosophila* mechanosensory transduction channel. Science 2000; 287:2229–34.
97. Howard J, Hudspeth AJ. Mechanical reaction of the hair bundle mediated adaptation in mechanoelectrical transduction by the bullfrog's saccular hair cell. Proc Natl Acad Sci U S A 1987;84:3064–8.
98. Sole CK, Derfler BH, Duyks GM, Corey DP. Molecular cloning of myosins from the bullfrog saccular macula: a candidate for the hair cell adaptation motor. Auditory Neurol 1994;3:63–75.
99. Kemp DT. Stimulated acoustic emissions from within the human auditory system. J Acoust Soc Am 1978;64:1386–91.
100. Brownell WE, Bader CR, Bertrand D, de Ribaupierre Y. Evoked mechanical responses of isolated cochlear outer hair cells. Science 1985;227:194–6.
101. Kachar B, Brownell WE, Altschuler R, Fex J. Electrokinetic shape changes of cochlear outer hair cells. Nature 1986;322:365–7.
102. Dallos P, Evans BN, Hallworth R. Nature of the motor element in electrokinetic shape changes of cochlear outer hair cells. Nature 1991;350:155–7.
103. Kalinec F, Holley MC, Iwasa K, et al. A membrane-based force generation mechanism in auditory sensory cells. Proc Natl Acad Sci U S A 1992;89: 8671–5.

104. Zheng J, Shen W, He DZ, et al. Prestin is the motor protein of cochlear outer hair cells. Nature 2000; 405:149–55.
105. Ludwig J, Oliver D, Frank G, et al. Reciprocal electromechanical properties of rat prestin: the motor molecule from rat outer hair cells. Proc Natl Acad Sci U S A 2001;98:4178–83.
106. Oliver D, He D, Klöcker N, et al. Intracellular anions as the voltage sensor of prestin, the outer hair cell motor protein. Science 2001;292:2340–3.
107. Niedzielski AS, Safieddine S, Wenthold RJ. Molecular analysis of excitatory amino acid receptor expression in the cochlea [published erratum appears in Audiol Neurootol 1997;2:231]. Audiol Neurootol 1997;2:79–91.
108. Ryan AF, Brumm D, Kraft M. Occurrence and distribution of non-NMDA glutamate receptor mRNAs in the cochlea. Neuroreport 1991;2:643–46.
109. Niedzielski AS, Wenthold RJ. Expression of AMPA, kainate, and NMDA receptor subunits in cochlear and vestibular ganglia. J Neurosci 1995;15:2338–53.
110. Housley CD, Batcher S, Kraft M, Ryan AF. Nicotinic acetylcholine receptor subunits expressed in rat cochlea detected by polymerase chain reaction. Hear Res 1994;75:47–53.
111. Drescher DC, Upadhyay S, Wilcox E, Fex J. Analysis of muscarinic receptor subtypes in the mouse cochlea by means of PCR. J Neurochem 1992;59:765–7.
112. Eybalin M, Safieddine S. Neurotransmitters and neuro-modulators in the cochlea. Presented at the Inner Ear Neuropharmacology Symposium, Montpellier, France, 1994.
113. Drescher DC, Green GE, Khan KM, et al. Analysis of gamma-aminobutyric acid A receptor subunits in the mouse cochlea by means of polymerase chain reaction. J Neurochem 1993;61:1167–70.
114. Housley GD, Raybould NP, Taylor L. Variation in conductances activated in response to extracellular ATP and tonotopic-dependent expression of P2 purinoceptors in the sensory hair cells of the inner ear. Presented at the Inner Ear Neuropharmacology Symposium Montpellier, France, 1994.
115. Ryan AF, Simmons DM, Watts A, Swanson LW. Enkephalin mRNA production by cochlear and vestibular efferents in the gerbil brainstem. Exp Brain Res 1991;87:259–67.
116. Elgoyhen A, Johnson D, Boulter J, et al. Alpha 9: acetylcholine receptor with novel pharmacological properties expressed in rat cochlear hair cells. Cell 1994;18:705–15.
117. Ryan AF, Watts AC. Expression of genes coding for α and β isoforms of Na/K-ATPase in the cochlea of the rat. Cell Mol Neurosci 1991;2:179–87.
118. Fina M, Ryan AF. Expression of mRNAs encoding subunit isoforms of the Na, K-ATPase in the vestibular labyrinth of the rat. Cell Mol Neurosci 1994;5:604–13.
119. Killick R, Richardson G. Isolation of clones for the α-subunit of the epithelial Na channel from chick cochlear cDNA library. Presented at the Inner Ear Neuropharmacology Symposium, Montpellier, France, 1994.
120. Steel KP. The benefits of recycling. Science 1999; 285:1363–4.
121. Delpire E, Lu J, England R, et al. Deafness and imbalance associated with inactivation of the secretory Na-K-2Cl co-transporter. Nat Genet 1999;22:192–5.
122. Vetter D, Mann J, Wangemann P, et al. Inner ear defects induced by null mutation of the *isk* gene. Neuron 1996;17:1251–64.
123. Thalmann I, Suzuki H, McCourt D, et al. Partial amino acid sequences of organ of Corti proteins OCP1 and OCP2: a progress report. Hear Res 1993; 64:191–8.
124. Chen H, Thalmann I, Adams J, et al. Cloning, molecular analysis and tissue distribution of *Ocp2*, a gene encoding a putative transcription-associated factor predominantly expressed in the auditory organs. Genomics 1995;27:389–98.
125. Legan K, Killick R, Goodyear R, Richardson G. Expression of the 41 Kda chick tectorin MRNA is restricted to the striolar region within the lagenar macula [abstract]. Abstracts Inner Ear Biology 1994;31:48.
126. Davis JG, Oberholtzer JC, Burns FR, Greene MI. Molecular cloning and characterization of an inner ear-specific structural protein. Science 1995;267: 1031–4.
127. Harter C, Ripoll C, Lenoir M, et al. Expression pattern of mammalian cochlea outer hair cell (OHC) mRNA: screening of a rat OHC cDNA library. DNA Cell Biol 1999;18:1–10.
128. Ryan AF, Dulon D, Keithley EM, Erkman L. A Brn-3.1 promoter element directs gene expression limited to hair cells in postnatal transgenic mice [abstract]. Abstr Conf Mol Biol Hear Deafness 1998;3:166.
129. Zuo J, Treadaway J, Buckner TW, Fritzsch B. Visualization of alpha9 acetylcholine receptor expres-

sion in hair cells of transgenic mice containing a modified bacterial artificial chromosome. Proc Natl Acad Sci U S A 1999;96:14100–5.

130. Yoo TJ, Fujiyoshi T, Readhead C, Hood L. Restoration of auditory evoked potential by myelin basic protein (MBP) gene therapy in shiverer mice. In: Assoc Res Otolaryngol Mid-Winter Meeting; 1993 Feb 7–11 p. 147.
131. Dulon D, Ryan AF. The bacterial *neo* gene confers neomycin resistance to mammalian cochlear hair cells. Neuroreport 1999;10:1189–93.
132. Han JJ, Mahtre AN, Wareing M, et al. Transgene expression in the guinea pig cochlea mediated by a lentivirus-derived gene transfer vector. Hum Gene Ther 1999;10:1867–73.
133. Dazert S, Aletsee C, Brors D, et al. In vivo adenoviral transduction of the neonatal rat cochlea and middle ear. Hear Res. 2001;151:30–40.
134. Holt JR, Johns DC, Wang S, et al. Functional expression of exogenous proteins in mammalian sensory hair cells infected with adenoviral vectors. J Neurophysiol 1999;81:1881–8.
135. Staecker H, Gabaizadeh R, Federoff H, Van De Water TR. Brain-derived neurotrophic factor gene therapy prevents spiral ganglion degeneration after hair cell loss. Otolaryngol Head Neck Surg 1998; 119:7–13.
136. Suzuki M, Yagi M, Brown JN, et al. Effect of transgenic GDNF expression on gentamicin-induced cochlear and vestibular toxicity. Gene Ther 2000; 7:1046–54.
137. Ryan AF, Luo L, Baird A. Implantation of cells transfected with the FGF-1 gene induces middle ear mucosal proliferation. In: Lim DL, editor. Recent advances in otitis media with effusion. Amsterdam: Kugler; 1997. p. 248–50.
138. Staecker H, Kopke R, Lefebvre P, et al. The effects of neurotrophins on adult auditory neurons in vitro and in vivo [abstract]. In: Assoc Res Otolarynol Mid-Winter Meeting; 1995 Feb 5–9. p. 177.
139. Post JC, Ehrlich GD. The impact of the polymerase chain reaction in clinical medicine. JAMA 2000; 283:1544–6.
140. Melhus A, Ryan AF. Expression of cytokine genes during pneumococcal and nontypeable *Haemophilus influenzae* acute otitis media in the rat. Infect Immun 2000;68:4024–31.
141. Ehrlich GD, Post JC. Susceptibility to otitis media: strong evidence that genetics plays a role. JAMA 1999;282:2167–9.
142. Ryan AF, Luo L. Expression of nitric oxide synthase in the mouse middle ear mucosa during immune-mediated chronic otitis media. In: Tos M, Thomsen J, Balle V, editors. Recent advances in otitis media. Amsterdam: Kugler; 2000. p. 112–6.
143. Ohtani F, Furuta Y, Iino Y, et al. Amplification of RNA from archival human temporal bone sections. Laryngoscope 1999;109:617–20.
144. Andersson B, Svanborg-Eden C. Attachment of *Streptococcus pneumoniae* to human pharyngeal epithelial cells. Respiration 1989;55 Suppl 1:49–52.
145. Coleman T, Grass S, Munson R. Molecular cloning, expression, and sequence of the pilin gene from nontypeable *Haemophilus influenzae* M37. Infect Immun 1991;59:1716–22.
146. Kar S, To SC, Brinton CC. Cloning and expression in *Escherichia coli* of LKP pilus genes from a nontypeable *Haemophilus influenzae* strain. Infect Immun 1990;58:903–8.
147. Bakaletz LO, Ahmed MA, Kolattukudy PE, et al. Cloning and sequence analysis of a pilin-like gene from an otitis media isolate of nontypeable *Haemophilus influenzae*. J Infect Dis 1992;165 Suppl 1:S201–3.
148. Sirakova T, Kolattukudy PE, Murwin D, et al. Role of fimbriae expressed by nontypeable *Haemophilus influenzae* (NTHi) in the pathogenesis of and protection against otitis media and relatedness of the fimbrin subunit to outer membrane protein A. Infect Immun 1994;62:2002–20.
149. St. Geme JW, Falkow S, Barenkamp SJ. High-molecular-weight proteins of nontypeable *Haemophilus influenzae* mediate attachment to human epithelial cells. Proc Natl Acad Sci U S A 1993;90:75–9.
150. Brinton CC, Carter M, Derber DB, et al. Design and development of pilus vaccines for *Haemophilus influenzae* diseases. Pediatr Infect Dis J 1989;Suppl 8:54–61.
151. Bakaletz LO, Tallan BM, Hoepf TM, et al. Frequency of fimbriation of nontypeable *Haemophilus influenzae* and its ability to adhere to chinchilla and human respiratory epithelium. Infect Immun 1988; 56:331–5.
152. Spinola SM, Griffiths GE, Shanks KL, Blake MS. The major outer membrane protein of *Haemophilus ducreyi* is a member of the OmpA family of proteins. Infect Immun 1993;61:1346–51.
153. DeMaria TF, Yamaguchi T, Lim DJ. Quantitative cytologic and histologic changes in the middle ear

after the injection of non-typeable H*aemophilus influenzae* endotoxin. Am J Otolaryngol 1989; 10:261–6.
154. Weiser JN, Lindberg AA, Manning EJ, et al. Identification of chromosomal locus for expression of lipopolysaccharide epitopes in *Haemophilus influenzae*. Infect Immun 1989;57:3045–52.
155. Weiser JN, Williams A, Moxon ER. Phase-variable lipopolysaccharide structures enhance the invasive capacity of *Haemophilus influenzae*. Infect Immun 1990;58:3455–7.
156. Weiser JN, Maskell DJ, Butler PD, et al. Characterization of repetitive sequences controlling phase variation of *Haemophilus influenzae* lipopolysaccharide. J Bacteriol 1990;172:3304–9.
157. Berry AM, Paton JC, Hansman D. Effect of insertional inactivation of the genes encoding pneumolysin and autolysin on the virulence of *Streptococcus pneumoniae* type 3. Microb Pathog 1992;12:87–93.
158. Lock RA, Hansman D, Paton JC. Comparative efficacy of autolysin and pneumolysin as immunogens protecting mice against infection by *Streptococcus pneumoniae*. Microb Pathog 1992;12:137–43.
159. Ogra PL, Barenkamp SJ, Mogi G, et al. Microbiology, immunology, biochemistry and vaccination; Recent Advances in Otitis Media. Report of the Fifth Research Conference. Ann Otol Rhinol Laryngol 1994; Suppl 164 103:8:27–43.
160. Karma PH, Bakaletz LO, Giebink GS, et al. Present views on possibilities of immunoprophylaxis of acute otitis media in infants and children. In: Immunological aspects of otitis media. Int J Fed Otolaryngol 1995;32:127–34.
161. Nelson MB, Munson RS, Apicella MA, et al. Molecular conservation of the P6 outer membrane protein among strains of *Haemophilus influenzae*: analysis of antigenic determinants, gene sequences, and restriction fragment length polymorphisms. Infect Immun 1991;59:58–63.
162. Deich RA, Metcalf BJ, Finn CW, et al. Cloning of genes encoding a 15,000-dalton peptidoglycan-associated outer membrane lipoprotein and an antigenically related 15,000-dalton protein from *Haemophilus influenzae*. J Bacteriol 1988;170:489–98.
163. Sikkema DJ, Murphy TF. Molecular analysis of the P2 porin protein of nontypeable *Haemophilus influenzae*. Infect Immun 1992;60:5204–11.
164. Martin D, Munson R, Grass S, et al. Mapping of B-cell epitopes on the outer membrane P2 porin protein *Haemophilus influenzae* by using recombinant proteins and synthetic peptides. Infect Immun 1991;59:1457–64.
165. Haasse EM, Yi K, Morse GD, Murphy TF. Mapping of bactericidal epitopes on the P2 porin protein of nontypeable *Haemophilus influenzae*. Infect Immun 1994;62:3712–22.
166. Green BA, Metcalf BJ, Quinn-Dey T, et al. A recombinant non-fatty acylated form of the Hi-PAL (P6) protein of *Haemophilus influenzae* elicits biologically active antibody against both nontypeable and type b *H. influenzae.* Infect Immun 1990;58:3272–8.
167. Murphy TF, Kirkham C. Amino acid sequence of a surface-exposed epitope of the P6 outer membrane protein of nontypable *Haemophilus influenzae* [abstract]. In: Lim DJ, editor. Abstracts of the Fifth International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1993. p. 120–l.
168. El-Adhami W, Kyd JM, Bastin DA, Cripps AW. Characterization of the gene encoding a 26-kilodalton protein (OMP26) from nontypeable *Haemophilus influenzae* and immune responses to the recombinant protein. Infect Immun 1999; 67:1935–42.
169. Deich RA, Anilionis A, Fulginiti J, et al. Antigenic conservation of the 15,000-dalton outer membrane lipoprotein PCP of *Haemophilus influenzae* and biologic activity of anti-PCP antisera. Infect Immun 1990;58:3388–93.
170. Barenkamp SJ, Bodor FF. Development of serum bactericidal activity following nontypeable *Haemophilus influenzae* acute otitis media. Pediatr Infect Dis J 1990;9:333–9.
171. Barenkamp SJ, Leininger E. Cloning, expression, and DNA sequence analysis of genes encoding nontypeable *Haemophilus influenzae* high-molecular-weight surface-exposed proteins related to filamentous hemagglutinin of *Bordetella pertussis.* Infect Immun 1992;60:1302–13.
172. Alexander JE, Lock RA, Peeters CC, et al. Immunization of mice with pneumolysin toxoid confers a significant degree of protection against at least nine serotypes of *Streptococcus pneumoniae*. Infect Immun 1994;62:5683–8.
173. Swiatlo E, Brooks-Walter A, Briles DE, McDaniel LS. Oligonucleotides identify conserved and variable regions of pspA and pspA-like sequences of *Streptococcus pneumoniae*. Gene 1997;188:279–84.
174. Okamoto Y, Kudo K, Ishikawa K, et al. Presence of respiratory syncytial virus genomic sequences in

middle ear fluid and its relationship to expression of cytokines and cell adhesion molecules. J Infect Dis 1993;168:1277–81.
175. Giebink GS. Studies of *Streptococcus pneumoniae* and influenzae virus vaccines in the chinchilla otitis media model. Pediatr Infect Dis J 1989;8(1 Suppl):S42–4.
176. Heikkinen T, Ruuskanenn 0, Waris M, et al. Influenza vaccination in the prevention of acute otitis media in children. Am J Dis Child 1991;145: 445–8.
177. Robinson HL, Hunt LA, Webster RG. Protection against a lethal influenza virus challenge by immunization with hemoagglutinin-expressing plasmid DNA. Vaccine 1993;11:957–60.
178. Montgomery DL, Shiver JW, Leander KR, et al. Heterologous and homologous protection against influenza A by DNA vaccination: optimization of DNA vectors. DNA Cell Biol 1993;12:777–83.
179. Needham CA. *Haemophilus influenzae*: antibiotic susceptibility. Clin Microbiol Rev 1988;1:218–27.
180. DeGraff J, Elwell OP, Falkow S. Molecular nature of two beta-lactamase-specifying plasmids isolated from *Haemophilus influenzae* type b. J Bacteriol 1976;126:439–46.
181. Mendelman PM, Chaffin DO, Stull TL, et al. Characterization of non-beta-lactamase-mediated ampicillin resistance in *Haemophilus influenzae*. Antimicrob Agents Chemother 1984;26:235–44.
182. Clairoux N, Picard M, Brochu A, et al. Molecular basis of the non-beta-lactamase-mediated resistance to beta-lactam antibiotics in strains of *Haemophilus influenzae* isolated in Canada. Antimicrob Agents Chemother 1992;36:1504–13.
183. Malouin F, Bryan LE. DNA probe technology for detection of *Haemophilus influenzae*. Mol Cell Probes 1987;1:221–32.
184. Roberts MC, Swenson CD, Owens LM, Smith AL. Characterization of chloramphenicol-resistant *Haemophilus influenzae*. Antimicrob Agents Chemother 1980;18:610–5.
185. de Groot R, Campos J, Moseley SL, Smith AL. Molecular cloning and mechanism of trimethoprim resistance in *Haemophilus influenzae*. Antimicrob Agents Chemother 1988;32:477–84.
186. Klugman KP. Pneumococcal resistance to antibiotics. Clin Microbiol Rev 1990;3:171–96.
187. Dowson CG, Hutchinson A, Spratt BG. Extensive remodeling of the transpeptidase domain of penicillin-binding protein 2B of a penicillin-resistant South African isolate of *Streptococcus pneumoniae*. Mol Microbiol 1989;3:95–102.
188. Klugman KP, Coffey TJ, Smith A, et al. Cluster of an erythromycin-resistant variant of the Spanish multiply resistant 23F clone of *Streptococcus pneumoniae* in South Africa. Eur J Clin Microbiol Infect Dis 1994;13:171–4.
189. Soares S, Kristinsson KG, Musser JM, Tomasz A. Evidence for the introduction of a multiresistant clone of serotype 6B *Streptococcus pneumoniae* from Spain to Iceland in the late 1980s. J Infect Dis 1993;168:158–63.
190. Owen RJ. Chromosomal DNA fingerprinting — a new method of species and strain identification applicable to microbial pathogens. J Med Microbiol 1989;30:89–99.
191. Loos BG, Bernstein JM, Dryja DM, et al. Determination of the epidemiology and transmission of non-typeable *Haemophilus influenzae* in children with otitis media by comparison of total genomic DNA restriction fingerprints. Infect Immun 1989; 57:2751–7.
192. Dickinson DP, Loos GB, Dryja DM, Bernstein JM. Restriction fragment mapping of *Branhamella catarrhalis*: a new tool for studying the epidemiology of this middle ear pathogen. J Infect Dis 1988; 158:205–8.
193. Salo P, Ortqvist A, Leinonen M. Diagnosis of bacteremic pneumococcal pneumonia by amplification of pneumolysin gene fragment in serum. J Infect Dis 1995;171:479–82.
194. Oehl H, Ryan A. Expression of genes encoding growth factor receptors in middle ear mucosa [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 70.
195. DeFoire-Hyrmer J, Bakaletz LO. Evidence obtained by confocal microscopy that links phosphorylation of tyrosine with nontypeable *Haemophilus influenzae* adherence and actin nucleation [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 175.
196. Sudhoff H, Dazert S, Gonzales AM, et al. Angiogenesis and angiogenic growth factors in middle ear cholesteatoma. Am J Otol 2000;21:793–8.
197. Huang CC, Chen CT, Huang TS, Shinoda H. Mediation of signal transduction in keratinocytes of human middle ear cholesteatoma by ras protein. Eur Arch Otorhinolaryngol 1996;253:385–9.

198. Ishibashi T, Shinogami M, Kaga K, Fukaya T. Keratinocyte growth factor and receptor mRNA expression in cholesteatoma of the middle ear. Acta Otolaryngol (Stockh) 1997;117:714–8.
199. Shinoda H, Huang CC. Expressions of c-jun and p53 proteins in human middle ear cholesteatoma: relationship to keratinocyte proliferation, differentiation, and programmed cell death. Laryngoscope 1995;105:1232–7.
200. Schilling V, Lang S, Rasp G, et al. Overexpression of tenascin in cholesteatoma and external auditory meatal skin compared to retroauricular epidermis. Acta Otolaryngol (Stockh) 1996;116:741–6.
201. Park K, Park HJ, Chun YM, Lee DH. Expression of PLC-gamma 1 in experimental cholesteatoma [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 80.
202. Hebda PA, Alper CM, Doyle WJ, et al. Up-regulation of messenger RNA for inflammatory cytokines in middle ear mucosa in a rat model of acute otitis media. Ann Otol Rhinol Laryngol 1998; 107:501–8.
203. Melhus Å, Forseni M, Ryan AF. Molecular markers for bone formation during bacterial otitis media in the rat. In: Lim DJ, et al, editors. Recent advances in otitis media with effusion. Amsterdam: Kugler; 2000.
204. Bikhazi P, Luo L, Ryan AF. Expression of immunoregulatory cytokines during acute and chronic middle ear immune response. Laryngoscope 1995;105:629–34.
205. John EO, Nam TTK. Concentration of nitric oxide in middle ear effusion [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 83.
206. Juhn S, Li W, Flaagan J, Javel E. Inflammatory NO involvement in auditory function [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 88.
207. Li JD, Xu H, Wang B, et al. Signal transduction mechanisms involved in bacteria-induced mucin overproduction [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 77.
208. Li JD, Feng WJ, Gallup M, et al. Activation of NF-kappaB via a Src-dependent Ras-MAPK-pp90rsk pathway is required for *Pseudomonas aeruginosa*-induced mucin overproduction in epithelial cells. Proc Natl Acad Sci U S A 1998;95:5718–23.
209. Xu H, Wang B, Kim JH, et al. Oxygen radicals as second messengers for bacteria-induced up-regulation of mucin gene transcription [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 255.
210. Chun YM, Li JD, Lim DJ. MEK1/2-MAPK pathway is involved in IL-1-induced down-regulation of surfactant B [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 256.
211. Moon SK, Yoon JH, Lee HK, et al. Effects of retinoic acid, triiodothyronine and hydrocortisone on mucin and lysozyme expression in passage-2 normal human middle ear epithelial cells [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft Lauderdale (FL): 1999. p. 257.
212. Lin J, Kawano H, Tsuboi Y, et al. Alterations of mucin gene expression in the transition of acute to mucoid otitis media. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 72.
213. Hutton DA, Fogg FJ, Kubba H, et al. Heterogeneity in the protein cores of mucins isolated from human middle ear effusions: evidence for expression of different mucin gene products. Glycoconjugate J 1998;15:283–91.
214. Jung HH, Lee JH, Lim HH, Chung C. Expression of mucin genes in otitis media [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 251.
215. Severn TL, Hutton DA, Birchall JP, Pearson JP. Mucin gene expression in human middle ear [abstract]. In: Lim DJ, editor. Abstracts of the 7th International Symposium on Recent Advances in Otitis Media. Ft. Lauderdale (FL): 1999. p. 253.
216. Wang B, Lim DJ, Han J, et al. Novel cytoplasmic proteins of nontypeable *Haemophilus influenzae* up-regulate human *MUC5AC* mucin transcription via a positive p38 MAP kinase pathway and a negative PI 3-kinase-Akt pathway. J Biol Chem 2002; 277:949–57.
217. Shuto T, Xu H, Wang B, et al. Activation of NF-kappa B by nontypeable *Hemophilus influenzae* is

mediated by toll-like receptor 2-TAK1-dependent NIK-IKK alpha/beta-I kappa B alpha and MKK3/6-p38 MAP kinase signaling pathways in epithelial cells. Proc Natl Acad Sci U S A 2001;98: 8774–9.
218. Lim DJ, Chun YM, Lee HY, et al. Cell biology of tubotympanum in relation to pathogenesis of otitis media — a review. Vaccine 2000;19 Suppl 1:S17–25.
219. Lehrer RI, Ganz T. Antimicrobial peptides in mammalian and insect host defense. Curr Opin Immunol 1999;11:23–7.
220. Ganz T, Lehrer RI. Defensins. Pharmacol Ther 1995;66:191–205.
221. McNamara N, Van R, Tuchin OS, Fleiszig SM. Ocular surface epithelia express mRNA for human beta defensin-2. Exp Eye Res 1999;69:483–90.
222. Schroder JM, Harder J. Human beta-defensin-2. Int J Biochem Cell Biol 1999;31:645–51.
223. Takemura H, Kaku M, Kohno S, et al. Evaluation of susceptibility of gram-positive and -negative bacteria to human defensins by using radial diffusion assay. Antimicrob Agents Chemother 1996;40:2280–4.
224. O'Neil DA, Porter EM, Elewaut D, et al. Expression and regulation of the human beta-defensins hBD-1 and hBD-2 in intestinal epithelium. J Immunol 1999;163:6718–24.
225. Boe R, Silvola J, Yang J, et al. Human beta-defensin-1 mRNA is transcribed in tympanic membrane and adjacent auditory canal epithelium. Infect Immun 1999;67:4843–6.
226. Lim DJ, Liu YS, Birck H. Secretory lysozyme of the human middle ear mucosa: immunocytochemical localization. Ann Otol Rhinol Laryngol 1976;85(1 Pt 1):50–60.
227. Vorland LH, Ulvatne H, Andersen J, et al. Lactoferricin of bovine origin is more active than lactoferricins of human, murine and caprine origin. Scand J Infect Dis 1998;30:513–7.
228. Qiu J, Hendrixson DR, Baker EN, et al. Human milk lactoferrin inactivates two putative colonization factors expressed by *Haemophilus influenzae*. Proc Natl Acad Sci U S A 1998;95:12641–6.
229. Hanamure Y, Lim DJ. Normal distribution of lysozyme- and lactoferrin-secreting cells in the chinchilla tubotympanum. Am J Otolaryngol 1986; 7:410–25.
230. Park K, Lim DJ. Development of secretory elements in murine tubotympanum: lysozyme and lactoferrin immunohistochemistry. Ann Otol Rhinol Laryngol 1993;102:385–95.
231. Lu J. Collectins: collectors of microorganisms for the innate immune system. Bioessays 1997;19:509–18.
232. Malhotra R, Lu J, Holmskov U, Sim RB. Collectins, collectin receptors and the lectin pathway of complement activation. Clin Exp Immunol 1994;97 Suppl 2:4–9.
233. Mondain M, Restituito S, Vincenti V, et al. Adenovirus-mediated in vivo gene transfer in guinea pig middle ear mucosa. Hum Gene Ther 1998;9:1217–21.

SUGGESTED READING

Bockaert J. G proteins and G-protein coupled receptors: structure, function and interactions. Curr Opin Neurobiol 1991;1:32–42.

Click RB, Pasternak JJ. Molecular biotechnology: principles & applications of recombinant DNA. Washington (DC): ASM Press; 1994.

Darnell J, Lodish H, Baltimore D. Molecular cell biology. 2nd ed. New York: Scientific American Books; 1990.

Davis L, Kuehl M, Battey J. Basic methods in molecular biology. 2nd ed. New York: Appleton & Lange; 1994.

CHAPTER 4

Physiology of the Auditory and Vestibular Systems

Brenda L. Lonsbury-Martin, PhD, Glen K. Martin, PhD, Anne E. Luebke, PhD

AUDITORY SYSTEM

General Principles

Over the past three decades, major advances have occurred in our knowledge about how the ear achieves its high sensitivity, sharp frequency tuning, vast dynamic range, and precise temporal resolution. This accumulated wisdom has led to a new understanding about the unique functions of the inner hair cells (IHCs) and outer hair cells (OHCs) of the cochlea. Traditionally, the transfer of information about sounds from the environment to the higher centers of analysis in the central auditory nervous system was considered to be entirely a passive process. According to conventional thinking, the salient features of sound, including frequency, magnitude, and timing attributes, were principally encoded by peripheral processes and then simply passed relatively unaltered along the ascending system, from one structure to another, in a forward-moving manner. With the discoveries that the healthy ear can generate sounds in the form of otoacoustic emissions (OAEs) and that OHCs mechanically vibrate in response to depolarizing stimuli, the role of the cochlea in analyzing acoustic signals is now considered to represent an active process. Given the assumption that the vibromechanical aspects of OHC function underlie the production of OAEs and the awareness that the major portion of the cochlear efferent system innervates this particular class of auditory sensory cell, it is clear that central auditory centers may also modify peripherally generated responses in an active manner. Thus, rather than conceptualizing the role of the auditory system as a passive analyzer of environmental sounds, the modern view is to consider it as an active participant in controlling acoustic information so that the most meaningful features are registered.

Auditory Apparatus

A traditional approach toward understanding the primary sensitivity, frequency tuning, and timing functions of the peripheral auditory system is to divide the periphery into three discrete parts. In this manner, the unique contributions that the external, middle, and inner ear (ie, cochlea) make to the overall analysis of sound can best be appreciated.

External Ear

The external ear or pinna of the mammal is regarded as a simple funnel that collects and crudely filters sound. Given the immobility of the human external ear, it is assumed that this frequency tuning function is performed passively. However, evidence from experimental studies suggests that the human pinna serves two functions: (1) it aids in sound localization, especially front-to-back and high-to-low distinctions, where interaural time differences provide no clues, and (2) along with the external ear canal, it increases acoustic pressure at the tympanic membrane in the 1.5 to 5 kHz range, which is the frequency range most important for speech perception. Evidence for the sound-localizing function of the human external ear includes the demonstration of accurate sound localization in patients with monaural hearing and the loss of this localization ability when the pinna of the hearing ear is strapped to the head. In addition, abnormally poor sound localization around an "imperfectly" remodeled pinna has been reported.[1]

In summarizing available data on the role that the external ear plays in boosting sound pressure at the tympanic membrane, Shaw described the contributions of the distinct parts of the head and neck and external ear by successively adding their different components in a model.[2] According to this clas-

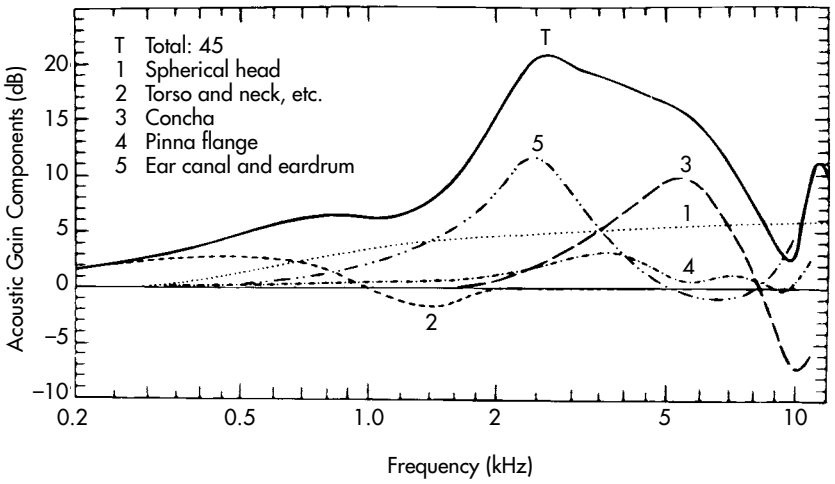


FIGURE 4–1. Pressure gain contributed by the individual components of the outer ear in humans. The total (T) curve plots the overall gain based on the addition of the various components in the order listed with the sound source positioned 45 degrees from straight ahead. Reprinted with permission from Shaw EAG.[2]

sic analysis, shown in Figure 4–1, most of the individual components of the external ear provide complementary gains, resulting in a significant increase in sound pressure from approximately 2 to 7 kHz. When the overall additive effect is appreciated ("T" in Figure 4–1), it can be seen that at certain frequencies this gain is substantial, resulting in sound pressure increases on the order of 20 dB.

Studies of models of the human external ear suggest that the pinna extracts information about sound location by altering the transmission properties of different frequencies according to the location of the sound source relative to the pinna. For example, when the sound source is behind the ear, interference of directly transmitted sound with sound waves scattered off the pinna flange or helix alters the response in the 3 to 6 kHz range. Thus, the external ear modifies the spectrum of the incoming sound, allowing an individual to make judgments about the location of unknown sound sources. This localization ability suggests that the central part of the auditory system can use very subtle spectral cues in the analysis of environmental sounds.[3]

Middle Ear

Figure 4–2 outlines the anatomy of the middle and inner ears. The middle ear ossicles form a transmission pathway that conducts sound energy from the tympanic membrane, at the interface of the external and middle ear, to the oval window of the cochlea. The discussion of middle ear function is separated into two categories: (1) "impedance matching" between the air of the external environment and the fluids (perilymph and endolymph) of the cochlea and (2) the acoustic reflex of the middle ear muscle system.

The widespread use of surgical modification of middle ear structures to improve hearing and the usefulness of immittance audiometry make the concept of acoustic impedance and its relationship to middle ear function important to the clinician. Accordingly, the discussion of the middle ear begins by reviewing some basic principles of acoustic impedance.

Transmission of Acoustic Energy Through the Middle Ear Figure 4–3, A, diagrams the route for the transmission of sound energy through the middle ear into the cochlear portion of the inner ear. The ossicular chain (see Figure 4–2), consisting of the malleus, incus, and stapes, can be thought of as a lever system. The tympanic membrane moves the manubrium or handle of the malleus. In turn, the long process of the incus and manubrium move together because the malleoincudal joint is essentially fixed. In contrast, the joint between the incus and the stapes is flexible. Therefore, because the stapes is fixed at its posteroinferior border, movement of the tympanic membrane causes it to rock in and out of the oval window. The changes in acoustic pressure caused by the stapes moving in and out of the oval window are transmitted instantaneously by the perilymph through the cochlear partition and then to the round window. This pressure transmission through the cochlear partition causes it to move either upward or downward, depending on the direction of the pressure change. The pressure

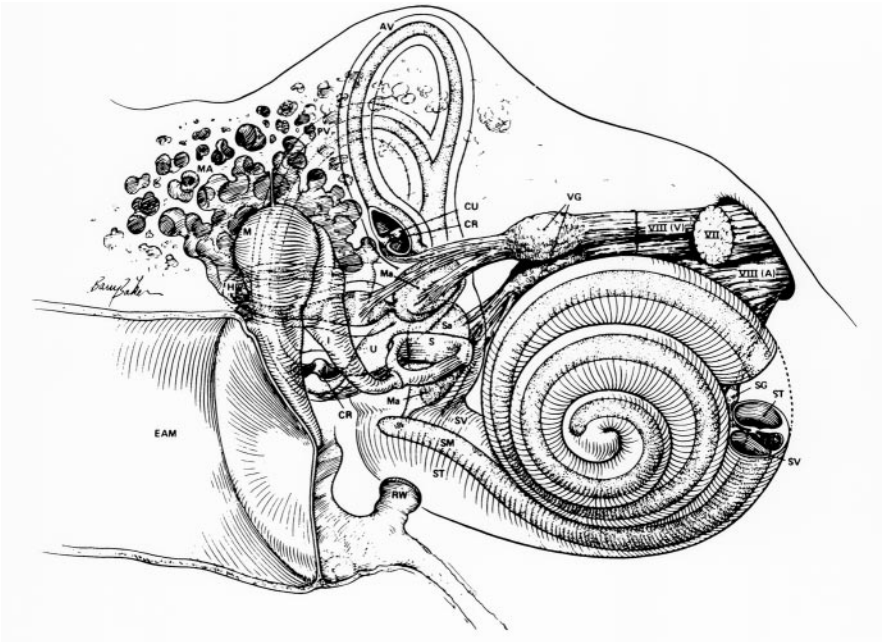


FIGURE 4–2. Schematic drawing of the inner ear (labyrinth), which consists of a series of tunnels within the petrous portion of the temporal bone. The osseous labyrinth (outer tunnel) is clear, and the membranous labyrinth (inner tunnel) is stippled. SV = scala vestibuli; ST = scala tympani; SM = scala media; SG = spiral ganglion, containing the cell bodies of the auditory nerve; VG = vestibular ganglia, containing the cell bodies of the vestibular nerve; VIII (A) = auditory part of the eighth cranial nerve; VIII (V) = vestibular part of the eighth cranial nerve; VII = facial nerve; Ma = macula; CU = cupula; CR = crista; Sa = sacculus; U = utriculus. Semicircular canals are labeled AV (anterior vertical), PV (posterior vertical), and H (horizontal). EAM = external auditory meatus; RW = round window; M = malleus; I = incus; S = stapes; MA = mastoid air cells in temporal bone.

change initiates a mechanical traveling wave, shown in Figure 4–3, B, that reaches a maximum at some point on the basilar membrane depending on the frequency of the stimulating sound. The mechanical traveling wave moves from the base to the apex of the cochlea largely owing to a reduced stiffness of the basilar membrane in the apical direction. As discussed below, this traveling wave disturbance causes the hair cells in the organ of Corti to stimulate the dendritic endings of the cochlear nerve, thus signaling to the central auditory system that a sound stimulus has occurred.

At high sound levels of ~100 to 110 dB SPL (sound pressure level), the mode of vibration of the ossicular chain changes. Thus, instead of rotating about its short axis, as shown in Figure 4–3, A, the stapes footplate turns about its long axis.[4] Because this change results in less efficient sound transmission through the middle ear, it likely serves a protective function. Interestingly, the change in vibration mode occurs at the threshold of feeling, thus suggesting that the somatic sensation caused by excessive sounds may be caused by the detection of the altered ossicular vibration by middle ear bone and tendon receptors.

Impedance Matching by the Middle Ear Hearing by terrestrial animals requires transmission of sound from an air to a fluid environment. A useful way to appreciate the problem in conducting sound effectively between two distinct media is to recall how difficult it is to listen to sounds produced even a few inches above the surface while swimming underwater. Thus, direct transmission of sound across an air/water boundary is extremely inefficient because the specific acoustic impedances of air and water differ greatly. Moreover, whenever energy is transmitted between media with different specific impedances, much of the energy is reflected back from the boundary between the two media. To help solve this problem, the middle ear matches reasonably well the impedances of the air with those of the

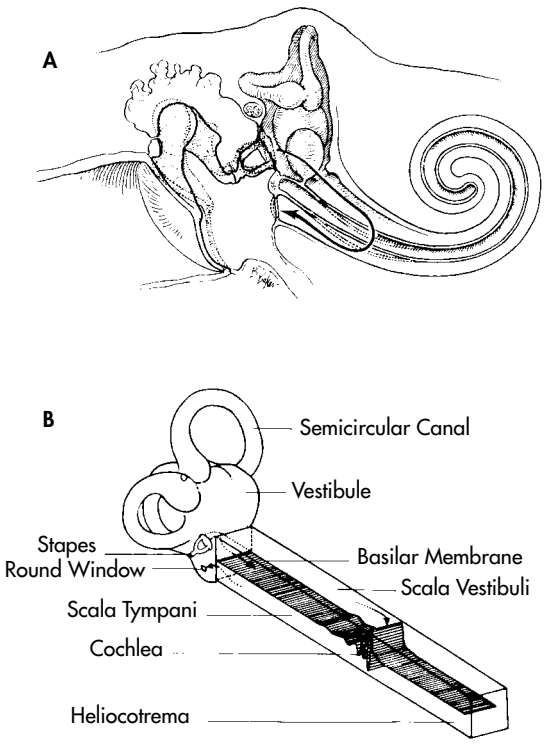


FIGURE 4–3. A, Transmission of tympanic membrane movement to the cochlea via the ossicles. The system at rest is unstippled. The stippled ossicles and dashed cochlear partition illustrate the system when the tympanic membrane is pushed inward by a sound wave. *B*, Inward pressure (*arrows*) initiates a traveling wave that migrates toward the apex of the cochlea (toward the helicotrema). This propagation depends largely on fluid coupling and the fact that the basilar membrane changes in stiffness, with the traveling wave moving from a region of highest stiffness (base) to a point of lower stiffness (apex). Reproduced with permission from Zweig G, Lipes R, Pierce JR. The cochlear compromise. J Acoust Soc Am 1976;59:975–82.

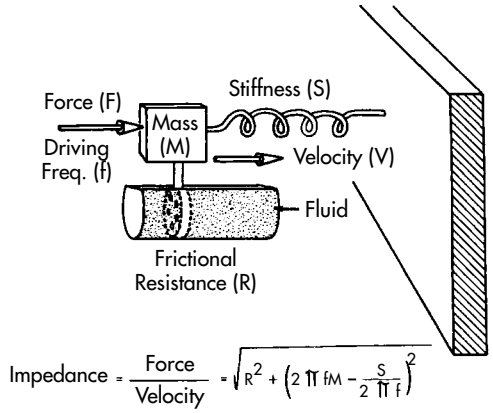


FIGURE 4–4. Principles of mechanical impedance. Frictional resistance is represented by a "dashpot"—a perforated piston operating inside a fluid-filled cylinder. The diagram could be converted to a representation of acoustic impedance by interposing a cylinder, or diaphragm, between the driving force (F) and the driven mass (M) and expressing the displacing input as pressure (ie, force per unit area).

cochlea and thereby greatly increases the efficiency of transmitting acoustic energy from the ambient environment into the cochlea.

The term impedance describes the opposition of a system to movement. Thus, the more force required to move a mechanical system at a given speed, the greater its impedance. Figure 4–4 illustrates the principles of mechanical impedance. The impedance of a mechanical system involves a complex relationship between the three physical parameters illustrated in Figure 4–4. Together, mass, stiffness, and resistance (ie, friction) determine the mechanical impedance of the middle ear system. Friction, the resistive component of impedance, consumes energy and is independent of the driving frequency. Stiffness and mass store energy; thus, they comprise the reactive component of impedance. For example, once the "fluid"-filled cylinder in Figure 4–4 is set in motion, it tends to continue because of inertia, and if the spring representing stiffness is compressed, it tends to push backward.

Acoustic impedance represents a special type of mechanical impedance in which force is replaced by pressure (ie, force per unit area) and the system is driven by sound. Thus, Figure 4–4 could be converted into a diagram of an acoustic system by interposing a piston, or membrane, between the force and the mass.

When air conducts sound, the stiffness component of its acoustic impedance is determined by the elastic coupling between air molecules, the mass component is determined by the mass of the air molecules, and the frictional component is determined by frictional resistance between the molecules. Because water is much denser and less compressible than air, it might seem at first that mass and stiffness create the principal difference between the acoustic impedance of the cochlea and that of air. However, Figure 4–4 demonstrates that transmission of energy into the cochlea does not

involve compression of the cochlear fluid itself. In addition, the elastic restorative forces of the cochlear partition and round window tend to cancel out the effect of the fluid's mass. Thus, the effective acoustic impedance of the cochlea is primarily resistive.[5]

As noted earlier, a primary function of the middle ear is to provide a means of "matching" the low impedance of air with the high cochlear impedance resulting from fluid flow and the distensibility of the cochlear membranes that separate the endolymph from the perilymph. Impedance matching by the middle ear is achieved by three factors: (1) the area of the tympanic membrane relative to the oval window, (2) the lever action of the middle ear ossicles, and (3) the shape of the tympanic membrane. The principles behind these factors are depicted in the three diagrams of Figure 4–5. By focusing the incident sound pressure from the large area of the tympanic membrane onto the small area of the oval window (Figure 4–5, A), the effectiveness of energy transfer between the air of the external ear canal and the fluids of the cochlea is increased greatly. In the human, the area ratio of the tympanic membrane to the oval window is about 20 to 1. However, the tympanic membrane does not vibrate as a whole.[6] Thus, the effective area ratio is only about 14 to 1. The ossicular chain also contributes to the transformer role of the middle ear by a levering action that increases the vibration amplitude (Figure 4–5, B). The ossicular chain lever ratio is around 1.3 to 1. A final factor influencing the efficiency of energy transfer depends on the conical shape of the tympanic membrane that allows a buckling action to occur (Figure 4–5, C). The buckling motion of the tympanic membrane results in an increased force and decreased velocity to produce approximately a fourfold increase in the effectiveness of energy transfer.[7] Together, these actions of the middle ear system result in an estimated overall transformer ratio of 73 to 1.

Middle Ear Muscles Mammals have two small skeletal muscles, the tensor tympani and the stapedius, which are attached to the ossicular chain. In primates, the stapedius muscle, which attaches to

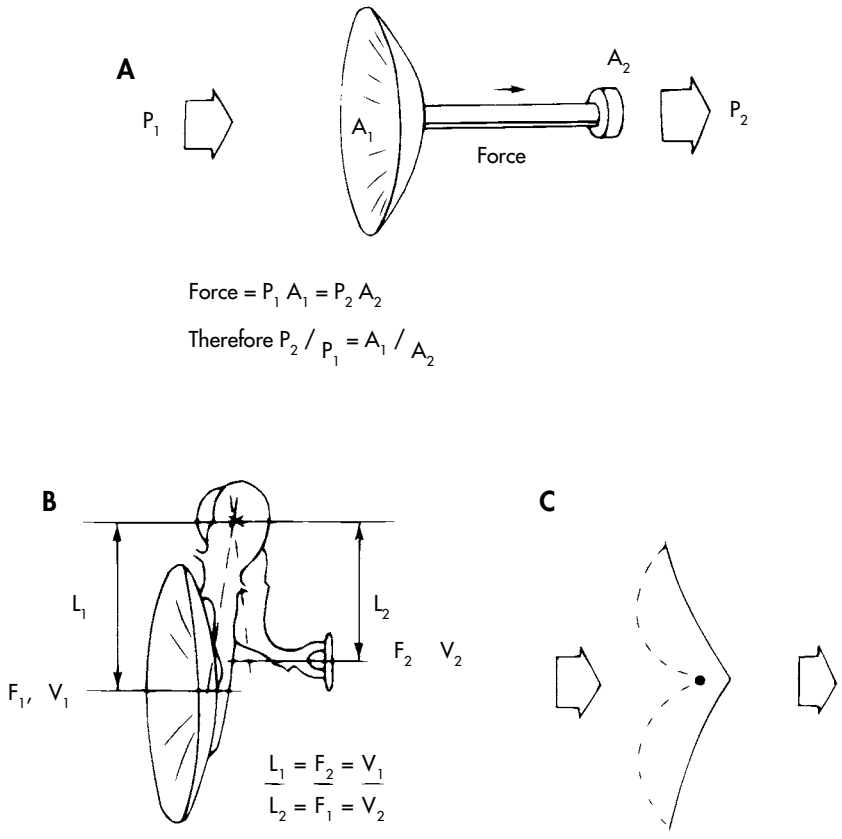


FIGURE 4–5. Illustration of the three principles of impedance matching performed by the middle ear. *A*, The primary factor is the ratio of the area of the tympanic membrane to that of the oval window. *B*, The lever action acts to increase the force and decrease the velocity. *C*, Buckling motion of the tympanic membrane also increases the force while reducing the velocity. Reproduced with permission from Pickles JO.[3]

the stapes and is innervated by the stapedial branch of the facial nerve, contracts reflexively in response to intense sound stimuli. However, the tensor tympani muscle, which attaches to the malleus and is innervated by the trigeminal nerve, probably does not.[8] In laboratory animals such as the cat, rabbit, and guinea pig, both muscles contract in response to loud sound. However, the threshold of the tensor tympani is often higher than that of the stapedius muscle. A four-neuron reflex arc consisting of the afferent fibers of the auditory nerve, neurons of the ventral cochlear nucleus, neurons of the medial superior olive, and facial motoneurons comprises the stapedius reflex pathway illustrated in Figure 4–6 for the rabbit. The reflex arc for the tensor tympani muscle is slightly different in that neurons of the ventral nucleus of the lateral lemniscus are also involved.[9] From the neural pathway depicted in Figure 4–6, it is clear that clinical abnormalities of the acoustic reflex primarily implicate pathology at the level of the lower brainstem.

In Figure 4–7, the sensitivity of the stapedius reflex in humans is plotted as a function of sound frequency. The reflex threshold curve evoked by pure tones parallels the audibility threshold curve but is about 80 dB above it.[10] Not surprisingly, because of their greater overall energy levels, broadband stimuli (eg, white noise) elicit the reflex more effectively than do pure tones.[11] Experimental evidence indicates that the stapedius reflex threshold decreases with increasing stimulus duration, with a time constant of about 200 ms.[12] This value approximates the time constant of temporal summation for the percepts of loudness and sensitivity. These findings, along with the clinical observation that the stapedius reflex exhibits "recruitment" in patients with cochlear hearing loss,[13] suggest that the acoustic reflex threshold correlates more with subjective loudness than with absolute stimulus intensity.

Contraction of the middle ear muscles can also be caused by nonauditory factors, including (1) spontaneous contractions, (2) body movements,[8] (3) vocalizations in which contractions begin prior to vocalization[8] (ie, prevocalization contractions), (4) movements of facial muscles involving only the tensor tympani,[8] (5) stimulation of the external ear canal,[14] and (6) voluntary contractions.[15]

As diagrammed in Figure 4–8, the stapedius muscle moves the stapes footplate backward and into the oval window, whereas the tensor tympani muscle pulls the manubrium inward. The effects of these contractions on the transmission of pure tones through the middle ear are illustrated in Figure 4–9.[15,16] In summary, the transmission of low-frequency sounds is attenuated by contraction of either muscle,[11] but the stapedius is probably a somewhat better attenuator than the tensor tympani.

The function of the human acoustic reflex has been studied by recording (1) gross muscle potentials via the electromyogram,[8] (2) pressure changes in the external auditory canal,[17] and (3) acoustic immittance changes.[11,13] The clinical application of acoustic impedance measurements is known as immittance audiometry, with the terms "impedance" and "immittance" being interchangeable. In Figure 4–10, the time course is shown for the human stapedius reflex as recorded by the immittance change method. The latency or time to the first detectable change generally varies from 10 to 30 ms, whereas the reflex decays over about 500 ms after the onset of the stimulus. The initial reduction in impedance is frequency dependent.

To date, acoustics-based clinical evaluations of middle ear function using immittance audiometry are an important part of the audiologic diagnostic test battery. However, they are not specific enough to represent independent diagnostic tests mainly because of limitations in the frequency range over

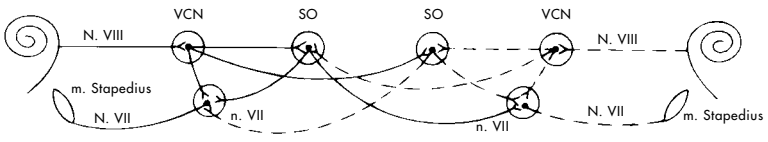


FIGURE 4–6. Neural pathway for the acoustic middle ear reflex in the rabbit. Reflex is mainly a chain of four neurons with a small ipsilateral three-neuron link. N. VIII = auditory nerve; N. VII = facial nerve; VCN = ventral cochlear nucleus; SO = superior olive; n. VII = facial nerve motoneuron. Reproduced with permission from Møller AR. Auditory physiology. New York: Academic Press; 1983.

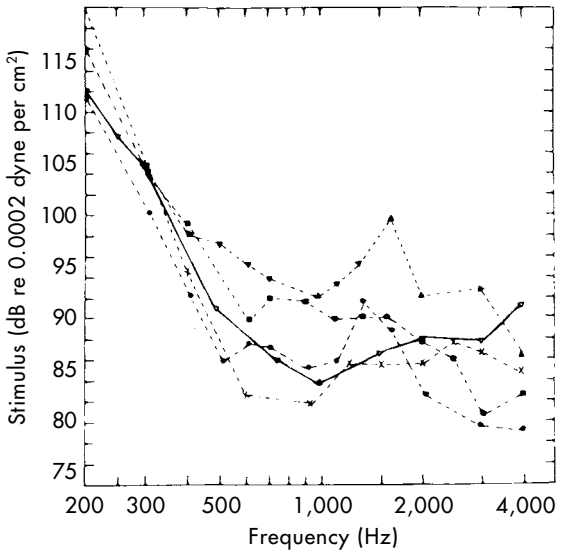


FIGURE 4–7. Sensitivity of the human acoustic reflex. *Dashed lines* from four subjects plot stimulus levels required to elicit an acoustic reflex with 10% of the maximum obtainable amplitude (measured with an immittance technique). The *solid line* is the threshold of audibility raised 80 dB. Reprinted with permission from Møller AR.[10]

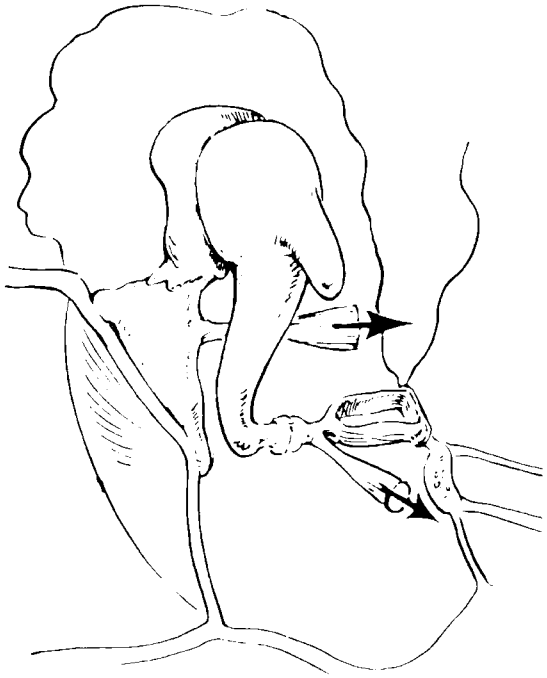


FIGURE 4–8. Action of the middle ear muscles. The view is from behind the head. The muscles are not drawn to scale. The tensor tympani muscle (*top arrow*) attaches to the handle of the malleus and pulls it backward, tensing the tympanic membrane. The stapedius muscle (*bottom arrow*) attaches to the neck of the stapes and pulls the posteroinferior border of the stapes down and into the oval window.

which impedance is easily measured at the tympanic membrane and the wide variations in their normal values. Currently, several new techniques, ear canal reflectance[18] and tympanic membrane velocity measurements,[19] are being investigated to determine if they can act as more specific indicators of middle ear dysfunction.

One major function of the middle ear muscles is to support and stiffen the ossicular chain.[11] In addition, because loud sounds are attenuated by the actions of the acoustic reflex, it is likely that another function of the reflex is to protect the inner ear against the damage that can be caused by overexposure to excessive sounds. This notion is supported by the results of a study investigating the amount of temporary threshold shift in patients suffering from acute Bell's palsy, which represents a disease in which the stapedius muscle is completely paralyzed.[20] These patients showed a greater threshold shift in the affected ear after exposures to low-frequency noise than in the opposite ear with a normal stapedius reflex. However, whether this protective effect is a "true" function of the stapedius reflex has been criticized on the basis that the continuous loud sounds against which the reflex is supposed to protect do not exist in nature.[21]

An alternative middle ear muscle function could be to attenuate low-frequency masking sounds that might otherwise interfere with auditory function. Contractions during chewing and other facial and body movements would attenuate the resultant internal body sounds, which are largely low frequency, while preserving sensitivity to high-frequency external sounds. The low-frequency attenuation produced by contractions prior to vocalization may also be functionally important. Observations that support a middle ear muscle role in vocalization and in speech discrimination are as follows: (1) patients with otosclerosis show significant deficits when administered the delayed feedback test for malingering,[22] (2) stutterers have a deficit in prevocalization middle ear muscle contraction,[23] and (3) absence of the stapedius reflex results in decreased speech discrimination when the level of speech is raised above 90 dB SPL.[23]

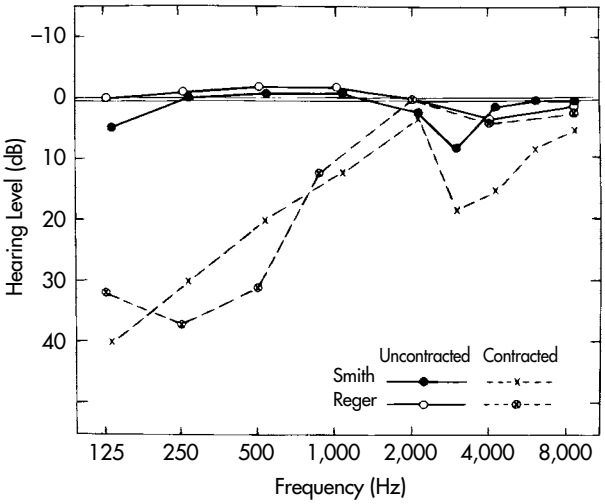


FIGURE 4–9. Effect of middle ear muscle contractions on pure-tone thresholds. The curves were obtained from normal subjects with the ability to voluntarily contract their middle ear muscles.[15,16] Although the effect of voluntary contraction may differ from the effect of normal involuntary contraction, the general observation that middle ear muscle contractions preferentially attenuate low frequencies is probably valid.

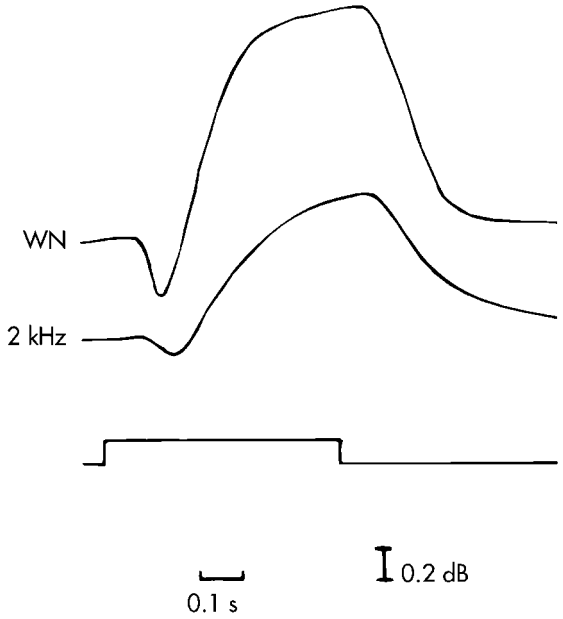


FIGURE 4–10. Time course of human stapedius contractions in response to white noise (WN) and 2 kHz pure-tone stimuli. *Bottom trace* shows stimulus time course. The stimuli were 90 dB above normal threshold and were delivered to the right ear. The muscle contractions were recorded from the left ear by the acoustic immittance change method. Records courtesy of Dr. James Jerger.

COCHLEA

The cochlea performs two basic functions as (1) a transducer that translates sound energy into a form suitable for stimulating the dendritic endings of the auditory nerve and (2) an encoder that programs the features of an acoustic stimulus so that the brain can process the information contained in the stimulating sound. Each of these functions is considered below.

Transducer Function

ANATOMY. As shown in Figure 4–11, A, the cochlea is divided into three tubes or scalae. The middle tube or scala media is the cochlear extension of the membranous labyrinth and is filled with a potassium (K^+)-rich, sodium (Na^+)-poor electrolyte fluid called endolymph. The outer two tubes, the scala vestibuli and the scala tympani, constitute the osseous labyrinth. They are separated by the scala media and filled with perilymph, a Na^+-rich, K^+-poor electrolyte fluid. When the cochlea is activated by sound, the scala media and its contents, bounded superiorly by Reissner's membrane and inferiorly by the basilar membrane, tend to move as a unit. This space is referred to as the "cochlear partition."[4]

The contents of scala media are illustrated in Figure 4–11, B. The auditory nerve fibers synapse at the bases of the hair cells of the organ of Corti. The dendrites of the auditory nerve enter the scala media through the habenulae perforatae, where they lose their myelin sheaths. The reticular lamina and the tectorial membrane are two membranes in the organ of Corti that are particularly critical in stimulus transduction. The reticular lamina, supported by the rods of Corti, resembles a stiff net with its webbing enmeshing the apical surfaces of the hair cells. Together, the rods of Corti and the reticular lamina provide the skeletal support of the organ of Corti.

The boundary separating the high-K^+ and high-Na^+ electrolytes is formed by the reticular lamina. As diagrammed in Figure 4–12, tight junctions, which are generally thought to present a barrier against ionic diffusion, are observed by scanning electron microscopy (SEM) between all cells that face the endolymphatic space. Figure 4–12 also indicates that the junctions between nonsensory epithe-

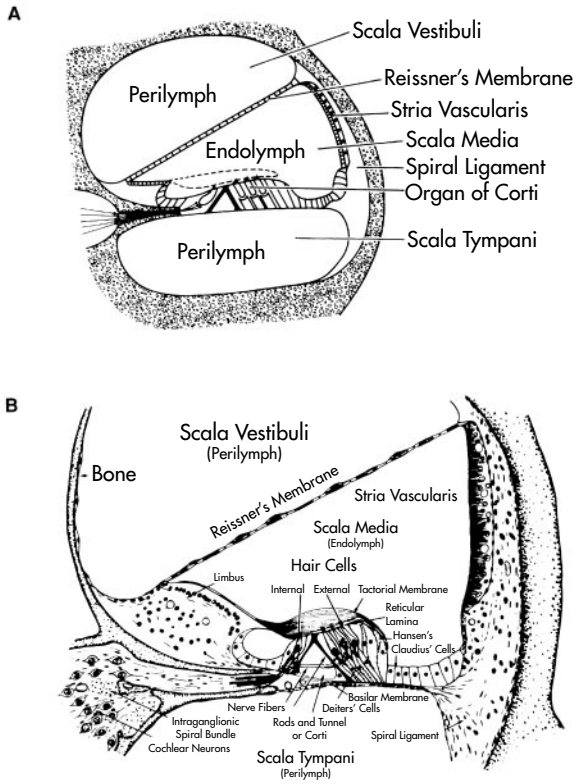


FIGURE 4–11. A, Simplified drawing depicting a cross-section through one turn of the cochlea showing the three compartments (scalae). The middle compartment (scala media) is filled with K^+-rich endolymph. The other two compartments (scala tympani and scala vestibuli) are enclosed in the osseous labyrinth. These compartments are filled with Na^+-rich perilymph. *B*, Semidiagrammatic representation of a cross-section of the guinea pig cochlear duct. *A* reproduced with permission from Salt AN, Konishi T. The cochlear fluids: perilymph and endolymph. In: Altschuler RA, Hoffman DW, Bobbin RP, editors. Neurobiology of hearing: the cochlea. New York: Raven Press; 1986. *B* reproduced with permission from Davis H. Advances in the neurophysiology and neuroanatomy of the cochlea. J Acoust Soc Am 1962;34:1377.

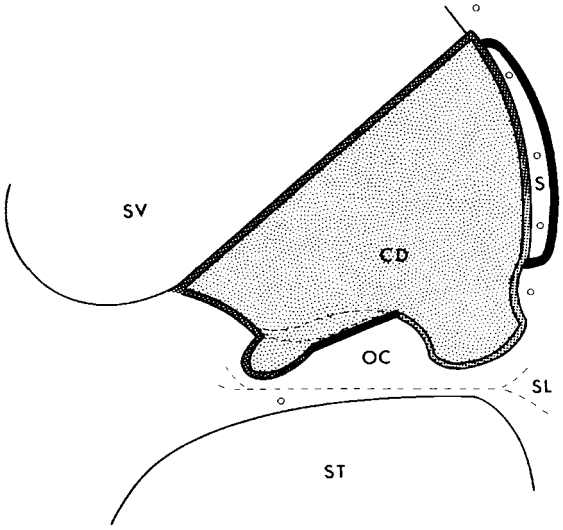


FIGURE 4–12. The perilymph/endolymph barriers around the scala media and stria vascularis. The *dotted line* represents "intermediate-to-tight" tight junctions, whereas the *solid line* represents "very tight" tight junctions. The *circles* represent blood vessels, which all have tight junctions that separate them from the intracochlear spaces. SV = scala vestibuli; ST = scala tympani; CD = cochlear duct; OC = organ of Corti; S = stria vascularis; SL = spiral ligament. Reproduced with permission from Jahnke K.[24]

lium (eg, the cells of Hensen and Claudius, the cells lining the spaces of Nuel, and the cells of Reissner's membrane) are not as "tight" (dotted line) as the junctions between the sensory cells and the basal cells (solid line) of the stria vascularis.[24] The fluid within the organ of Corti (ie, between the basilar membrane and the reticular lamina) has been termed cortilymph.[25] However, the fluid in this space is in free communication with the perilymph of scala tympani via relatively large channels through the basilar membrane.[26] Thus, cortilymph is probably identical to perilymph with respect to its electrolyte content. The tectorial membrane resembles a rather stiff, oval, gelatinous tube. It is attached to the limbus by a flexible membranous band that allows it to move up and down like the cover of a book.[27]

Mammals have three rows of OHCs. A fourth row of OHCs is seen in an occasional cross-section, especially in primates, but there is no well-defined fourth row.[28] There is only one row of IHCs. The fine finger-like projections present on the apical surface of the hair cells are called stereocilia. The detailed arrangements of the stereocilia on the hair cell apices are unique to each class of receptor and are illustrated in the SEM of Figure 4–13. The IHC stereocilia form a longitudinally oriented, relatively shallow curve, whereas the OHC cilia make a radially oriented "V," with a notch at the apex marking the site of the missing kinocilium that degenerates embryonically during development of cochlear hair

cells. The OHC stereocilia vary systematically in length in that they are longest at the apex of the V and shorten progressively from the apex to the distal limbs. Whereas the longest OHC stereocilia attach to the undersurface of the tectorial membrane, it is probable that the shorter cilia do not.[29] However, an interconnecting network of fine fibrils links the entire bundle of stereocilia, both laterally and from tip to tip, so that it moves as a unit.[30]

Until recently, little was known about stereocilia bundles. However, their importance in the transduction process has stimulated considerable research directed toward understanding the morphologic and physiologic properties of hair cell bundles and their separate stereocilia. The detailed structure and arrangement of individual hair cell bundles vary across species and even as a function of position along the organ of Corti. However, a number of general features of stereocilia bundles and individual cilia have been described, including (1) each bundle consists of 30 to 150 stereocilia arranged in several rows of decreasing length, (2) individual cilia range from 0.8 to 0.2 μm in diameter and increase in length from cochlear base to apex, and (3) each stereocilium is covered with a charged cell coat material that has been postulated to keep individual stereocilia from fusing together.[31]

Using various techniques, Flock and coworkers demonstrated that stereocilia contained the proteins actin and fimbrin.[32,33] Whether stereocilia are associated with motile responses owing to their musclelike composition is still not fully resolved. Indeed, one current proposition is that stereocilia and their associated ion channels participate in an amplificatory process that sensitizes and sharpens hearing.[34] However, the unidirectional orientation of the actin filaments[35] suggests that these proteins serve a structural, rather than a contractile, role.

Functional studies of the mechanical properties of stereocilia[36] showed that the cilia are stiff and pivot at their bases. When too much force is applied, the stereocilia break as if they were brittle. Noise exposure studies have also demonstrated the vulnerability of stereocilia to damage.[31] That is, the tallest stereocilia are affected first by becoming floppy owing to alterations in their membrane and cytoskeleton. With continued noise exposure, the stereocilia fuse and then degenerate, resulting in permanent hearing loss. In fact, the earliest measurable hearing changes resulting from noise exposure probably involve alterations in the stereocilia of the cochlea's hair cells.[31]

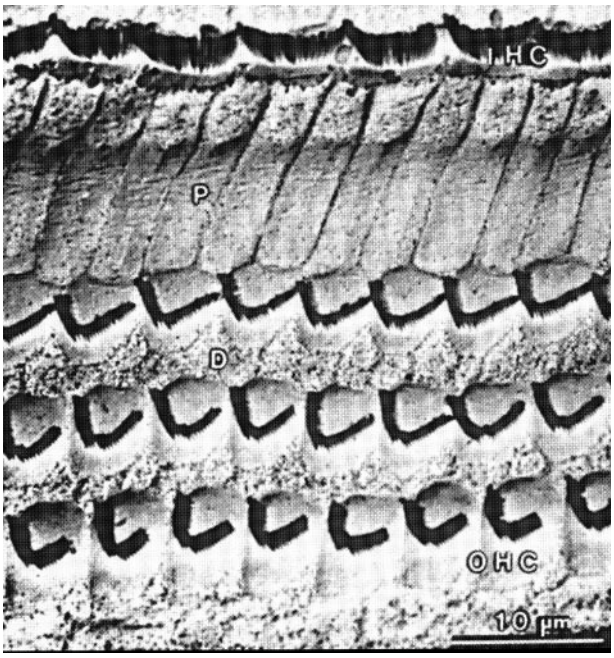


FIGURE 4–13. Scanning electron micrograph of the surface of the organ of Corti after removal of the tectorial membrane. A single row of inner hair cells (IHC) has cilia arranged in a linear (or shallow "U") pattern, whereas the three rows of outer hair cells (OHC) have stereocilia arranged in a "V" pattern. D = Deiters' cells; P = pillar cells. Reproduced with permission from Harrison RV, Hunter-Duvar IM. An anatomical tour of the cochlea. In: Jahn AF, Santos-Sacchi J, editors. Physiology of the ear. New York: Raven Press; 1988.

Mechanical Transduction. The final mechanical event in the cochlear transduction process is the bending of the stereocilia, as illustrated in Figure 4–14. Basilar membrane deformation causes a shearing action between the reticular and tectorial membranes. Because the long OHC cilia are attached to both membranes, they are bent (Figure 4–14, B). In contrast, the IHC cilia, and possibly also the shorter OHC cilia, which are not attached to the tectorial membrane, bend in response to some mechanism other than displacement shear. One proposition is that this process may involve fluid streaming between the sliding parallel plates formed by the reticular and tectorial membranes.[37] Fluid streaming would result from differences between the relative

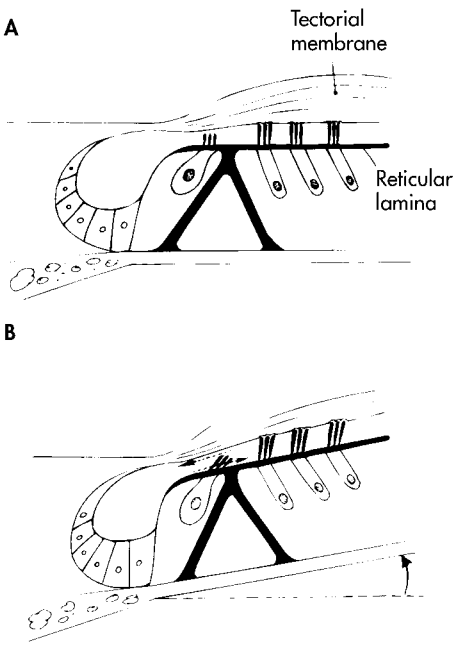


FIGURE 4–14. Translation of basilar membrane displacement (as shown in Figure 4–3) into bending of the hair cell cilia. *B*, Basilar membrane displacement (*arrow*) indicates the shearing action between the tectorial membrane and the reticular lamina bends the OHC cilia, which are attached to both structures. Streaming movement, imparted to the fluid between the reticular lamina and the tectorial membrane, may bend the inner hair cell cilia, which are not attached to the tectorial membrane. Inner hair cell cilia deflection may be longitudinal (ie, perpendicular to the page) rather than radial, as depicted.

velocities of the two membranes rather than by their relative displacements. Thus, the IHCs may be velocity sensors and the OHCs displacement sensors.[38] However, other notions about the unique functions of OHCs versus IHCs have also been proposed.[39] For example, von Békésy noted that the shearing action between the two membranes could reduce the displacement amplitude of the stimulating energy while increasing its force.[4,40] Thus, the shearing action may serve to match the impedances of fluid and solid transmission media, just as the middle ear matches air and fluid impedances.

Electrical Potentials of the Cochlea Using gross microelectrode methods, three cochlear bioelectric events have been studied extensively: (1) the endocochlear potential (EP), (2) the cochlear microphonic (CM), and (3) the summating potential (SP). Whereas the EP is present at rest, the CM and SP appear only when sound stimulates the ear.

Endocochlear Potential. The EP is a constant (DC), +80 mV potential that can be recorded with an electrode in scala media.[41] The majority of evidence indicates that the stria vascularis generates the EP.[42] Because it is extremely sensitive to anoxia and chemical agents interfering with oxidative metabolism,[43] its existence likely depends on the active metabolic ion pumping processes of the stria vascularis. The anatomic distribution of the EP closely approximates the limits of the endolymphatic compartment formed by the tight junction boundaries shown in Figure 4–12. The augmentation of the voltage drop cross the apical ends of the hair cells that the EP provides is thought to be critically important in cochlear transduction.[44]

Cochlear Microphonics and Summating Potentials. When the appropriate stimulus is applied, most sensory end-organs generate bioelectric events called receptor potentials. These potentials differ from action potentials in that (1) they are graded rather than all or none, (2) they have no latency, (3) they are not propagated, and (4) they have no apparent post-response refractoriness.[45] The SP and CM form the receptor potentials of the cochlea and can be recorded indirectly from gross electrodes on the round window or directly from the fluid spaces within the organ of Corti.

The CM reproduces the alternating (AC) waveform of the stimulating sound (hence the name "microphonic"[46]). The SP, illustrated in the lowest inset at the top left of Figure 4–15, represents the DC shift that follows the "envelope" of the stimulating sound.[47] As depicted in Figure 4–15, both the CM and SP are generated across the hair-bearing end of the hair cells.[41,48] Based on experimental findings,[49] it is assumed that the generator site for the receptor potentials is at the apical tips of the stereocilia.

The results of intracellular recordings from IHCs and OHCs suggest that IHCs generate the SP and the OHCs generate the CM.[39] Generation of the SP requires some form of rectification of the acoustic waveform (ie, alteration from a waveform that oscillates above and below the baseline to a waveform that is entirely above or below the baseline). It is reasonably straightforward to model the

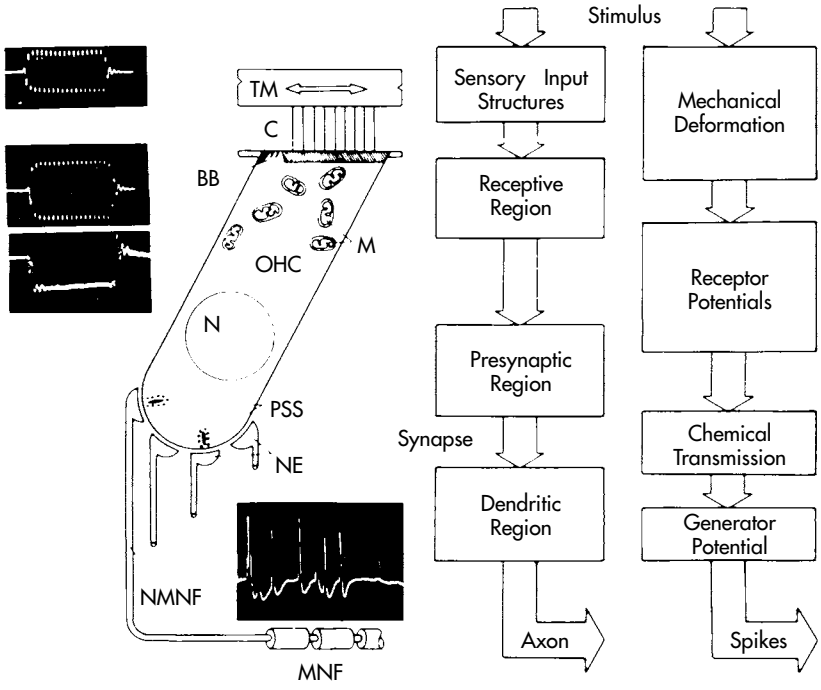


FIGURE 4–15. Generator sites of cochlear potentials (*upper left from top to bottom*: sound stimulus, cochlear microphonic, and summating potential) and functional diagram of the cochlear transducer mechanism (*right*). TM = tectorial membrane; C = cilia; OHC = outer hair cell body; N = cell nucleus; M = mitochondria; PSS = presynaptic structures; NE = afferent nerve endings; NMNF = nonmyelinated segment of nerve fiber; MNF = myelinated segment of nerve fiber; BB = basal body. Reproduced with permission from Dallos P. The auditory periphery. New York: Academic Press; 1973.

mechanical coupling of the free-floating IHC cilia in such a way as to produce the mechanical rectification. Thus, the concept that the IHCs generate the SP fits well with what is known about the morphology of this class of hair cell. Other evidence, however, derived from recording in the fluid spaces of the cochlea with OHCs selectively damaged by kanamycin,[50] suggests that both the IHCs and OHCs contribute to the volume-recorded SP and CM. Indeed, this latter notion represents the current consensus of the field.[51]

Generation of the CM can be understood by examination of the resistance battery model of Davis[52] depicted in Figure 4–16. In this conceptualization, the EP serves as the "battery" that provides the driving force to move current through the high resistance of the reticular lamina in which the apical ends of the hair cells are embedded. The traveling wave displaces the basilar membrane, resulting in the deflection of the stereocilia and changes in hair cell resistance. Thus, hair cells acting as variable resistors modulate current flow across the reticular lamina to produce a time-varying potential, the CM, which follows the waveform of the input stimulus.[3]

Role of the Cochlear Potentials in Stimulus Transduction. Wever and Bray's discovery of the CM[53] led to the hypothesis that auditory nerve fiber endings are stimulated electrically.[54] However, studies with the transmission electron microscope (TEM) revealed that the morphology of the interfaces between the afferent nerve endings and the hair cell bases was unmistakably that of a chemical synapse. The chemical synapse-like structures found there include (1) a synaptic cleft (ie, a uniform space between the hair cell and nerve-fiber membranes), (2) synaptic vesicles, and (3) a synaptic bar in the form of an electron-dense disk surrounded by vesicles that resembles the synaptic ribbon of the retina.[55]

The presence of subcellular organelles consistent with a chemical synapse at the hair cell/auditory nerve junction supports the current view that a chemical transmitter released by the hair cell stimulates the auditory nerve endings.[55] Thus, cochlear receptor potentials are probably either directly involved in the cause-and-effect chain leading to chemical stimulation of the auditory nerve or are intimately related to a process that is directly involved but does not stimulate the auditory nerve fibers electrically.[52]

Current thinking about cochlear transduction, which originates primarily from the studies of Hudspeth and colleagues,[56] is illustrated in Figure

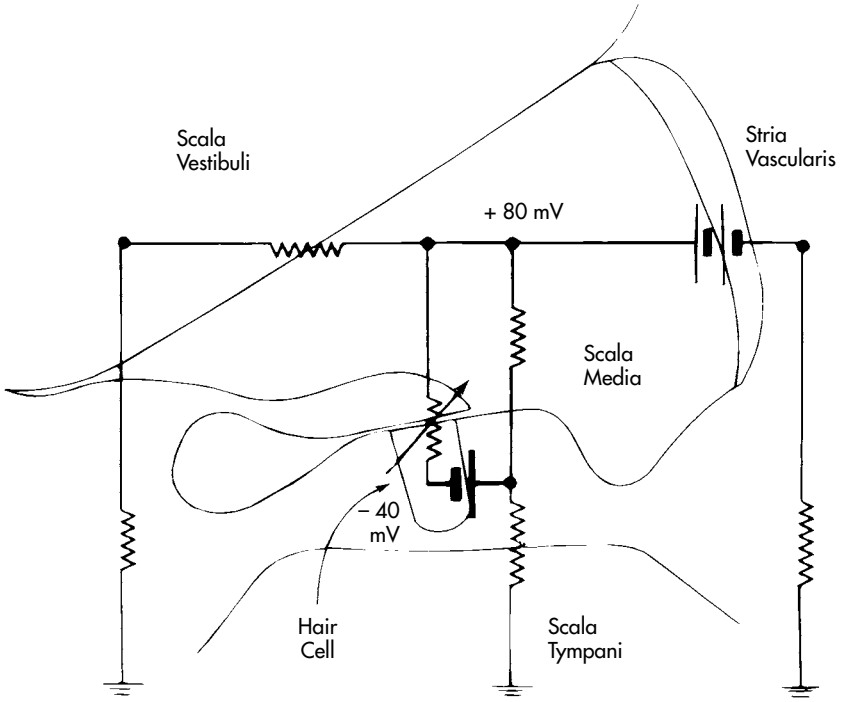


FIGURE 4–16. Diagram of Davis' battery model of transduction in which the positive endocochlear potential and negative intracellular potential provide the force to drive current through the variable resistances at the top of the hair cells. Reprinted with permission from Pickles J.[3]

4–17. In this formulation, displacement of the hair cell bundle opens the transduction channels located at the tips of the stereocilia to allow K^+ to flow into the cell. This influx of K^+ depolarizes the cell, causing calcium (Ca^{2+}) channels at the base of the hair cell to open, thus admitting Ca^{2+} into the cell. The Ca^{2+} ions, in turn, stimulate the transmitter vesicles to fuse with the hair cell membrane and release transmitter into the synaptic cleft. Transmitter substance then diffuses across the synaptic space to initiate action potentials in the adjacent auditory nerve fibers.

Cochlear Transmitters One area of intense research interest in the study of cochlear transduction leading to intercellular communication has been the identification of the afferent transmitter substance that is released onto the primary afferent neurons at the bases of the IHCs and OHCs. In combination with these studies, other efforts have attempted to characterize the efferent neurotransmitter released on the hair cells, and on afferent endings terminating on hair cells, by efferent neurons originating in the brainstem. Recent evidence indicates that the afferent neurotransmitter is probably a single excitatory amino acid, or a structurally related compound, which is responsible for initiating auditory nerve action potentials. Besides this chemical transmitter substance, other chemicals, called neuromodulators,[57] that influence the action of the transmitter are also believed to be released into the synaptic cleft.[58]

In the search for neurotransmitters, several criteria have been established that are requisite for the identification of transmitter substances.[59] These criteria require that (1) the transmitter must produce the same response when applied to the synapse as the natural stimulation of presynaptic elements, (2) extraneous substances that alter natural synaptic transmission such as blocking agents should produce the same effect on the transmitter candidate, (3) stimulation of presynaptic elements should release the transmitter substance, (4) the transmitter substance must be shown to exist presynaptically, (5) enzymes responsible for the synthesis of the transmitter candidate must be present, and (6) a mechanism must be demonstrated that can deactivate the transmitter once it has been released into the synaptic cleft. To date, concrete evidence for the auditory transmitter in the mammalian cochlea is scanty when compared with the findings of other studies of the central nervous system. That is to say, all of the criteria have yet to be met for any candidate afferent transmitter substance. However, based on our present ability to satisfy the above criteria, one of the

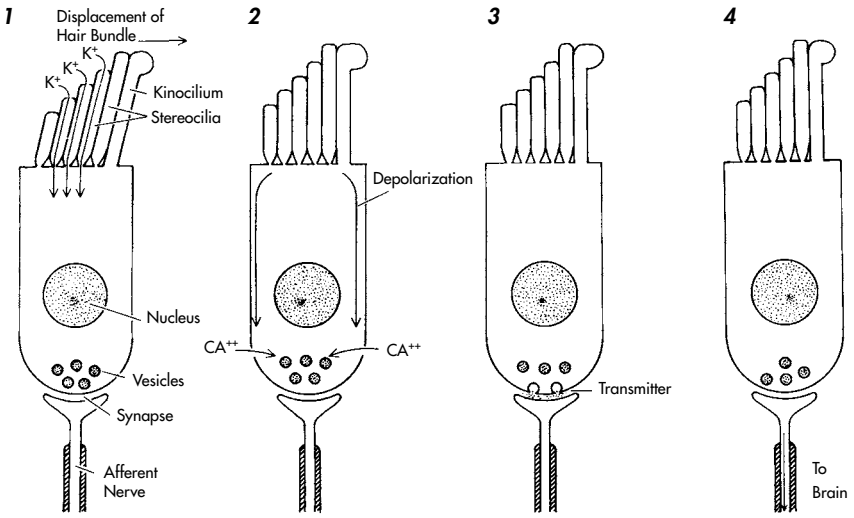


FIGURE 4–17. Schematic representation of the major steps involved in stimulus transduction in hair cells. Deflection of the hair bundle (1) opens the transduction channels to allow K^+ to flow into the hair cell. This results in a reduction of potential difference or depolarization. Depolarization spreads instantly to the lower part of the cell (2), causing Ca^{2+} channels to open. Ca^{2+} ions (3) cause transmitter vesicles to fuse with the basal part of the cell membrane. Fusing vesicles release transmitter substance into the synaptic cleft. (4) Transmitter diffuses across the synaptic cleft to initiate an action potential in the adjacent auditory nerve fiber. Reproduced with permission from Hudspeth AJ.[56]

most likely afferent transmitter substances is believed to be the excitatory amino acid glutamate,[60] and aspartate, too, represents a more recently documented potential candidate.[61]

Documentation regarding the efferent transmitter in the mammalian cochlea is considerably stronger with the most persuasive evidence favoring acetylcholine (ACh). The strongest support for ACh comes from a set of experiments demonstrating that anticholinergic compounds block the effects of efferent stimulation but do not influence afferent cochlear activity.[62] From other histochemical and immunostaining studies of the mammalian cochlea, there is considerable evidence for a GABAergic (ie, γ-aminobutyric acid) efferent innervation of the OHCs, as well as the IHC afferents, including (1) measurements of the uptake of tritiated GABA,[63] (2) immunostaining for glutamic acid decarboxylase,[64] and (3) immunostaining for GABA.[64] Interestingly, evidence that isolated OHCs stain for GABA[65] and that GABA application to isolated OHCs induces membrane hyperpolarization and membrane elongation[66] infers that the functional consequences of this candidate transmitter substance is inhibition of the electromechanical transduction process. For a more in-depth treatment of studies of cochlear transmitters, readers are encouraged to consult an excellent review.[67]

Coding in the Cochlea The cochlea must encode acoustic features into properties of neural activity. The principal acoustic parameters to be encoded are frequency, intensity, and temporal pattern, whereas the basic biologic variables available for neural encoding are place (ie, the location of the activated cell), amount of neural firing, and temporal pattern of firing. Because they are the percepts most studied, the cochlear encoding of frequency and intensity is addressed below. For a recent discussion of temporal processing, see Frisina.[68]

Frequency Coding. In the late nineteenth century, two opposing theories of frequency coding in the auditory periphery were proposed. These classic "place" and "frequency" theories have influenced subsequent thinking about cochlear frequency coding.[3] The place theory of Helmholtz held that the basilar membrane acts as if it were a series of tuned resonators, analogous to a set of piano strings. Each tuned resonator vibrates sympathetically to a different frequency and thus selectively stimulates a particular nerve fiber. Rutherford's frequency theory,

later termed "telephone" theory, proposed that all frequencies activate the entire length of the basilar membrane, which transmits, essentially unchanged, the temporal pattern of the auditory stimulus. According to the telephone theory, it remains for more central neural structures to "decode" these temporal patterns to educe the features of the acoustic stimulus.

Evidence for Place Coding: Mechanical and Neural Tuning Curves. Von Békésy used optical methods to make the first direct observations of the mechanical place analysis of stimulus frequency in the cochlea.[4] Subsequent observations using more sensitive techniques (eg, Mössbauer radioactive source, capacitive probe, laser interferometry) confirmed the general nature of von Békésy's results but, as discussed below, modified his conclusions with respect to several important details.[69]

The principal features of the mechanical responses that von Békésy observed are illustrated in Figure 4–18. Each pure-tone cycle elicits a traveling wave that moves along the cochlear partition from base to apex.[4] As it travels, the wave's amplitude increases slowly, passes through a maximum, and then declines rapidly (Figure 4–18, A). As the frequency of the stimulating tone is increased, the maximum of the traveling wave moves basally toward the oval window.

The inset at the bottom of Figure 4–18, B, illustrates the tuned behavior of the vibration of a specific locus on the cochlear partition that results from

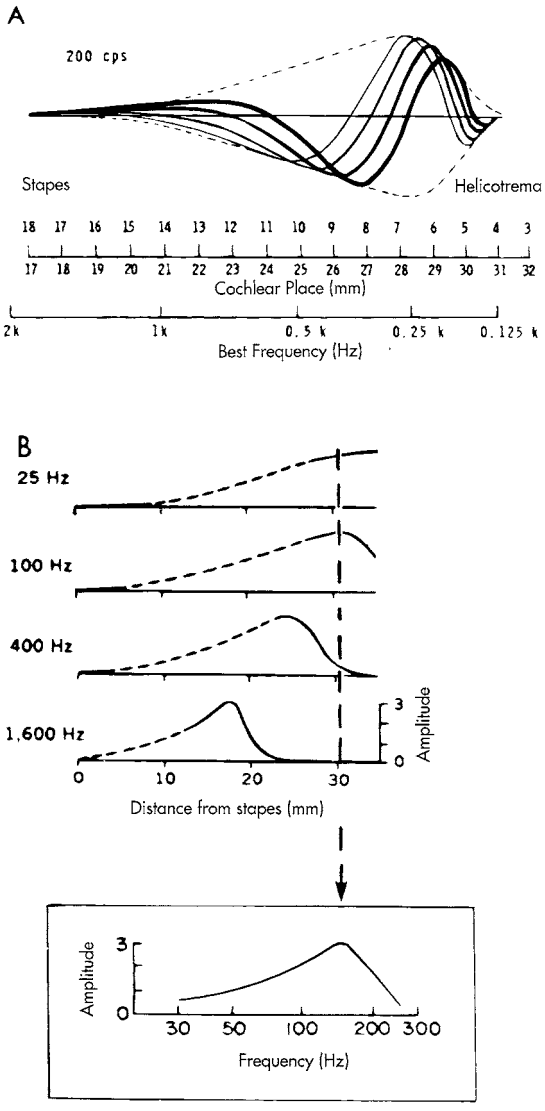


FIGURE 4–18. Mechanical place coding in the cochlea. *A*, The cochlear partition is "uncoiled," showing the traveling wave response to a pure-tone stimulus. Vertical distance represents the amount of cochlear partition displacement. The four *progressively darker lines* show cochlear partition positions at three successive instants during one cycle of a 200 Hz stimulation tone. *Darker lines* represent later points in time. The *fine dashed line* shows the "envelope" of cochlear partition displacement. Scales at the bottom show linear distance along the cochlear partition measured from helicotrema (*upper scale*), from stapes (*middle scale*), and also in terms of one commonly used cochlear partition "frequency map" (*bottom scale*). Reproduced with permision from Greenwood DD. Critical bandwidth and the frequency coordinates of the basilar membrane. J Acoust Soc Am 1961;33:1344. Note the expected peak in displacement pattern envelope at about the 0.2 kHz point on the frequency map. *B*, The top four curves are envelopes of traveling wave responses to pure-tone stimuli of varying frequencies. Each envelope depicts a point on the partition approximately 30 mm from the stapes (indicated by the *vertical dashed line*). Only the upper half of the envelope traced in *A* is shown. At bottom, the response of a single cochlear point is plotted against frequency. This is the mechanical tuning curve of this point. Reproduced with permission from von Békésy G.[4]

the traveling wave patterns generated by stimuli of different frequencies (four plots above). The bottom plot of Figure 4–18, B, called a cochlear mechanical tuning curve, was obtained by plotting the vibration amplitude of a single cochlear partition point against the exciting frequency. The frequency that generates the maximum vibration of a specific place on the cochlear partition is that point's characteristic frequency (CF). Thus, the plot at the bottom of Figure 4–18, B, illustrates the mechanical tuning curve for a place on the cochlear partition with a CF of 150 Hz. Mechanical tuning curves at all cochlear partition locations have the same general shape, that is a rapid falloff for frequencies above the CF (> 150 Hz) and a gradual drop-off for frequencies below the CF (< 150 Hz).

Microelectrode recordings from single hair cells and auditory nerve fibers yield an analog of the cochlear mechanical tuning curve. Figures 4–19 and 4–20, respectively, show examples of neural tuning curves obtained from primary auditory nerve fibers with traditional glass micropipets[70] and mechanical tuning curves for the 9 kHz point on the basilar membrane made by a laser-Doppler vibrometer for a number of vibration criteria.[71] Essentially, neural tuning curves are plots of response threshold against frequency (see Figure 4–19), and similar tuning curves can be recorded from single neurons throughout the auditory neural pathway. Note the relatively similar tuning capabilities of the nerve fiber with the 9 kHz CF (middle panel at the right of Figure 4–19), the most sensitive membrane velocity response (thick solid line of Figure 4–20) with respect to frequency selectivity (ie, the Q_{10dB} "tuning," at 10 dB above threshold, of ~2), and the tip-to-tail distance of about 80 dB.

The tuning curves of primary auditory nerve fibers have the same basic shape (ie, steep high-frequency slope, shallow low-frequency slope) as the mechanical tuning curves. Thus, the mechanical place code of the cochlea is clearly imprinted on the auditory nerve's neural response pattern.

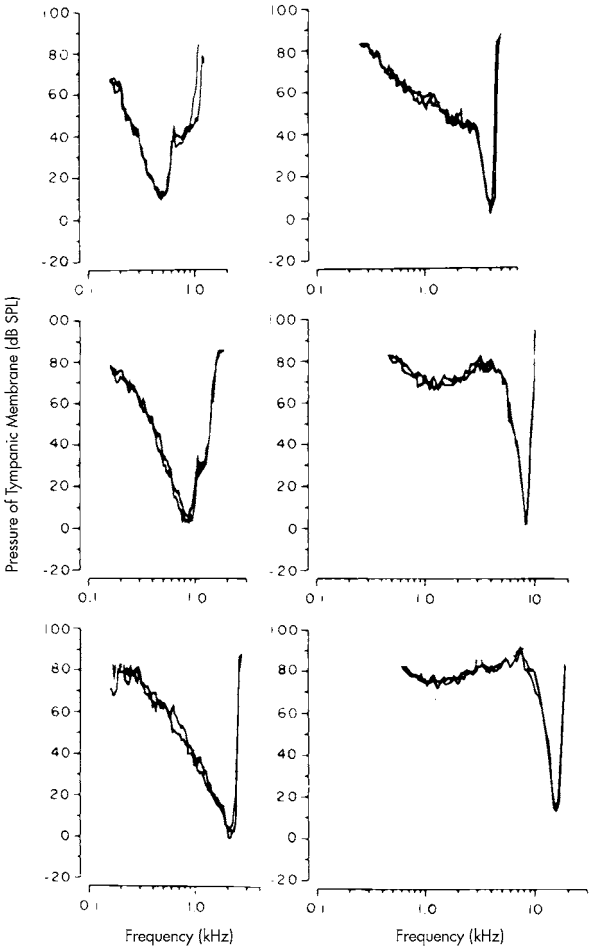


FIGURE 4–19. Representative tuning curves (frequency threshold curves) of cat single auditory nerve fibers are shown for six distinct frequency regions. In each panel, two fibers from the same animal, of similar characteristic frequency and threshold, are shown, indicating the constancy of tuning under such circumstances. Reproduced with permission from Liberman MC and Kiang NYS.[70]

Cochlear Tuning. Two characteristics of cochlear tuning are critical to the determination of its location and mechanism. First, the process by which filtering takes place is too rapid to permit a neural delay; thus, there is no possibility that tuning is sharpened by some sort of neural "lateral inhibition" analogous to that occurring at higher levels in the auditory pathway (see below).[72] The second important characteristic of cochlear filtering is that it is physiologically vulnerable. Almost all damaging agents, including hypoxia,[73] ototoxic drugs,[74] local mechanical damage,[75] and acoustic trauma,[76] detune the neural tuning curves so that they closely approximate the broader mechanical tuning curves. As Figure 4–21 shows, the detuning by damaging agents occurs not only for neural tuning curves but also for mechanical tuning curves, SP tuning curves measured with electrodes placed intracellularly in IHCs, and psychophysical tuning curves from humans with

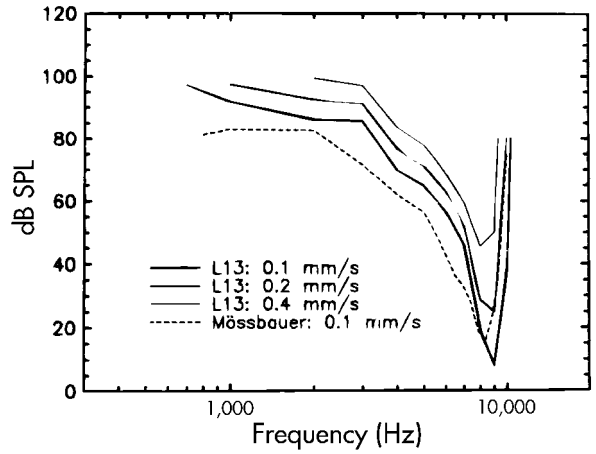


FIGURE 4–20. Isoresponse contours (tuning curves) for basilar membrane velocity responses to tone pips recorded from the chinchilla with the laser-vibrometer (*solid lines*). The ordinate indicates the sound pressure level required at any particular frequency to elicit a given velocity amplitude (0.1, 0.2, or 0.4 mm/s, indicated as the parameter). For the laser-vibrometry responses, sound pressure levels were interpolated logarithmically using a series of isointensity contours for basilar membrane responses to tone pips. For comparison, a 0.1 mm/s tuning curve is shown, which is representative of results obtained using the Mössbauer technique (*dashed line*) at an approximately equivalent basilar membrane site in normal chinchilla. Reproduced with permission from Ruggero MA and Rich NC.[71]

cochlear deficits. Until recently, the vulnerability of mechanical tuning curves was a controversial point. The first published "modern" mechanical tuning curves measured using Mössbauer and capacitive probe techniques were tuned about as poorly as von Békésy's.[77] However, even at this early stage, the sharpness obtained by different laboratories differed.[78]

It has now become clear that the sharpness of cochlear mechanical tuning is extremely vulnerable and that, even when great care is taken, the surgical and other manipulations necessary to obtain mechanical tuning curves in experimental preparations unavoidably cause broadening of the mechanical frequency response.[69] Cochlear mechanical tuning curves obtained under conditions in which extreme precautions were taken to minimize cochlear damage have demonstrated tuning that is as sharp as that of neural tuning curves.[71] In addition, the existence of nonlinear behavior in the normal cochlear mechanical response has been confirmed. Thus, the sharpness of tuning observed at the primary neuron level is already accomplished at the level of the mechanical traveling wave.

The question remains, however, of how the physiologic vulnerable sharpening of mechanical tuning is accomplished. One critical consideration in answering this question is the calculation by Kim and coworkers that mechanically tuning a location on the basilar membrane requires the local addition of mechanical energy.[79] The discovery that OAEs are produced by the healthy cochlea[80] and can be recorded simply from human and animal ear canals provided indirect evidence for the presence of such a mechanical energy generator within the cochlea. The results of further studies on OAEs in vertebrates and more recent investigations of the biophysical properties of isolated OHCs have implicated micromechanical processes at the level of the OHCs as being responsible for the normal frequency selectivity of the cochlea.

The OAEs are traditionally classified into two general categories: those that are elicited with deliberate acoustic stimuli (evoked OAEs) and those that occur in the absence of stimulation (spontaneous OAEs). All OAEs can be simply recorded using conventional averaging techniques by a sensitive subminiature microphone assembly inserted snugly into the external auditory canal. Over the past several decades, four types of emissions, illustrated in Figures 4–22 and 4–23, have been studied extensively. These include the spontaneous OAE (SOAE), which is measured in the absence of acoustic stimulation (see Figure 4–22, A); the stimulus frequency OAE (SFOAE), which is evoked by a low-level, continuous pure tone (see Figure 4–22, B); the transient evoked OAE (TEOAE), which is elicited by clicks or tone pips (see Figure 4–22, C) and distortion-product OAE (DPOAE), which is elicited by two simultaneously applied pure tones (see Figure 4–23) known as the lower- (f_1) and higher-frequency (f_2) primary tones.

The following facts support the notion that OAEs come from the cochlea and are related to the active frequency filtering process: (1) the frequency of the TEOAE varies considerably from subject to subject, and across subjects there is a systematic relationship between the emission's delay (ie, latency) and its frequency.[81] This suggests that the origin of OAEs is the cochlear bandpass filtering mechanism; (2) stim-

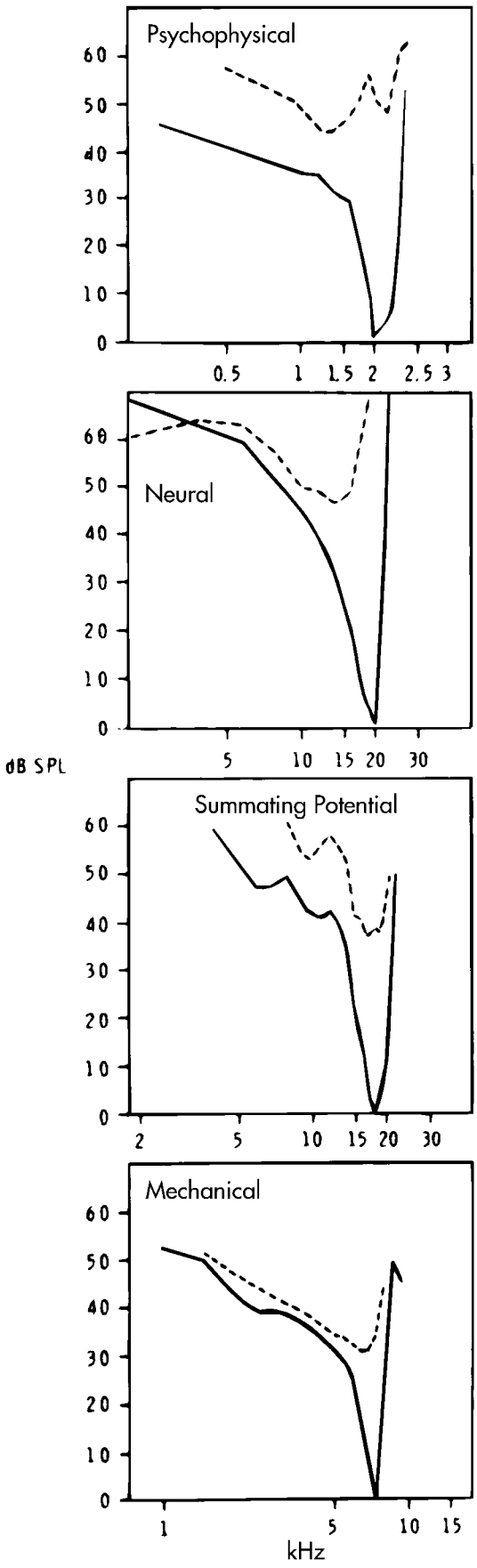


FIGURE 4–21. Collection of tuning curves from several authors illustrating "detuning" with damage. The tuning curves illustrated were obtained from humans (psychophysical) and from animals at the primary auditory neuron (neural), cochlear receptor potential (summating potential), and basilar membrane (mechanical) levels. The psychophysical tuning curves were obtained by a tone-on-tone masking procedure. Reproduced with permission from Gulick WL. Hearing, physiology and psychophysics. New York: Oxford University Press; 1971. The *dashed curve* is from a hearing-impaired listener; the *solid curve* is from a normal listener. The "notch" in the detuned hearing-impaired curve may be a technique-related artifact created by the detection of combination tones or beats made by combining masker and test tones. The neural tuning curves were obtained from guinea pigs before (*solid curve*) and 20 minutes after (*dashed curve*) acoustic trauma.[76] The summating potential curves are 10 μV isoamplitude curves obtained from intracellular hair cell recordings. The *dashed line* is an example of an "insensitive" cell (presumably damaged in the course of exposure); the *solid line* is from a "sensitive" cell.[39] The mechanical tuning curves illustrate the range of results obtained from different animals in the course of Rhode's Mössbauer technique measurements. Reproduced with permission from Rhode WS. Cochlear partition vibration—recent views. J Acoust Soc Am 1980;67:1696.

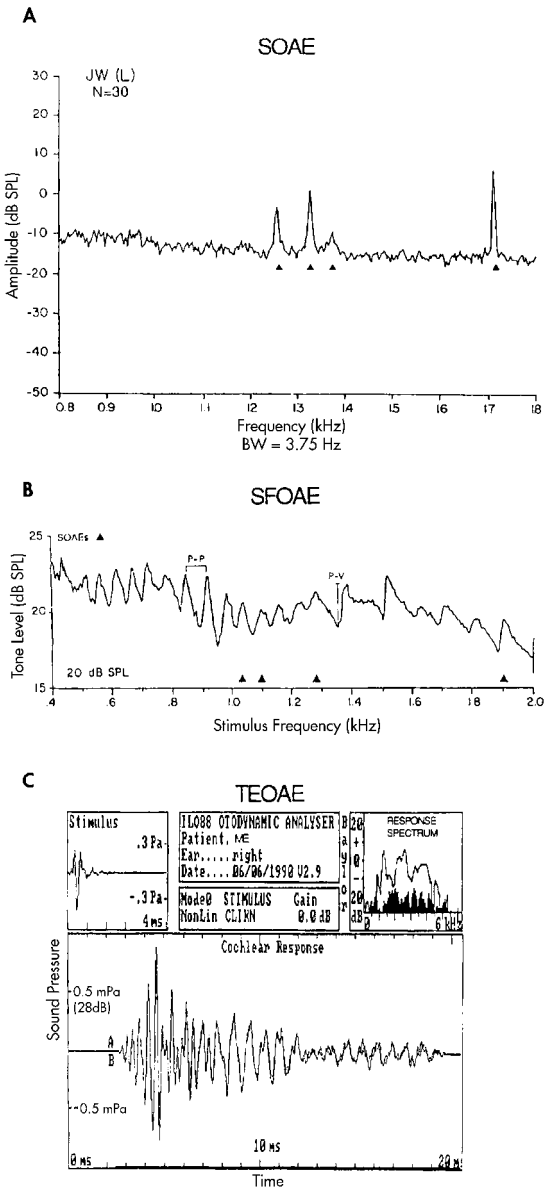


FIGURE 4–22. Examples of three types of otoacoustic emissions (OAEs). *A*, Spectrum of the sound pressure level in the ear canal of a 26-year-old woman showing four spontaneous OAEs (SOAEs) within the 0.8 to 1.8 kHz frequency range. Spectral average was based on 30 samples. *Arrowheads* = frequencies of SOAEs. *B*, Ear canal record for a 23-year-old man evoked by a continuous 20 dB SPL pure tone swept slowly (150 s) from 0.4 to 2 kHz. Peaks (P) and valleys (V) in the trace indicate regions of stimulus frequency OAE (SFOAE) activity in which the elicited returning emitted response moves in and out of phase with the sweeping tone. *Arrowheads* = frequencies of coexisting SOAEs. *C*, An example of the ILO88 Otodynamic Analyser record displayed on screen and in hard-copy form, which includes visual representations of the temporal course of two averages of the poststimulus response acquired by separate buffers (*below*), the form of the eliciting stimulus (*above left*), in this case a click based on an 80 μs rectangular pulse with a peak level of about 82 dB SPL, and a spectrum (*above right*) of the response showing both the emission (*unshaded*) and background noise (*shaded*) components. Basic information concerning the patient and stimulus mode is noted (*above center*). Reproduced with permission from Lonsbury-Martin BL, Martin GK, Balkany T. Clinical applications of otoacoustic emissions. In: Lucente FE, editor. Alexandria (VA): Am Acad Otolaryn Head Neck Surg; 1994.

ulation of the crossed olivocochlear bundle (OCB), which preferentially provides efferent innervation to the OHCs, modulates the magnitude of OAEs[82,83]; and (3) the emission is extremely sensitive to the detrimental effects of cochlear pathology, acoustic trauma, and ototoxic drugs.[83,84] The most parsimonious interpretation of the existence of the various OAE types, their duration, bandwidth, and delayed-onset characteristics and their relationship to the activity of the cochlear efferent system is that they originate in the OHCs, which possess a mechanical energy that generates the frequency-selective output of the cochlea.

Brownell and Zenner and their colleagues were the first to show in in vitro preparations of mammalian hair cells that OHCs display electromotility.[85,86] Using mechanical trituration to dissociate single OHCs from other organ of Corti tissues and video-enhanced imaging to visualize the isolated OHC directly, these investigators used both intracellular and transcellular electrical stimulation[85] or pharmacologic (K^+, cholinergic chemicals) agents[86] to elongate and shorten OHCs from their resting lengths. Because it has been demonstrated that hair cells possess both actin and myosin,[87] it was originally presumed that some aspect of the active motile mechanism is mediated, as in muscle, by interactions between these molecules. However, as more details developed about OHC electromotility, including the

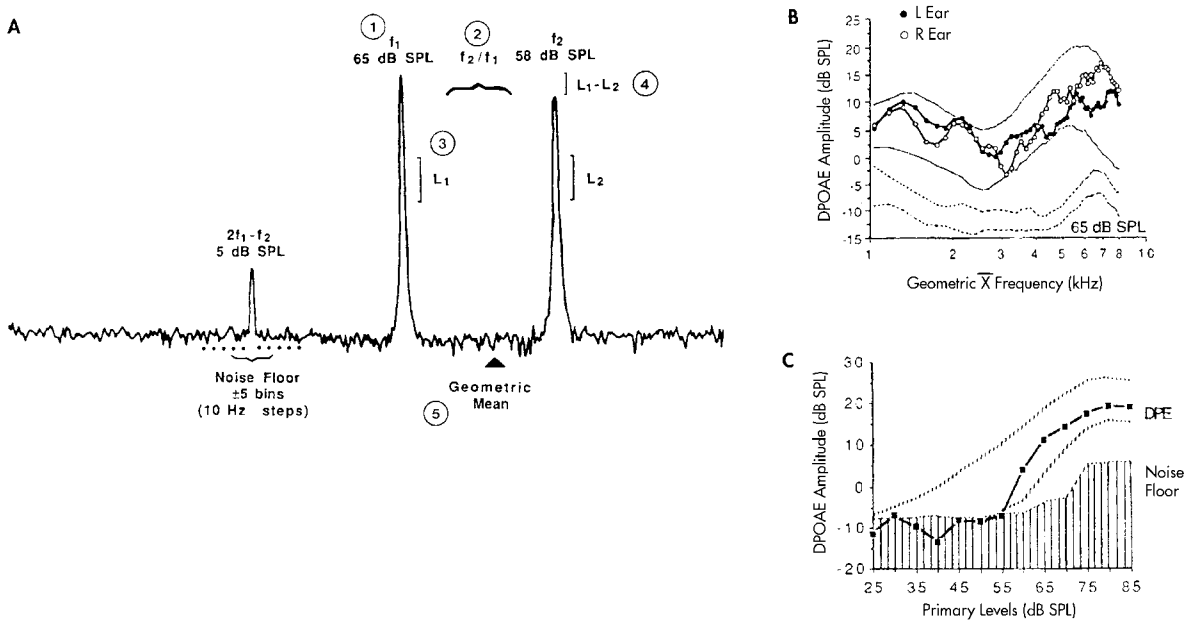


FIGURE 4–23. Distortion-product otoacoustic emissions (DPOAEs). *A*, DPOAE at $2f_1$-f_2 is shown in the spectrum of the ear canal sound. Note the position of the geometric mean frequency (5) with respect to the primary tones (f_1, f_2) and the frequency range over which the noise floor (NF) is measured relative to that of the emission (*dotted region* surrounding the DPOAE frequency). *B*, A typical DPOAE frequency/level function or "DP-gram" elicited by 75 dB SPL primaries is shown for the right (*open circles*) and left (*solid circles*) ears of a normal-hearing individual. The pairs of long (*above*) and short (*below*) *dotted lines* represent the ranges of average activity and their related NF levels, respectively, associated with normal function. *C*, A typical response/growth or input/output curve depicts the growth of the DPOAE (DPE) with progressive increases in the levels of the primary tones. The *pair of dotted lines* above depicts the normal range of emission levels, whereas the *striped lines* below represent the average NF. Reproduced with permission from Lonsbury-Martin BL, Whitehead ML, Martin GK. Clinical applications of otoacoustic emissions. J Speech Hear Res 1991;34:964–81.

knowledge that they mechanically vibrate at frequencies up to at least 30 kHz, it became clear that a very fast-acting motor molecule,[88] or some instantaneous physical phenomenon induced by an electrokinetic process,[89] formed the basis of the ability of the OHC to vibrate in response to changes in the receptor potential. Most recently, Zheng and colleagues tentatively identified a protein they called prestin as the motor protein of the OHC.[90] The detailed follow-up studies of these investigators further support the proposition that this membrane protein is the motor molecule responsible for the electromotility of OHCs.[91]

Inner versus Outer Hair Cell Function. Figure 4–24 illustrates the salient characteristics of IHC and OHC anatomy, and Figures 4–25 and 4–26 illustrate their afferent and efferent innervation, respectively. There are many differences between IHCs and OHCs in terms of their morphology, biochemistry, physiology, and afferent and efferent innervation patterns. Phylogenetically, OHCs are much younger than IHCs, being present only in the mammalian cochlea. In keeping with their relative phylogenetic youth, OHCs develop later embryologically,[92] are more easily compromised by various damaging agents, and have many unique characteristics that distinguish them from hair cells in other mechanoreceptor systems.

It seems clear that the OHCs and IHCs are functionally different, but our concept of the nature of this difference is changing. The classic view that the relatively insensitive IHC system carries the frequency place code, whereas the OHCs comprise a sensitive low-level detector system that has poor place-coding ability, has been abandoned in favor of more recent notions based on the active biomechanical functioning of the organ of Corti. That the OHCs possess effector abilities supports the pro-

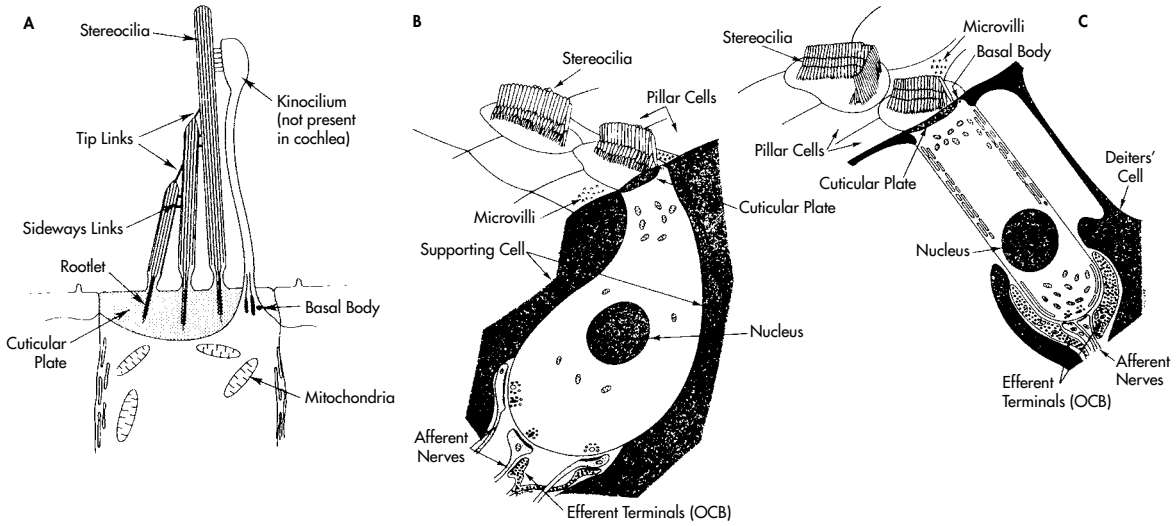


FIGURE 4–24. Schematic drawing of cochlear hair cells. *A*, The common structures on the apical portion of acousticolateral hair cells include rows of stereocilia that are graded in height and joined by cross-links. The tip links may be involved in transduction by opening the transducer channels. The kinocilium is not present in the mature cochlea, although it is present in vestibular hair cells. *B* and *C*, Inner hair cells are shaped like a flask (*B*), and outer hair cells are shaped like a cylinder (*C*). OCB = olivocochlear bundle. Reproduced with permission from Pickles JO. An introduction to the physiology of hearing. 2nd ed. New York: Academic Press; 1988.

posal of Kim that these receptors are bidirectional transducers that are capable of converting acoustic energy into neural energy (ie, mechanoelectrical transduction) and electrical into mechanical energy (ie, electromechanical transduction).[93] Kim elaborated on this notion by postulating the existence of two distinct but parallel cochlear subsystems that use OHCs as the modulators and IHCs as the carriers of auditory information.[93] Table 4–1 presents a summary of a comparison of the two subsystems.

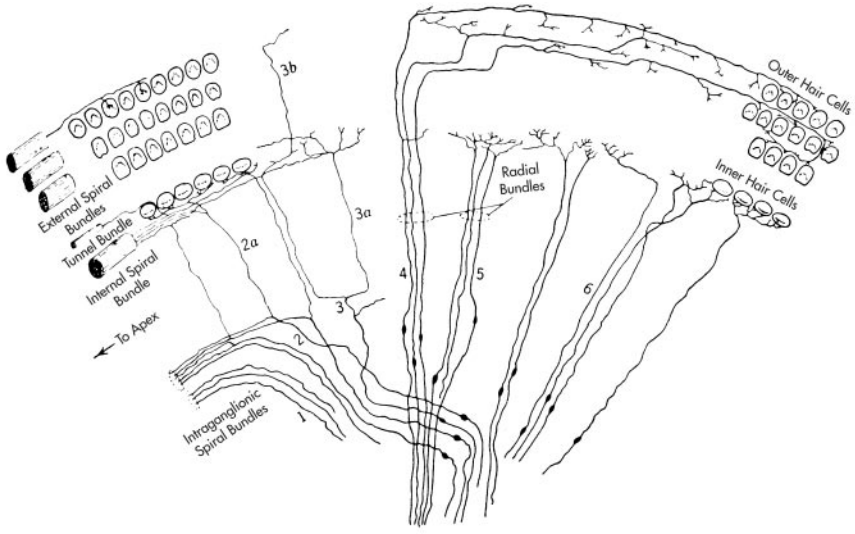


FIGURE 4–25. Diagram of the cochlea's afferent innervation pattern. The view is through Reissner's membrane, looking down on the organ of Corti. Principal fiber bundles are 1 and 2, intraganglionic spiral bundles (fibers labeled "1" are efferent olivocochlear fibers); 2a and 3a = internal spiral fibers; 4 = external spiral fibers, traveling in radial bundle to innervate outer hair cells (OHCs); 5 and 6 = radial fibers, innervating inner hair cells (IHCs). The "V" shape of cilia pattern on the OHCs and the shallower "U" pattern on the IHCs are shown in the upper corners of the diagram. Reproduced with permission from Wever EG.[94]

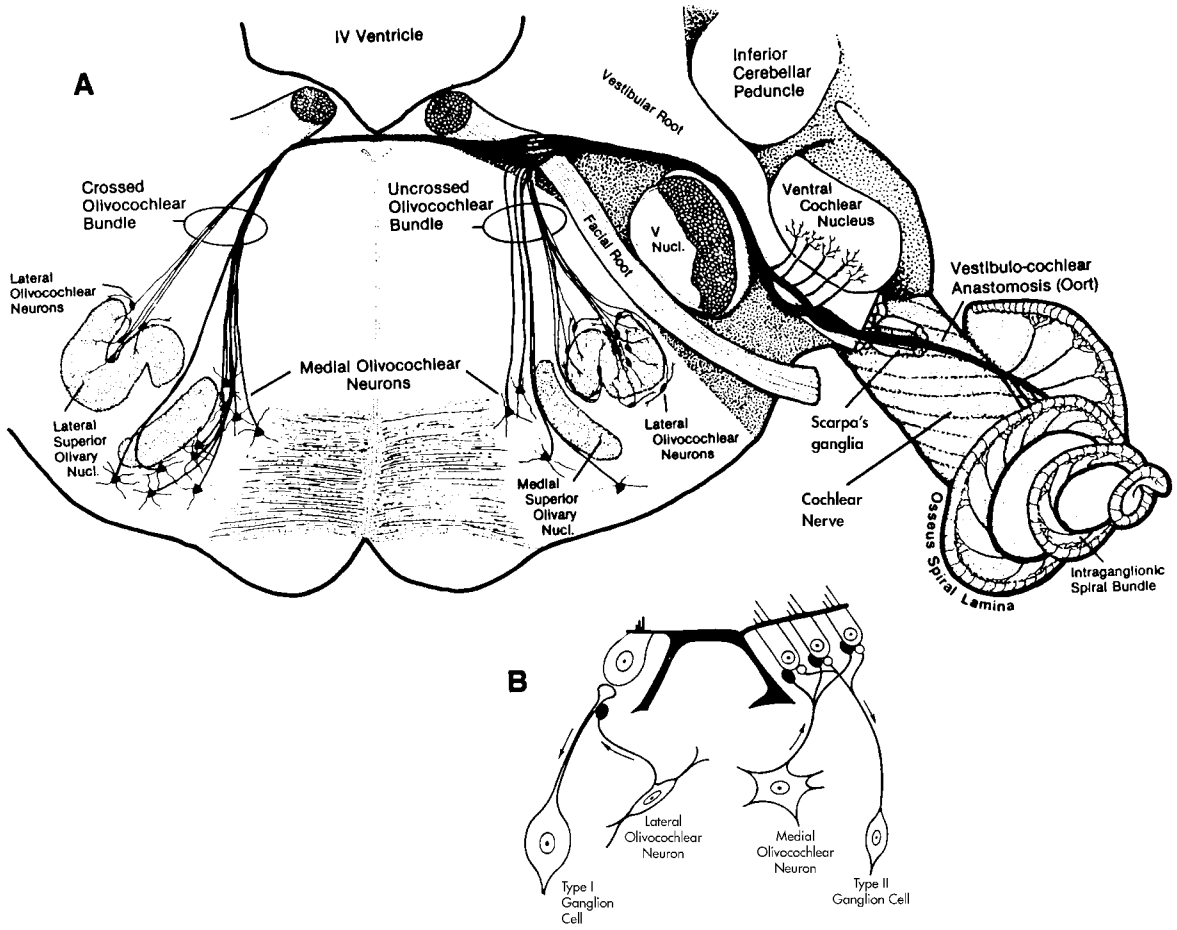


FIGURE 4–26. The descending auditory system. *A*, Origination in the brainstem, distribution, and termination within the cochlea of the olivocochlear bundle efferent fibers of the cat. *B*, Organization of the lateral and medial olivocochlear (OC) systems. Lateral OC neurons project to the region beneath the inner hair cells, where they form axodendritic contacts on the dendrites of type I spiral ganglion cells. Medial OC neurons project to the region beneath the outer hair cells and synapse directly with them. *A* reproduced with permission from Warr WB et al.[95] *B* reproduced with permission from Spangler KM and Warr WB.[96]

There is good evidence that the vulnerable filter sharpening described previously involves some kind of interaction between the IHCs and OHCs including the following: (1) selective damage to OHCs causes detuning of neural tuning curves (Figure 4–27),[97] which presumably are from IHC fibers; (2) crossed OCB stimulation, which presumably affects only OHCs (see below), strongly depresses the auditory nerve's click-evoked, whole-nerve, or compound action potential (CAP), which apparently originates entirely from IHC fibers; (3) crossed OCB stimulation also reduces the amplitude of the receptor potential recorded intracellularly from IHCs[98]; (4) crossed OCB stimulation as well reduces the firing rate and synchronization index[99] and shifts the dynamic range of auditory nerve fibers[100]; and (5) a trapezoid acoustic stimulus elicits from a auditory nerve fiber a spike train with a complex time course, suggesting combined IHC and OHC influences, whereas selective damage to OHCs by kanamycin treatment simplifies the trapezoidal response pattern in a way compatible with isolated IHC influence.[38]

The precise mechanism by which OHCs influence the response of IHCs is unknown at this time. However, because the physiologically vulnerable cochlear filter-sharpening process appears to involve basilar membrane vibrations, the mechanism requires that the OHCs directly influence the mechanical vibration of the basilar membrane. A corollary to this conclusion is that the OAEs recorded from the human ear canal reflect OHC function. Indeed, one intensely studied subfield of

TABLE 4–1. Differences Between Inner and Outer Hair Cells

	Inner Hair Cells	*Outer Hair Cells*
Anatomy (see Figure 4–24)		
Cell shape	Flask (35 μm × 10 μm)	Cylinder [25 μm (base) to 45 μm (apex) × 7 μm]
Relation to supporting cells	Closely approximated to inner phalangeal cell processes	Contact Deiters' supporting cells only at apical and basal ends and are surrounded by large perilymph-filled spaces of Nuel
Cilia	Longitudinally oriented shallow curve	Radially oriented "V"-shaped curve
	Not obviously attached to tectorial membrane	Attached to tectorial membrane
Location of cell nucleus	Central	Basal
Organelles	Resemble hair cells in other systems	Several unique features including submembrane cisternae, numerous mitochondria parallel to cell membrane, Hensen's bodies
Afferent innervation (see Figure 4–25)	Convergent via radial fibers	Divergent via spiral fibers
	95% of fibers in auditory nerve come from IHCs	
	All published nerve-fiber studies are probably from IHCs	No convincing neurophysiologic demonstration of population of nerve fibers from OHCs
	Nerve endings plentiful and show typical chemical synapse morphology	Nerve endings relatively sparse; morphology differs from chemical synapse
Efferent innervation (see Figure 4–26)	Small neurons near LSO	Large neurons near MSO
	Primarily ipsilateral via uncrossed olivocochlear bundle	Primarily contralateral via crossed olivocochlear bundle
	Small endings primarily on afferent nerve terminals	Large endings on OHCs
	Acetylcholinesterase activity poorly visualized	Acetylcholinesterase activity easily visualized
	Enkephalin-like immunoreactivity	Aspartate aminotransferase immunoreactivity
	Distributed evenly along cochlear length	Distributed preferentially in middle and basal cochlea
Physiology		
Resting membrane potential	–35 to 45 mV	–70 mV
Basal cells		
AC receptor potential	0.6 mV	3 mV
DC receptor potential	12 mV	Immeasurable

Continued

TABLE 4–1 ***Continued.*** **Characteristics of Duct Responding Units in Vestibular Nucleus**

	Inner Hair Cells	*Outer Hair Cells*
Apical cells		
AC receptor potential	10 mV	5 mV
DC receptor potential	5 mV	3 mV
Biochemistry		
Intracellular glycogen	Scarce	Plentiful

IHCs = inner hair cells; OHCs = outer hair cells; LSO = lateral superior olive; MSO = medial superior olive.

OAE research is focused on using emitted responses to evaluate the normality of central auditory processing based on the influence of the descending efferent system on OHC-generated activity.[101,102]

Clinical Consequences. The existence of a vulnerable cochlear filtering process and a related cochlear emission that can be recorded noninvasively have some important implications for our understanding of pathophysiologic processes in the cochlea. First, although it is unlikely that the majority of cases exhibiting severe tinnitus are related to spontaneous OAEs, emissions may elucidate the mechanism of at least some kinds of tinnitus.[103] Moreover, by separating the sensory from the neural aspects of ear dysfunction, OAEs may reveal the critical underlying causes of some common but puzzling otologic diseases such as Meniere's disease[104] and sudden idiopathic sensorineural hearing loss.[105] Second, as shown in Figure 4–28, detuning of the neural response in cochlear pathologic conditions may also explain recruitment.[106] As previously described, detuning caused by cochlear pathology eliminates the low-threshold, sharply tuned tip region of the tuning curve but preserves the high-threshold tail region. Elimination of the tip region raises threshold, but preservation of the tail region preserves neural responsiveness at high intensities. Thus, the loudness function (bottom right plot) is made abnormally steep because threshold is elevated. However, at high intensities, loudness is normal because a normal number of neurons are responding.

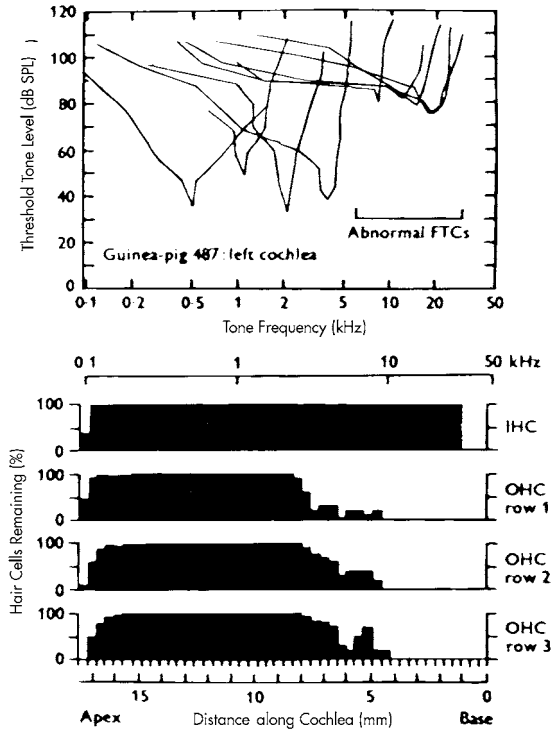


*FIGURE 4–27. C*orrelation between detuning of neural tuning curves and outer hair cell (OHC) damage. The results are shown from a kanamycin-treated guinea pig in which the inner hair cells (IHCs) were intact throughout the cochlea. In the high-frequency region, however, OHCs were damaged, and tuning curves (FTCs), presumably from fibers coming from this region, were detuned. This result suggests that the OHCs participate in the process by which cochlear tuning curves are sharpened. Reproduced with permission from Evans EF and Harrison RV.[97]

Evidence for Telephone Coding. The aforementioned evidence for place coding has not invalidated Rutherford's telephone theory because there is also considerable evidence for a telephone code at low frequencies. For example, cochlear tuning becomes progressively poorer as frequency is lowered, and below about 100 Hz, there are no cochlear partition amplitude maxima and no tuned auditory units.[107] Thus, the physiologically observed frequency place code becomes progressively worse as frequency is lowered.

A neural telephone code has been demonstrated that, in contrast to the place code, becomes

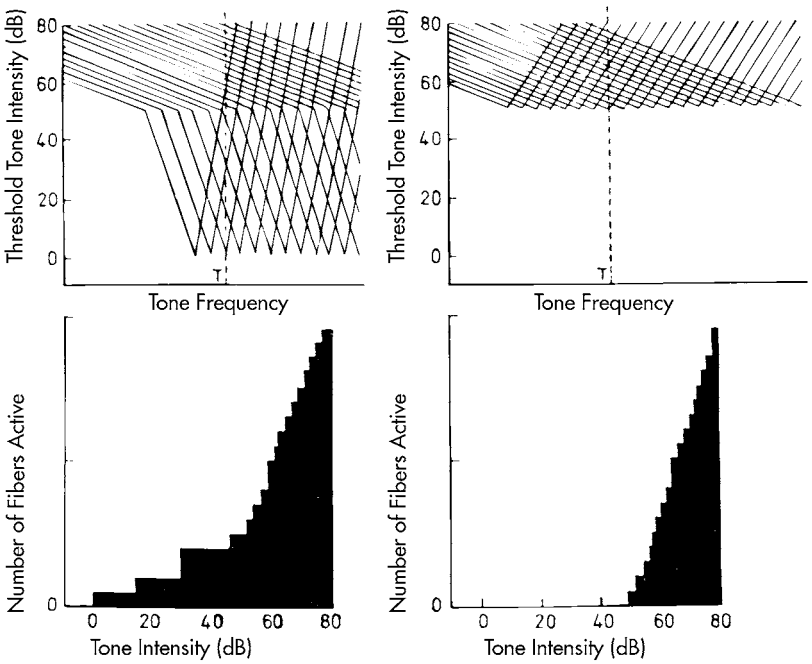


FIGURE 4–28. Possible explanation of recruitment (elevated threshold and abnormal growth of loudness) based on detuning with pathology. *Solid diagrams* at bottom depict increase in number of active auditory nerve fibers as tone level is increased in normal (*left*) and abnormal (*right*) ears. The tone's frequency (T) is indicated by the *dashed line* in the *upper diagram*. Because the effect of disease is to remove the sharply tuned "tips" but leave the broadly tuned "tails" relatively unaffected, the threshold of the pathologic cochlea is increased. However, as tone level is increased into the normal tail regions, the number of active fibers responding becomes normal; hence, loudness grows abnormally rapidly from the elevated threshold. Reproduced with permission from Evans EF.[107]

progressively better as frequency is lowered. Analyzing the spike discharges of single auditory nerve fibers to low-frequency stimulation demonstrated "phase-locking" to the individual cycles of the eliciting tone that preserves the temporal firing pattern.[108] Compiling large numbers of single-unit spikes into spike discharge histograms showed an impressive ability of single auditory unit discharges to reproduce the waveform of the stimulating sound. The upper limit of this phenomenon is generally estimated at about 4 kHz, but critical inspection of quantitative single-unit data suggests that phase-locking becomes poor at around 2.5 kHz.[109]

Another line of evidence supporting the importance of the encoding of phase or timing information comes from the outcome of studies of the responses of auditory nerve fibers to speech sounds. The work of Kiang addressed the limitations of fiber discharge rate place encoding in neurally representing the frequency components in speech signals, especially those of moderate to high intensities.[110] More recently, the single-unit population studies of Sachs and Young on the representation of speech sounds in hundreds of auditory nerve fibers have demonstrated that the temporal pattern of fiber discharges in the form of phase-locking provides considerable information about the frequency content of the stimulus at all levels of stimulation.[111] These population measures of temporal synchrony were based on Fourier transforms of period histograms from fibers with CFs near each frequency component of the sound. Such measures were combined to provide discharge rate profiles for groups of active nerve fibers. Because information concerning rate, timing, and place is represented in these measures, these spectra provide details concerning temporal fine structure from a localized cochlear region and thus represent a type of temporal place code.

"Telephone Place" Theory of Frequency Coding. One shortcoming of the telephone theory became apparent when neural refractoriness was discovered. This process imposed a physical limitation on nerve fiber firing at the very short time intervals associated

with high-frequency stimuli. Thus, to reconcile in part the effects of neural refractoriness on the telephone theory, Wever published an auditory frequency coding theory, which he termed the volley theory, but which has since become known as the "telephone place" or "frequency place" theory.[94] Wever advanced an interesting evolutionary argument in support of the telephone place code theory. Primitive ears (as found in fish, amphibians, and lower reptiles) are telephone coders that cannot analyze high frequencies because of the inherent limitation of the rate at which nerve fibers can respond. Therefore, as evolution progressed, place coding had to be added to allow analysis of high frequencies, which are important in sound localization and directionality.

Our use of sounds in communication is greatly aided by our keen sensitivity to the frequency region between 1 to 4 kHz, and although our sensitivity falls off beyond 4 kHz, we still depend on the higher frequencies for many of our discriminations of consonants and sharp transients. The appearance and elaboration of the place principle for frequency representation, therefore, was a major event in the evolution of the ear.[94]

Intensity Coding Loudness is the approximate subjective correlate of the physical dimension of sound intensity. It is generally assumed that the neural correlate of loudness is the amount of nervous activity, with the amount factor meaning the total number of action potentials delivered by a population of nerve fibers over a given time period. Thus, loudness is encoded as a combination of the number of fibers firing and the rate at which they are firing.

Figure 4–29 compares plots of stimulus level versus single auditory nerve fiber firing rates (left) and stimulus level versus subjective loudness (right). The agreement in the general shapes of the curves supports the view that the amount of auditory nerve firing encodes loudness. On closer scrutiny, however, it is apparent that firing rate changes only over a 20 dB sound intensity range, whereas loudness varies over a 100 dB range. This limited dynamic range is a general property of all primary auditory neurons.[112]

The obvious way to solve this loudness encoding problem is to assume that additional nerve fibers are "recruited" as stimulus level is increased. This process would increase the total number of dis-

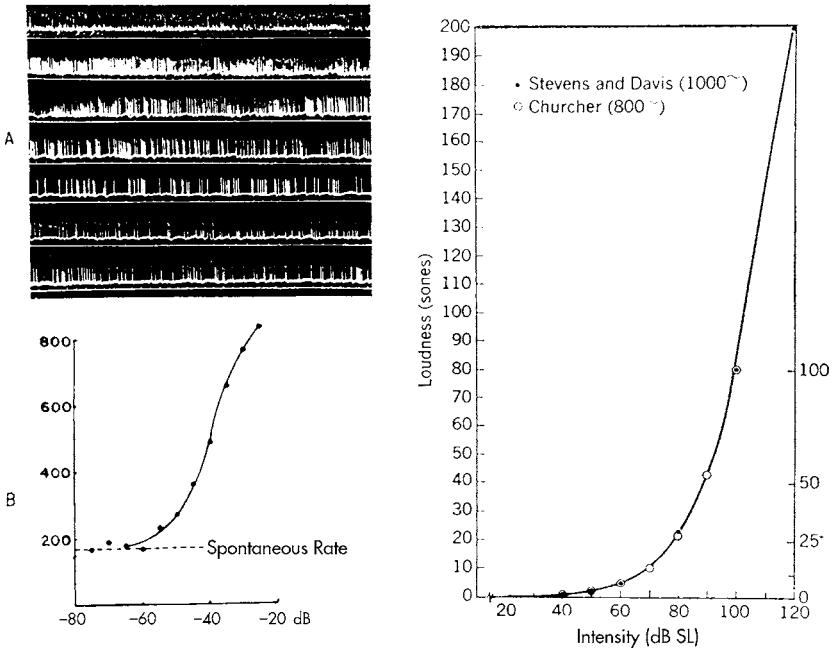


FIGURE 4–29. Effects of sound intensity on subjective loudness (*right*) and auditory nerve fiber firing rate (*left*). The shapes of the curves are similar, but the dynamic range of the single auditory nerve fiber is much narrower than that of the auditory system as a whole. Above the auditory nerve firing rate curve (*B*) are single-fiber responses to sounds of progressively increasing levels (*A*). Subjective loudness curve from Gulick (1971) noted in the legend for Figure 4–21. Auditory nerve curve reprinted with permission from Katsuki Y, Sumi T, Uchiyama H, Watanabe T. Electric responses of auditory neurons in cat to sound stimulation. J Neurophysiol 1958;21:569.

charges per unit time by increasing the number of fibers firing rather than the firing rate for an individual fiber. However, most studies of auditory nerve single units agree that the thresholds of the nerve fiber population fall within a relatively restricted stimulus range, that is, within 20 to 30 dB of behavioral threshold.[113] Thus, there is no possibility of recruiting additional fibers as sound level is increased more than 60 dB above threshold. The fiber recruitment hypothesis also creates problems for frequency place coding because adding new responding fibers as stimulus level is increased would quickly require the spread of activity to fibers of adjacent CFs. This tradeoff between the ability to recruit new fibers and the ability to place code is especially bothersome for explaining the perception of complex sounds for which multiple frequencies must be distinguished simultaneously. Thus, the dynamic range problem becomes apparent.[114] The dynamic range of auditory nerve fibers, that is, the range over which the spike rate changes as stimulus level is altered, is too narrow to account for (at least in simple terms) the extent over which the human ear discriminates sound levels.

Several subsequent observations shed light on the dynamic range problem. Evans and Palmer discovered that approximately two-thirds of the cells in the dorsal cochlear nucleus (DCN) had dynamic ranges that extend to well over 100 dB and thus could easily "code" loudness over the necessary sound intensity range.[115] However, the problem of the input pathway to these DCN cells remains to be resolved. In addition, some auditory nerve units have been found that do, after all, have wide dynamic ranges, but these are only a minority of the total auditory nerve fiber population.[116] Other studies have also demonstrated that relative degrees of phase-locking[117] of different frequency components in complex sounds could code the relative intensity levels of the frequency components. However, because phase-locking does not occur above 3 to 4 kHz, this mechanism could not account for the loudness coding of high frequencies.

Another approach to the dynamic range problem is the concept of an automatic gain control operating at the input of the auditory system.[114] Such a control mechanism would feed back a decreased gain command signal to the input to maintain the cochlear output within an acceptable operating range. Possibilities for such a control system include the middle ear reflex and the olivocochlear efferent system.

Finally, Liberman classified auditory nerve fibers into three functional subclasses according to spontaneous discharge rate as low-, medium-, and high-firing units.[118] Whereas high spontaneous rate (SR) fibers exhibited low thresholds and restricted dynamic ranges of about 20 dB, the medium and low SR units displayed higher thresholds and dynamic ranges as great as 60 dB. Thus, it is possible that although fewer in number than the high SR fibers, the low and medium SR fibers are available to encode the higher-intensity sounds.[119]

Viemeister concluded from his review of the psychophysical and physiologic literature relevant to intensity encoding in auditory nerve fibers that a localized rate code "seems theoretically possible and, at present, appears to be the best candidate for a general intensity code."[120] The conclusion that rate information from a frequency-specific population of active nerve fibers provides the neural basis for a code of stimulus intensity is consistent with the findings discussed above concerning the encoding of high-level speech sounds by a population of auditory nerve fibers that possibly exhibit low to intermediate spontaneous firing rates.

Efferent Auditory System *ANATOMY.* Rasmussen established the existence of a chain of descending auditory neurons that links auditory cortex to hair cells and that parallels the classic afferent projection pathway (discussed below).[121] The OCB, the final link in the descending chain, originates in the superior olivary complex (SOC).

Figure 4–26, A, summarizes schematically the origins, course, and distribution of the OCB to one cochlea, whereas Figure 4–26, B, analyzes the principal relationships between the two afferent and the two efferent innervations of the organ of Corti.[95] The OCB is divided into crossed and uncrossed components. The crossed component is composed primarily of relatively large myelinated fibers that originate in large globular neurons surrounding the medial region of the SOC.[122] Most of the uncrossed component is composed of small fibers,[123] a large percentage of which may be unmyelinated and which originate in the small fusiform neurons in the lateral region of the SOC.[122] The majority of the large crossed fibers innervate OHCs, whereas most of the small uncrossed fibers innervate IHCs.[122]

The crossed olivocochlear fibers decussate just beneath the floor of the fourth ventricle and then enter the vestibular nerve. At this point, they are joined by the uncrossed fibers. The combined crossed and uncrossed fibers travel in the vestibular nerve and cross into the cochlear nerve via the vestibulocochlear anastomosis. Within the cochlea, efferent fibers have been identified among the external spiral fibers (3b in Figure 4–25), the tunnel spiral fibers, and the internal spiral fibers.[124]

Physiology. In cases in which olivocochlear axons have been labeled histochemically and traced following single nerve fiber recordings, the labeled fibers terminate beneath OHCs in cochlear regions with best frequencies corresponding to the CFs of the fiber.[125] Such direct recordings from efferent axons within the periphery have demonstrated that these units possess thresholds and tuning curve properties that are essentially identical to those of primary afferents with similar CFs.[125] Although the proportion of olivocochlear neurons that can be excited by ipsilateral and contralateral stimuli is consistent with the predominant efferent innervation to a given cochlea, many units are binaurally responsive.[126]

The effects of stimulating the crossed OCB are almost certainly mediated by the release of ACh as the major neurotransmitter at the OHCs,[127] as noted above. In turn, ACh acts on ACh receptors (AChRs) located on the OHC postsynaptic membrane to modulate the electrical properties of the OHCs. Elgoyhen and associates established that ACh acts on nicotinic AChRs (nAChRs) containing the α9 subunit, which mediates synaptic transmission between the cholinergic olivocochlear fibers and the OHCs.[128] It has long been known that stimulating the crossed OCB in the floor of the fourth ventricle causes depression of the auditory nerve's response to sound stimulation (ie, the CAP) and a simultaneous increase in the amplitude of the CM.[129] It is likely that the α9 nAChR subunit mediates such efferent inhibition because an α9 knockout mouse model showed no suppression of either CAPs or DPOAEs by electrically stimulating the OCB at the floor of the fourth ventricle.[130] It was further shown that the α9 nAChR protein opens its ion channel via associated Ca^{2+}-activated K^+ channels.[131]

Much less is known about the uncrossed OCB. Stimulation of the uncrossed bundle decreases the click-evoked action potential amplitude, but the effect is much less than for the crossed bundle.[99] In addition, unlike the crossed bundle, the terminations of the uncrossed bundle do not show high concentrations of acetylcholinesterase,[132] and curare does not block the effect of stimulating the uncrossed bundle (although strychnine does).[127] Thus, the uncrossed system may not be cholinergic.

Functional Significance The function of the olivocochlear efferent system is, at present, largely unknown. Several attempts to demonstrate functional deficits after sectioning the crossed OCB have failed.[133] A series of more recent studies has systematically examined the effects of electrically stimulating the crossed OCB on auditory nerve fiber dynamic ranges that have been compressed by broadband noise stimulation.[100] The restoration of dynamic range in the presence of background noise supports the notion that the cochlear efferent system may function to improve the ability to discriminate complex signals (ie, the signal-to-noise ratio). However, the hypothesis that the olivocochlear system primarily provides some sort of input gating mechanism may not be completely correct.

The most recent and unexpected effect of stimulating the crossed OCB is a decrease in the levels of OAEs that are generated mechanically by the cochlea and recorded in the ear canal. Mountain was first to report that electrical stimulation of the olivocochlear efferent system directly affected the active nonlinearities in OHC mechanics inferred from decreases in DPOAEs.[82] A number of other studies evaluating the influence of contralateral acoustic stimulation with wideband noise on OAEs recorded ipsilaterally reported small reductions in TEOAEs[101] and in $2f_1$-f_2 DPOAEs.[134] The influence of contralateral stimulation on OAEs also supports the notion that the cochlear efferent system is involved in the modulation of OHC micromechanics. These effects, along with the reductions in neural discharge rate and tuning, suggest that the function of the crossed OCB is to allow the central auditory system to govern the mechanical properties of the basilar membrane by controlling the vulnerable cochlear tuning mechanism previously discussed. Because detuning a frequency bandpass system increases damping, the detuning effect of the crossed OCB could be useful as a method of improving auditory temporal resolution (eg, to improve speech perception). The consensus view at present is that the active transduction process[135] is regulated in some way through efferent innervation of the

OHCs.[89] Also, as noted above, the olivocochlear efferent system may also contribute toward extending the dynamic range of the auditory system.[136]

In addition, other experimental results suggest that the olivocochlear system may help protect against acoustic trauma.[136] In this work, it was discovered that the amount of threshold shift induced by overstimulation of one ear could be reduced by the presentation of a simultaneous tone to the contralateral ear, thus inferring that activating the efferent system reduced the effects of overexposure. One intriguing line of current research on the traumatic effects of noise exposure is the notion of "training" the cochlea to become less susceptible to damaging sounds.[138,139] One common protocol to induce the training (or conditioning or toughening) effect is to apply daily, moderate-level sound exposures to reduce injury from a subsequent high-level noise exposure.[140] Although the underlying mechanism(s) responsible for the trained protective effect is, at present, unknown, the most likely site is the crossed cochlear efferent system.[141,142] Indeed, LePage demonstrated that loud sound induces a mechanical baseline shift in the position of the cochlear partition.[143] It is highly probable that such a gain control system is effected by cochlear efferents.

Central Auditory Pathway

Each of the five primary senses sends information into the brain via two separate pathways: a direct or specific pathway and a nonspecific pathway. The nonspecific pathway involves structures in the core of the neuraxis, collectively known as the reticular system. In the reticular system, all sensory modalities share the same gross neural structures (hence the name nonspecific). Ascent via the nonspecific structures is multisynaptic and hence is characterized by long delay times.

The direct pathways for each sensory modality are separate and involve long axonal processes, with a minimal number of synapses; consequently, compared with the indirect pathway, transmission along the direct pathways involves minimal delay times. The synapses of the direct pathways tend to congregate in well-defined neural structures called nuclei. Clinically, lesions of the central auditory system are localized according to their level in the direct projection pathway. Therefore, the following discussion emphasizes this pathway.

Anatomy Figure 4–30 diagrams the direct auditory projection pathway. The numbers in each nucleus indicate neuronal order (determined by the number of synapses). The auditory projection pathway is more complex than the pathways of other sensory systems, possibly because it developed relatively late on the phylogenetic scale and had to incorporate pieces of other already developed neuronal systems. Although the basic wiring diagram depicted in Figure 4–30 has remained relatively unchanged for many years, it should be emphasized that within the last 25 years, a number of advances have been made in the development of methods for describing neuronal connectivity.[144] These breakthroughs have relied on the discovery that, when certain amino acids, conjugated enzymes (eg, horseradish peroxidase–conjugated lectin from soy bean), sugars, or immunocytochemicals (eg, polyclonal or monoclonal antibodies) are injected into a neuronal or fluid (eg, cochlear duct) region of interest, they will be taken up by the cells or nerve fiber endings in this region and transported by the normal cellular process of axonal transport in both retrograde and anterograde directions. Other techniques make use of such

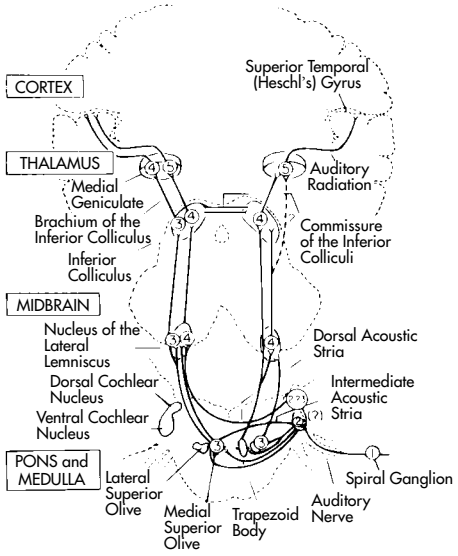


FIGURE 4–30. Diagram of the direct auditory projection pathway. *Numbers* indicate the approximate neuron order as determined by the number of synapses traversed. *Dashed lines* labeled with question marks indicate two areas of uncertainty: (1) whether the dorsal cochlear nucleus primarily contains second- or third-order neurons and (2) whether any nerve fibers bypass the inferior colliculus.

labels as lipophilic dyes (eg, dilinoleyl-tetramethylindocarbocyanine or Dil), which can be retrogradely transported in fixed tissue by diffusion.

Visualization of the location of cellular projections or the cellular/subcellular location of labeled substances at the light and electron microscope levels is achieved either by autoradiography, in the case of radioactive compounds, or by catalyzation of histochemical reactions, which are typically viewed under epifluorescence or darkfield optics. With recent refinement of these tract-tracing and cell component–labeling techniques using immunohistochemical markers, virtually all projection pathways and cell types of the auditory system have been described. Lagging far behind our description of the connections of the neural pathway for hearing is our understanding of how this network of interconnections interacts to produce our complex auditory sensations. Presented below is a brief description of the major projections of the auditory pathway sufficient for the clinician to understand the transfer of information within the auditory system, without consideration of the many lesser connections revealed by modern immunocytochemical staining techniques. The reader is referred to several comprehensive reviews that provide more details of the intricate interconnections of the central auditory system.[96,145]

Cochlear Nucleus. Central processing of the information carried in the auditory nerve begins in the cochlear nucleus (CN), the first obligatory synapse for all nerve fibers. On entering the CN, the auditory portion of the eighth nerve bifurcates into two branches, one that sends fibers to synapse in the anteroventral cochlear nucleus (AVCN) and one that synapses in the posteroventral cochlear nucleus (PVCN) and DCN. The distribution of fibers within the CN is not random but follows an orderly pattern of tonotopic projection throughout the nucleus, with low-frequency fibers projecting ventrally and high-frequency fibers distributed dorsally. Thus, both neurophysiologic and anatomic observations show that cochlear place is represented in an orderly manner throughout the projection pathways of the central auditory system.[146] The security of tonotopic organization is also apparent in the multiple representation of the stimulus frequency domain. For example, typically, along any one penetration of a microelectrode trajectory within a principal nucleus, there are two or more breaks in the orderly progression of best frequencies. Thus, as in the direct projection pathways of all sensory systems, multiple representation of the receptor surface occurs. It is likely that nerve fiber branching in the various nuclei, as described for the CN, is the anatomic basis of this multiple-frequency representation.[147]

All auditory nerve fibers display relatively uniform response characteristics to pure-tone stimuli compared to the activity of CN cells when categorized by the use of a poststimulus time histogram (PSTH).[113] The PSTH plots the number of nerve discharges that occur in small time bins within the period that begins slightly before and extends throughout the duration of the stimulus. Poststimulus time histograms are shown for an auditory nerve fiber and various CN cells in Figure 4–31. The CN consists of at least nine different cell types described for their anatomic characteristics revealed by various staining techniques.[148] It can be seen from the unit activity patterns of Figure 4–31 that the uniform response characteristics of auditory nerve fibers are soon elaborated on by the various cells of the CN to produce a variety of response types named after the patterns seen in the PSTH.[149] When examined in the frequency intensity domain, some CN cells exhibit complex response patterns that describe regions of excitation and inhibition produced by a pure-tone stimulus. An example of such complex response patterns by cells in the CN is shown in Figure 4–32.[149]

Three fiber pathways project information from the CN to higher brainstem centers. These fiber tracts include the ventral (trapezoid body), intermediate, and dorsal acoustic striae. The ventral acoustic stria, originating from AVCN and PVCN regions, projects ventrally and medially to send fibers ipsilaterally to the lateral superior olive (LSO) and the medial superior olive (MSO) and then to the medial nucleus of the trapezoid body (MTB). The fibers then cross the midline to terminate on the contralateral MSO and MTB but do not innervate the LSO on the opposite side. Thus, for any one ear, connections are bilateral only to the MSO and MTB.[150]

The intermediate acoustic stria, the primary output of the PVCN, sends some fibers ipsilaterally to a group of cells around the SOC called the periolivary nucleus (PON), whereas other fibers ascend in the lateral lemniscus (LL). This pathway also crosses the midline to innervate the same structures on the contralateral side. Fibers of the intermediate acoustic stria that enter the LL synapse within the nucleus.

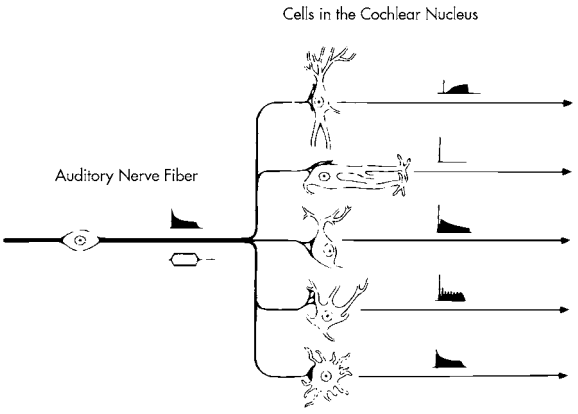


FIGURE 4–31. Diagram showing the diverse single-unit response types obtained from cells in the cochlear nucleus compared to the uniform response pattern found for fibers of the auditory nerve. The poststimulus time histograms were obtained by presenting short 25 to 50 ms tone bursts at the unit's characteristic frequency. Response types from top to bottom are pauser, on_1, primary like with notch, chopper, and primary like. Cell types presumably associated with these response patterns from top to bottom are pyramidal, octopus, globular, multipolar, and spherical. Reproduced with permission from Kiang NYS. Stimulus representation in the discharge patterns of auditory neurons. In: Tower DB, editor. The nervous system. Vol 3. Human communication and its disorders. New York: Raven Press; 1975.

The dorsal acoustic stria, the principal output of the DCN, bypasses the SOC to synapse in the contralateral dorsal nucleus of the LL and the inferior colliculus (IC).[150]

SUPERIOR OLIVARY COMPLEX. The SOC is composed of the LSO, MSO, MTB, and PON. The SOC receives fibers from both CNs and consequently receives information from both ears. This feature allows this group of nuclei to monitor the arrival time and level of sounds to both ears and provides the cues for the localization of sound in space based on the stimulus arrival time and intensity to both ears.[150] In fact, localization of sound in the horizontal plane provides the most straightforward relationships that have been observed between central auditory single-unit behavior and psychophysical function. As sound to one ear is made progressively louder or earlier than sound to the opposite ear, some SOC units change abruptly from inhibitory to facilitatory response patterns. By assuming a facilitatory contralateral input and an inhibitory ipsilateral input (or vice versa), the neurophysiologic behavior of these units can be explained. Thus, slight binaural differences in level and/or arrival time provide the auditory cues for localization of sound in the horizontal plane (sound lateralization). The acoustic image is located on the side of the louder or earlier sound. As a whole, the SOC represents the lowest

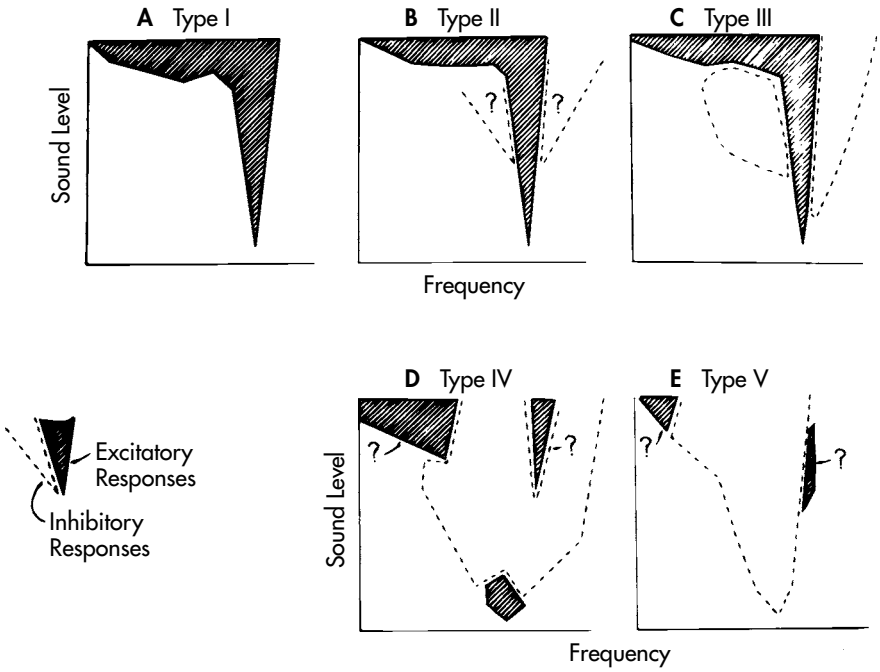


FIGURE 4–32. Tuning curves of excitation and inhibition in the cat cochlear nucleus are shown in order of increasing amounts of inhibition (*A–E*). Purely excitatory responses as in *A* are predominant in the anteroventral cochlear nucleus. Greater amounts of inhibition are found toward the dorsal cochlear nucleus (*D* and *E*). *Question marks* show variable or uncertain features. Reproduced with permission from Young ED.[150]

level in the auditory system at which binaural processing takes place.151 However, at all levels in the auditory projection pathway above the trapezoid body, there are units that are sensitive to binaural time and level differences.152,153

Lateral Lemniscus. The LL is the major ascending projection from the CN and SOC to the IC and contains both contralateral and ipsilateral fibers from lower auditory brainstem structures. Although there are three distinct nuclei within the LL, historical emphasis has been simply on its function as a connection between the SOC and IC. Recently, these nuclei have received more attention in attempts to define their role in auditory processing.[154]

Inferior Colliculus. The IC receives synapses from the majority, if not all, of the fibers projecting from the lower auditory nuclei. The three neuronal areas that make up the IC are the central, external, and pericentral nuclei. The predominant termination zone for the ascending auditory projections is in the ventrolateral region of the central nucleus. This region receives inputs from the SOC and a heavy contralateral input from the DCN. Other major projections come ipsilaterally from the MSO and bilaterally from the LSO. The function of the IC is far from completely understood, with many neurons exhibiting complex excitatory/inhibitory interactions. A simple view is that the IC integrates the frequency analysis features of the DCN with the localization abilities of the SOC.[150]

Medial Geniculate. The medial geniculate (MG) is the thalamic relay nucleus for auditory information. All auditory projections from the IC to the auditory cortex pass through the MG. This nucleus is also composed of three divisions, the ventral, dorsal, and medial nuclei.[155] The ventral nucleus receives heavy ipsilateral projections from the ventrolateral portion of the central nucleus of the IC. This portion of the MG projects to the AI, AII, and Ep regions of the auditory cortex (see below). Again, a wide variety of single-unit response types have been recorded in the MG. In attempting to understand the MG's role in auditory processing, special efforts have been made to examine MG single-unit responses to complex sounds. Thus, David and associates[156] and Keidel[157] demonstrated that MG neurons in the unanesthetized cat were responsive only to specific parameters of complex speech sounds. In general, it has been extremely difficult to attribute specific feature extraction capabilities to neurons in the higher auditory nuclei that cannot be explained by complex responses already present at the level of the CN.

Auditory Cortex. The auditory cortex has been most extensively studied in the cat and can be divided into three areas based on similarity of Nissl-stained cytoarchitectural details.[158] These include a primary area AI, a secondary area AII, and a remote projection region Ep. In human and nonhuman primates, the primary auditory projection area is located in the temporal lobe but hidden by the sylvian fissure. The ventral division of the MG projects almost entirely to AI,[159] which can be considered the primary auditory cortex. Surrounding auditory areas receive projections from all divisions of the MG. Like the auditory relay nuclei, the auditory cortex is also tonotopically organized. As might be expected by the many intricate interconnections of the auditory system prior to input arriving at the cortex, the understanding of cortical processing has been a complex and difficult task. One approach to solving the problem of cortical function has been the use of ablation studies in which the auditory cortex is removed after training an animal to perform a specific auditory task. These studies have demonstrated that cortical ablation does not result in a complete loss of function as do similar lesions in the visual system. In fact, for many simple tasks, no long-term deficits can be detected.[157] Based on these results, it is reasonable to assume that the auditory cortex is involved in numerous details of more complex auditory processing. Consequently, it is unlikely that one simple unifying concept will be uncovered that describes the functional role of the auditory cortex.

Summary Although there have been many studies of single-unit activity in the auditory nuclei over the past 50 years, they have thus far yielded few unifying principles about the data-processing mechanisms of the auditory nervous system. However, a number of recent studies suggest that, in many instances, the brain makes use of patterns of activity distributed over many cells to extract relevant information.[147] Consequently, if future studies focus on describing the responses of individual cells without viewing their participation, as a whole, in some functional

unit, they may be doomed to result in failure as a means of understanding central auditory processing.

Auditory-Evoked Electrophysiologic Responses

Modern averaging techniques have made it possible to record from humans, with surface electrodes, electrical responses reflecting the entire auditory pathway, from cochlea to cortex.[160,161] Clinical applications of this technique have expanded rapidly and are routine in most otolaryngology clinics today. In this section, some basic physiologic principles are discussed on which the clinical applications of auditory evoked potentials are based.

Two major classes of human auditory evoked potentials generated by acoustic transients (clicks) are used clinically. One class is recorded with an electrode located as close to the cochlea as possible, that is, either extratympanic (eg, located on the external ear canal skin) or transtympanic (eg, penetrating the eardrum to rest on the medial wall of the middle ear).[162] Tests based on this class of auditory evoked potentials have been termed electrocochleography (ECochG). The second class of auditory evoked potentials is recorded between one electrode located on the vertex and another near the external ear (eg, either on the mastoid or earlobe). This latter class of evoked potentials has been subdivided conventionally according to their onset latency range into early, middle, and late responses.

Electrocochleographic Responses: Whole-Nerve Compound Action Potential, Cochlear Microphonic, Summating Potential *Description of the Responses.* Figure 4–33 shows a typical example of cochlear and auditory nerve click-evoked potentials recorded from the ear canal. Evoked responses to clicks of opposite polarity are shown in Figure 4–33, A. The whole-nerve CAP has two or more ear canal negative peaks designated N_1, N_2, N_3 (N_2 and N_3 not labeled here), and so forth. Each peak lasts about 1 ms and, in normal ears, N_1 is always the larger peak. In Figure 4–33, B, the CM appears as a series of sinusoidal oscillations, typically about 3 kHz, on the leading edge of the N_1 peak. The SP in

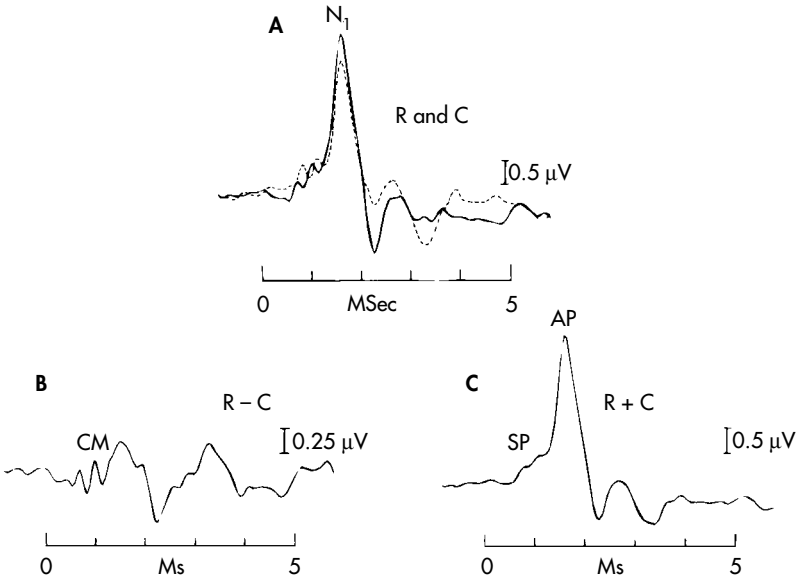


FIGURE 4–33. Examples of cochlear and auditory nerve electrical responses to clicks recorded from the human outer ear canal. The cochlear potentials are the cochlear microphonic (CM) and summating potential (SP). *A*, The whole-nerve compound action potential typically has a large negative peak (N_1). R and C are shown as separate condensation (*dashed line*) and rarefaction (*solid line*) rectangular-pulse click responses recorded from the outer ear canal with a nasion "reference" electrode. *B*, R – C is the waveform produced by subtracting C from R responses. *C*, R + C is the waveform produced by adding C and R responses. Upward deflections represent negativity at the ear canal, and time scale zeroes are set at the leading edge of the rectangular pulse driving the earspeaker. Click rate was 8/s; click level was 115 dB peak equivalent SPL. Each C and R response was the average of 1,000 single sweeps.

Figure 4–33, C, appears as an ear canal negative hump on the leading edge of the N_1 peak, with the CM oscillations superimposed.

The auditory nerve CAP (see Figure 4–33, A) is the sum of the synchronous firing of single-fiber auditory nerve fibers as seen by the distant recording electrode as illustrated in Figure 4–34.[163] The single-unit spikes are triggered by the passage of the cochlear traveling wave. Thus, spikes from basal (high-frequency) fibers appear at the recording electrode earlier than spikes from apical (low-frequency) fibers. Because the degree to which single neural unit responses contribute to the evoked potential is determined largely by their synchrony, high-frequency (basal) fibers contribute more effectively to the whole-nerve CAP and also to the evoked auditory brainstem response (ABR) than do low-frequency fibers.

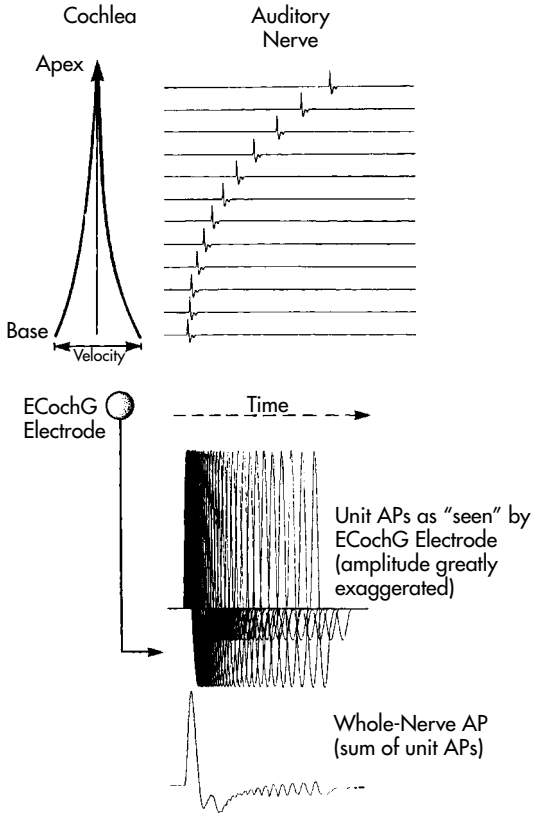


FIGURE 4–34. Summation of the single-unit spikes (amplitude exaggerated) picked up by a distant electrocochleographic (ECochG) electrode to form the whole-nerve compound action potential (CAP). Note that the highly synchronized spikes from the basal end of the cochlea sum more effectively than the poorly synchronized spikes from the apical end of the cochlea. The records of unit action potentials (APs) and the whole-nerve CAP at the bottom are from a simplified computer experiment. Reproduced with permission from Elberling C.[163]

This principle leads to several important conclusions for clinical applications: (1) broadband-click evoked CAPs preferentially reflect high-frequency fiber activity, and, as a corollary, high-frequency clicks are more effective generators of auditory evoked potentials than low-frequency clicks[164]; (2) high-frequency hearing loss caused by selective removal of basal units tends to bias the broadband click-evoked CAP and ABR toward more apical units. Therefore, high-frequency cochlear deficits prolong N_1 peak latencies[164]; and (3) a low-frequency deficit does not cause a corresponding basalward shift because of the already existing heavy bias toward basal units in the normal response. Therefore, low-frequency cochlear deficits have no apparent effect on CAP and ABR latencies. Because the CM response follows the waveform of basilar membrane vibration, it reverses polarity when click polarity is reversed, whereas the envelope-following SP does not. The neural response also maintains the same polarity. Thus, adding together responses to opposite-polarity clicks cancels out the polarity-reversing CM and allows a better view of the SP (see Figure 4–33, C). In contrast, subtracting the condensation and rarefaction responses preserves the CM, while canceling the SP and most of the CAP (see Figure 4–33, B).

The cochlear receptor potentials (CM, SP) can also be separated from the CAP because receptor potentials are nonrefractory and do not adapt, whereas neural responses do. Thus, masking and increasing click rate are often used in clinical testing to depress selectively the neural response while preserving the CM and SP.

Reliable methods for separation of the SP from the other cochlear potentials have made this response useful in the diagnosis of Meniere's disease.[165] A number of studies have demonstrated that the SP is enlarged in a certain percentage of patients with Meniere's disease.[166] Evidence supporting the notion that SP enlargement is specifically related to endolymphatic hydrops is provided by observations that manipulations expected to reduce fluid accumulation within the endolymphatic space (eg, administration of hyperosmotic agents such as glycerol) often shrink abnormally enlarged SPs.[167]

Vertex-Recorded Auditory Evoked Potentials

Description of the Responses Figure 4–35 shows examples of early (A), middle (B), and late (C) vertex-recorded auditory evoked potentials generated by broadband clicks. Commonly used peak designations are also shown. In all traces, upward deflection represents negative voltage at the vertex ("referred" to an electrode on the mastoid). Note the differences in the time and voltage scales. The early response, commonly termed the ABR, is the smallest and also is the most recently discovered.

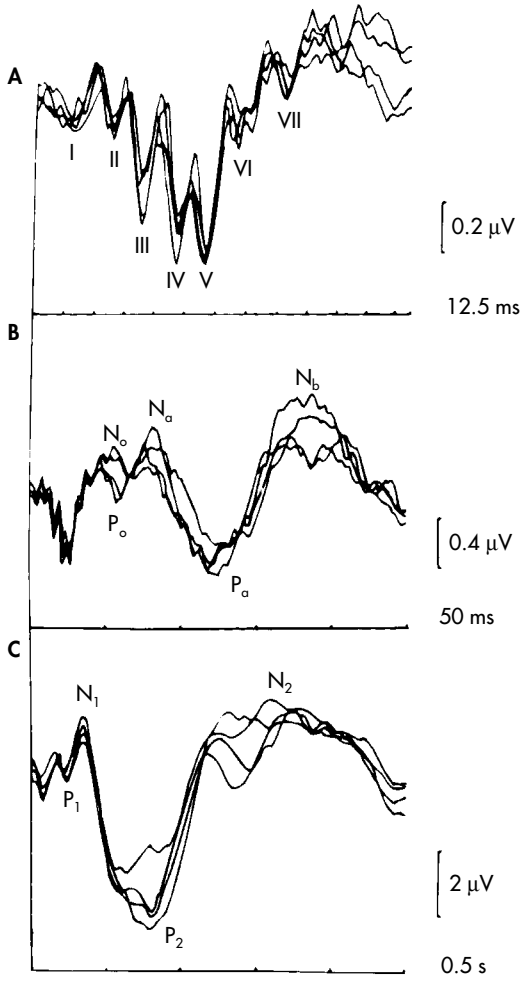


FIGURE 4–35. Averaged human click-evoked potentials, recorded with vertex-mastoid electrodes. Upward deflection represents negativity at the vertex. From top to bottom, the time base is slowed to demonstrate progressively later responses. *A*, Auditory brainstem response (with the polarity opposite to the response recorded with a nasopharyngeal electrode). *B*, Middle latency response. *C*, Late "vertex" auditory evoked response. Each tracing represents the average of 1,024 single responses. Several tracings are superimposed in each record to give an idea of variability. Roman numerals and letters identify individual peaks of the various responses according to accepted convention. Reproduced with permission from Picton TW et al.[161]

Auditory Brainstem Response. An example of an ABR is shown in Figure 4–35, A, in which the peaks of this response are labeled with Roman numerals I to VII. It can be seen that the ABR occurs approximately between 1 and 8 ms after stimulus onset. Wave V is typically the largest and most robust of the potentials, and waves beyond V are seldom used clinically. Studies in humans[168] and animals[169] suggest that wave V is generated at the LL. Thus, the ABR is probably not useful in detecting abnormalities at or above the level of the IC.[170] However, for the lower brainstem, the ABR is one of the best audiologic tests for detecting dysfunction.

The ABR is one of the most frequently used auditory evoked potential procedures because the variables that affect this response have been well described and it provides a good measure of cochlear sensitivity and retrocochlear status. Clinically, the response can be interpreted by quantitative measures of peak latencies, interpeak intervals, and interpeak latency differences. In addition, the presence or absence of the various waves is noted as well as waveform morphology.[171]

Middle Latency Response. The middle latency response (MLR), labeled N_0, P_0, N_a, P_a, and so forth (Figure 4–35, B), occurs between 8 and 40 ms after the auditory stimulus.[172] In animals, wave P_a appears to be generated in the primary auditory cortex by cortical elements on the side contralateral to stimulation.[173] However, in humans, P_a seems to be generated in both hemispheres, even with monaural stimulation. Human cortical mapping localizes this potential to the region of the sylvian fissure,[174] whereas studies based on clinical correlations indicate involvement of thalamocortical projections to the primary auditory area located along Heschl's gyrus in the genesis of this response. Site-of-lesion studies suggest that wave N_a originates in the midbrain, including such structures as the MG and thalamocortical projections.[174]

Although the MLR is not used nearly as frequently as ABR, useful clinical information can be obtained from this response, particularly in conjunction with ABR testing. The MLR is robust at all

frequencies including low frequencies below 1,000 Hz and, consequently, offers a complementary index of high- and low-frequency hearing.[175] The MLR can also be useful in assessment of central auditory system disorders, neurologic evaluation, and, most recently, as a tool for cochlear implant assessment.[176] The most significant disadvantage of MLR is that subject variables such as sleep state or sedation can severely reduce the amplitude of this potential. Thus, behavioral state must be controlled for to obtain meaningful results.

Late Auditory Evoked Response. The relatively high-amplitude late auditory evoked response (LAER), labeled P_1, N_1, and so forth (also called the vertex potential), occurs between 50 and 250 ms after the stimulus (Figure 4–35, C). It is a cortical response, but its long latency suggests that the LAER is not generated within the primary auditory cortex but rather that it originates in the associative cortex.[161] These potentials are often very dependent on the subject's state of alertness and consequently have received little use in clinical situations.

Auditory Brain Mapping One technique that is receiving considerable attention as a new and potentially powerful diagnostic tool is the use of auditory brain mapping. Methods for this procedure are highly similar to those used to obtain the traditional evoked responses already described. The primary difference is that brain mapping simultaneously records the electrical activity from an array of scalp electrodes. Although this mass of data could be viewed as many evoked potentials, the human observer cannot readily interpret such a massive amount of data. Therefore, brain mapping was developed using computer-processing techniques to provide a visual display of the potential fields among all of the electrodes simultaneously. This method reduces the data into a multicolored or shaded plot in which intense activity is usually given the most brilliant color or the darkest shading. An example of the typical electrode placements and the derivation of a simple brain map are shown in Figure 4–36. With this type of visual output, patterns of activity in normal patients can be established and abnormalities easily visualized in pathologic cases. Computer techniques also allow the clinician to view the patterns of activity throughout the duration of the stimulus to produce a motion picture of the evoked electrical activity that occurred over time. This method allows for the visualization of the spatiotemporal patterns of brain activity. Because the data are stored in the computer, a number of advanced mathematical processing methods, which permit further refinement of the data, can be applied

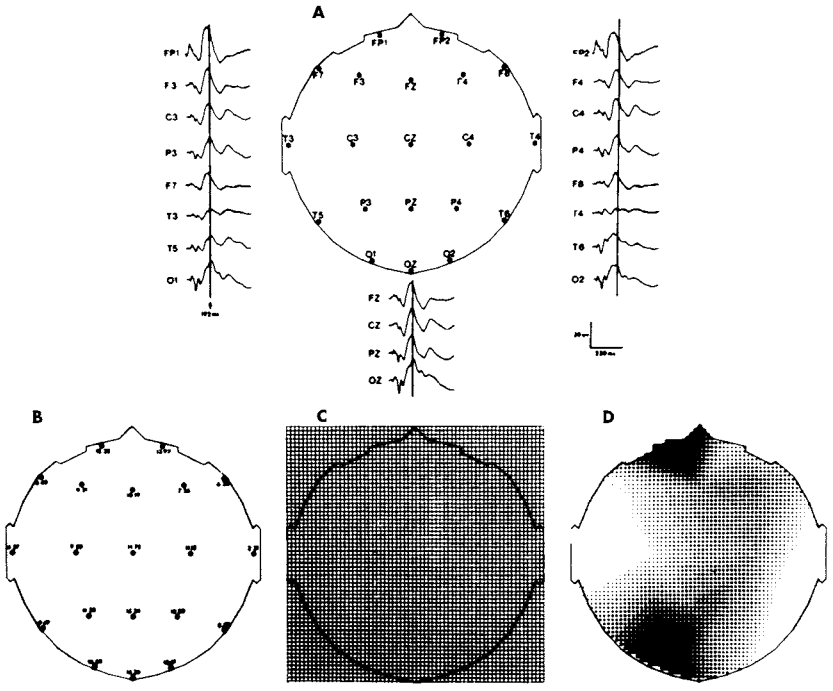


FIGURE 4–36. Example of the construction of topographic images from evoked potential data. *A*, Individual evoked potentials are shown at the electrode locations indicated in the diagram. *B*, Mean voltages are shown for the time interval 192 ms after the stimulus. *C*, Head region is divided into a 64×64 matrix to produce 4,096 spatial domains. Each spatial domain is assigned a voltage derived by linear interpolation from the three nearest recording locations. *D*, A visual image constructed by fitting a discrete-level, equal-interval intensity scale to the points of *C*. Although a visual evoked potential was used to create this topographic map, the same procedure is used for mapping auditory data. Reprinted with permission from Duffy DF.[177]

to aid in the detection of subtle abnormalities. For more information on this topic, the reader is referred to these comprehensive reviews.[177,178]

Magnetoencephalography Magnetoencephalography (MEG) involves the completely noninvasive recording of weak cerebral magnetic fields, which represent about one part in 109 of the earth's geomagnetic field, outside the head.[179] The neuromagnetic technique was made possible by the invention of SQUID (superconducting quantum interference device) magnetometers. Using a whole-head neuromagnetometer, magnetic brain signals are averaged by time-locking them to the onset of acoustic stimulation. It is assumed that the probable sources of cerebral magnetic fields are the electric currents in the synapses of synchronously activated cortical pyramidal neurons, and the sinks or volume currents, to complete the electrical circuit, are generated in the surrounding tissue. The development of multichannel systems increased the speed and convenience of neuromagnetic recording and made it feasible to apply MEG for clinical purposes.

Based on the accurate spatiotemporal resolution of whole-head MEG, Hari and Lounasmaa have been instrumental in applying MEG to investigate the activity of the auditory projection areas, particularly with respect to language-related sites.[179] A schematic illustration of a typical experimental setup for auditory measurements is illustrated in Figure 4–37 and shows the relationship of the SQUID to the patient's head. The inset above indicates the isocontours across the scalp for the component of the magnetic field, which is generated by the active region of the auditory cortex. An example of the averaged magnetic responses from measurement sites anterior (above) and posterior (below) to the sylvian fissure of

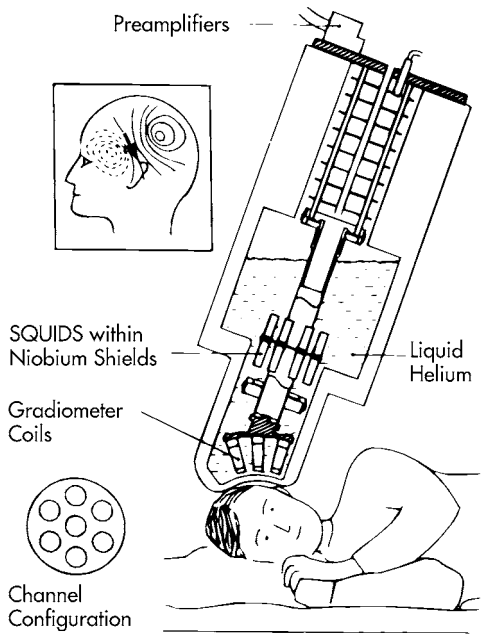


FIGURE 4–37. Schematic illustration of a typical experimental setup for auditory measurements using magnetoencephalography. The most important parts of the seven-channel DC SQUID instrument are shown, with the midpoints of the different channels separated by 36.5 mm. The *inset* above depicts isocontours across the scalp for the radial component of the magnetic field, generated by an active area in the auditory cortex. The *arrow* illustrates the location of the equivalent current dipole. Reproduced with permission from Hari R and Lounasmaa OV.[179]

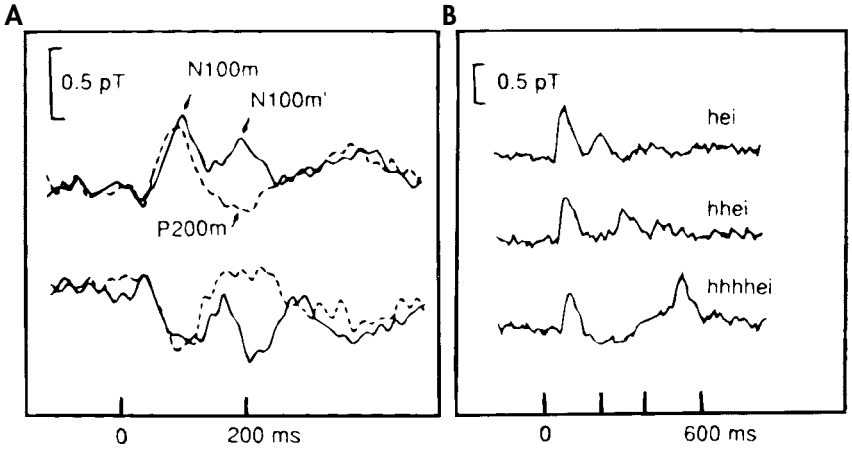


FIGURE 4–38. Averaged magnetic responses from the auditory cortex. *A*, Averaged magnetic response (n = 120) to "hei" words (*solid lines*) and noise bursts (*dashed lines*) from the right hemisphere in one subject; the upper curves are from an anterior and the lower ones from a posterior measurement location near the ends of the sylvian fissure. The passband was 0.05 to 70 Hz. *B*, Effect of increasing the duration of "h" in another subject. Reproduced with permission from Hari R et al.[180]

the right hemisphere (ie, along the superior surface of the temporal lobe) are shown for one subject for a spoken word (solid line) and noise bursts (dashed lines) on the left side of Figure 4–38. The right side of Figure 4–38 illustrates the effects of increasing the duration of the onset of the word in another subject. It is clear that sound evokes a complex magnetic waveform, which lasts several hundred milliseconds after stimulus onset. Interesting and clinically useful data have been obtained about the perception of speech of deaf patients with cochlear prostheses by examining the activity of their auditory cortices with MEG.[180] Information of this real-time sort implies that neuromagnetic recording may be used to assess functional disorders in more detail than what is possible by other clinical evidence; thus, it is likely to become an important diagnostic tool of the future.

Neuroimaging The beginning of in vivo imaging in awake and behaving subjects has revolutionized the ability to relate structure to function, particularly in the human brain. For example, three-dimensional magnetic resonance imaging (MRI) provides detailed static images of the in vivo brain, with spatial resolution on the order of cubic millimeters. Other functionally related techniques in neuroimaging such as positron emission tomography and functional MRI provide excellent spatial and temporal resolutions for recording changes in regional blood flow and cellular metabolism during auditory stimulation. These techniques have been used to study a number of central auditory processes, including speech perception,[181,182] revealing tinnitus-related abnormalities in brain function.[183,184] Although these individual functional imaging approaches still have some limitations, the results are converging to improve our understanding of the neural processes underlying both normal and pathologic auditory function.

Evoked Otoacoustic Emissions One of the principal benefits of OAEs with respect to clinical testing is that they provide an objective and noninvasive measure of cochlear activity, which is completely independent of retrocochlear activity. In particular, OAEs measure the functional responses of OHCs that are uniquely sensitive to agents that damage hearing. Kemp,[80] the pioneer investigator into the basic features of OAEs, has also been instrumental in establishing one class of evoked emissions, the click-based TEOAE (see Figure 4–22, C), as a potentially useful diagnostic indicator of various ear diseases including Meniere's disease[104] and acoustic neurinoma.[185] In addition, Kemp and Ryan have also established the use of the click-evoked OAE as a method of screening for hearing dysfunction in newborns.[186]

Because of the frequency specificity of the eliciting pure tones, DPOAEs (see Figure 4–23) also have a beneficial clinical applicability. This stimulus-related frequency specificity permits DPOAEs to be measured after averaging the responses to only a few stimuli. Thus, under computer control, the fine resolution of DPOAEs in both the stimulus frequency and level domains permits the precise determination of the boundary between normal and abnormal OHC function. Data from our laboratory illustrate this feature of DPOAEs in the plots of Figure 4–39, A,[187] which show the development of ototoxicity during a course of antineoplasm therapy with cisplatin. Comparing the OAE findings with behavioral hearing (top), note the abrupt change in hearing and emission activity between 2 and 3 kHz (open circles), which accurately followed the 40 dB loss in hearing sensitivity produced by the ototoxic agent.

Based on their ability to distinguish between the relative contribution that sensory and neural components of the cochlea make to a hearing loss, OAEs promise to facilitate identification of the anatomic substrate of complex ear diseases. For example, the records shown in Figure 4–39, B, represent a typical pattern of audiometric loss in early Meniere's disease, which is accompanied by thresholds that are > 40 dB HL and, not unexpectedly, no TEOAEs (bottom left) and abnormally low-level DPOAEs (solid circles). Together, these results imply that the OHC system is involved in this patient's disease. However, other patients with Meniere's disease, who show a similar hearing loss, can demonstrate rather robust DPOAEs, even for test frequencies for which behavioral thresholds are > 50 dB HL.[188] Such differential findings for patients with Meniere's disease suggest that OAEs can distinguish between disease states that involve either OHCs (see Figure 4–39, B) or other cochlear processes that are probably neural in origin (eg, degeneration of the unmyelinated dendritic endings of the cochlear nerve). A comparable example of this feature of OAE testing to contribute toward differentiating between cochlear versus noncochlear involvement in hearing disease is illustrated in Figure 4–40 for two patients diagnosed with acoustic neurinomas. The patient in Figure

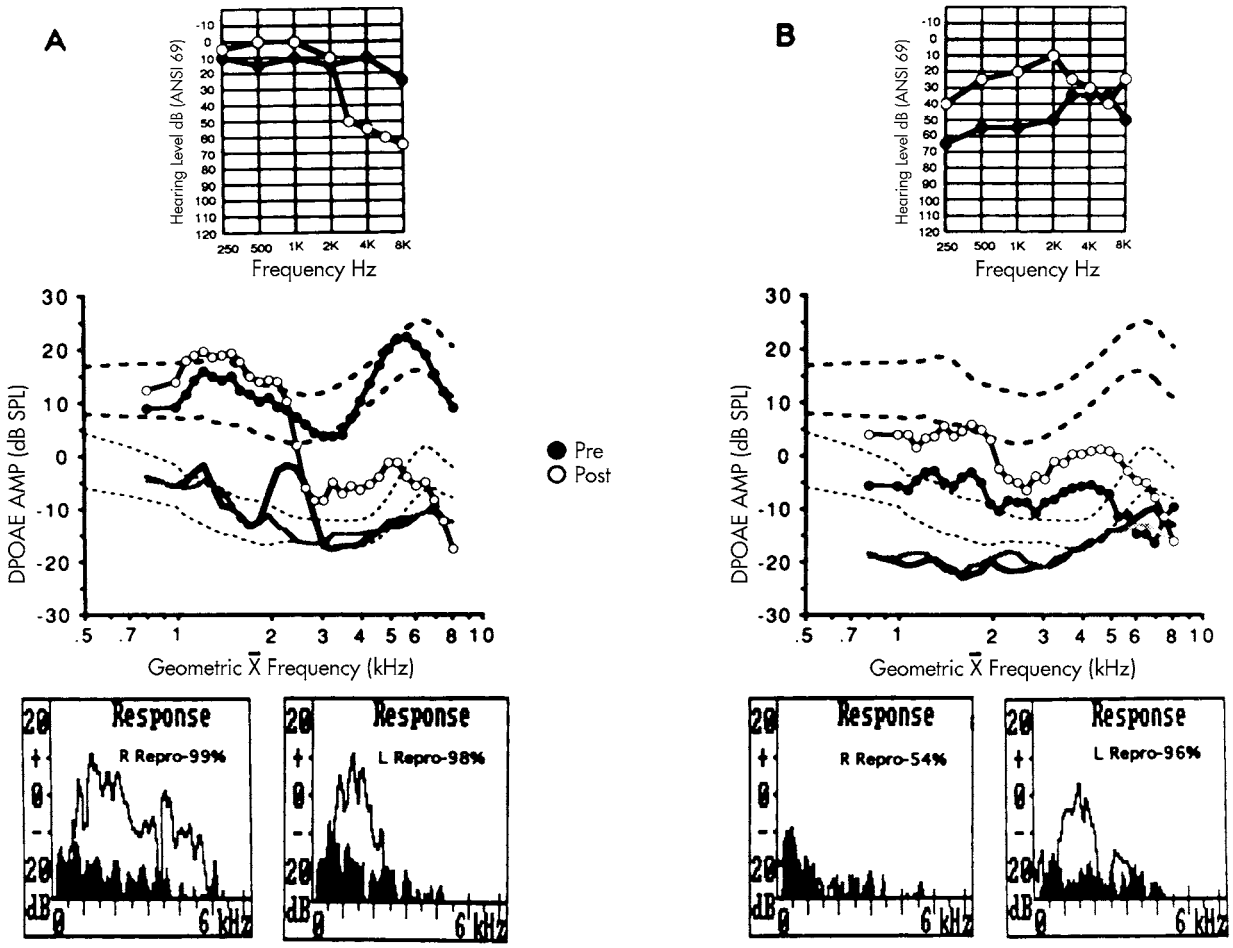


FIGURE 4–39. Conventional audiograms (*top*), distortion-product otoacoustic emission (DPOAE) frequency/level or DP-gram functions (*middle*) relating emission magnitude to stimulus frequency, in 0.1 octave steps from 0.8 to 8 kHz, for primary tone levels of 75 dB SPL, and click-evoked transient evoked otoacoustic emissions (OAEs) (*bottom*) elicited by the "default" mode of the ILO88 Otodynamic Analyser. *A*, Behavioral hearing and evoked OAE findings for the right ear of an 8-year-old girl comparing pre- (*solid circles*) versus post-treatment (*open circles*) responses following a course of antineoplasm therapy with cisplatin. Immittance and speech testing results were normal for the baseline session and were not tested during the post-treatment examination. *B*, Audiometric and OAE results for the left ear of a 49-year-old woman for two test sessions separated by about 2 weeks. The patient complained of a fluctuating hearing loss, tinnitus, aural fullness, and episodic dizziness of several months duration. Immittance findings were normal, whereas the speech discrimination score and speech reception threshold improved from 72 to 96% and 55 to 25 dB HL, respectively, at the last evaluation period (*open circles*). Reproduced with permission from Balkany TJ et al.[187]

4–40, A, displays the expected outcome for OAE testing in the presence of a retrocochlear disease. That is, in the presence of a moderate hearing loss on the tumor side (solid circles), both TEOAEs and DPOAEs are measurable. However, it is clear for the patient in Figure 4–40, B, diagnosed with a right-sided (open circles) acoustic neurinoma, that the tumor has modified cochlear function in a manner that mimics the frequency configuration of the hearing loss. Thus, in certain retrocochlear disorders such as tumors of the cerebellopontine angle, cochlear function can be adversely affected, probably owing to compromise of the vascular supply to the inner ear.

One benefit of OAE testing illustrated in Figure 4–39, B, is its sensitivity to changes in hearing status, which are mediated by the OHC system. In this instance, at the time the patient returned for follow-up testing (open circles), her hearing had improved, particularly over the low-frequency test range. The 20 to 30 dB improvement in low-fre-

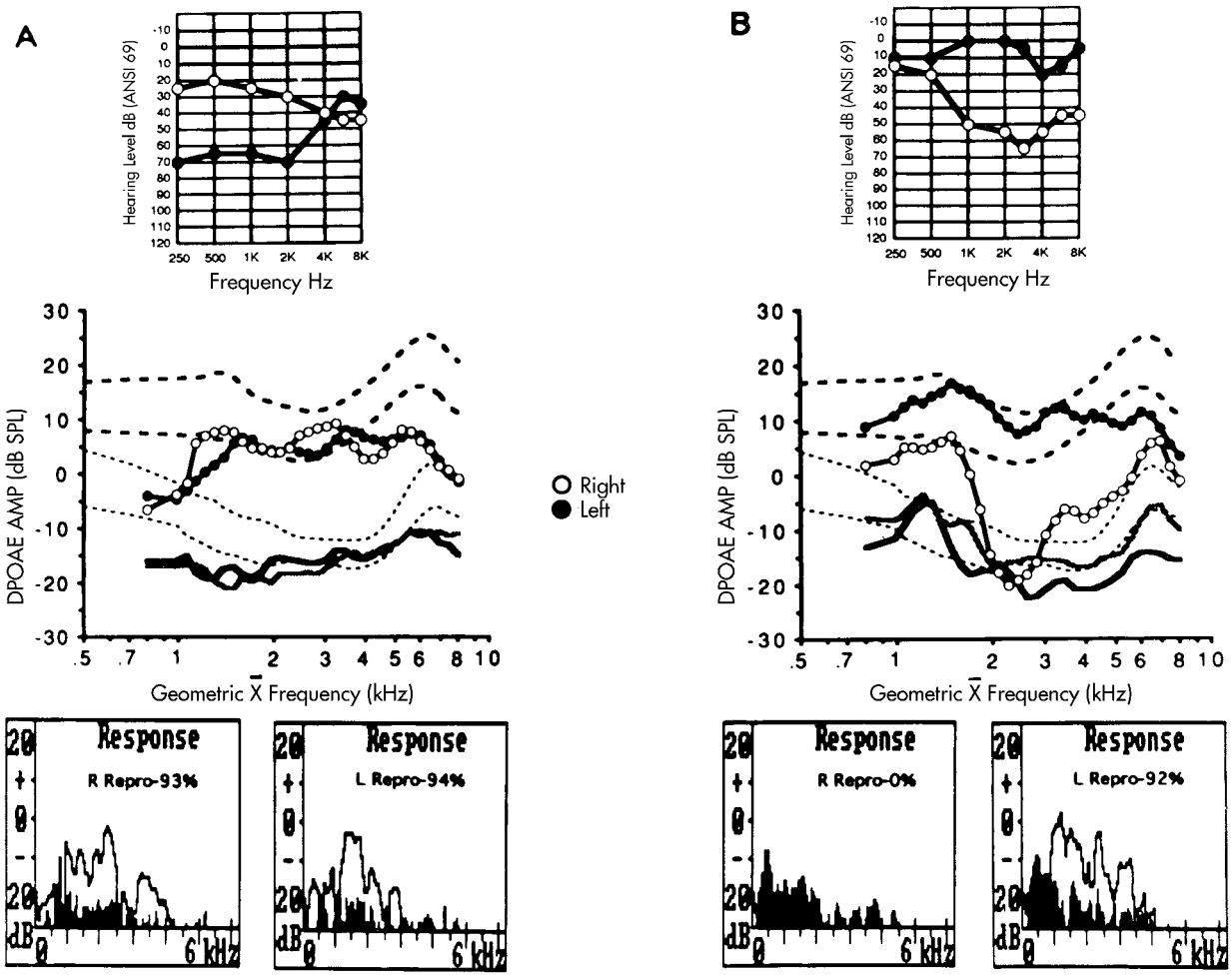


FIGURE 4–40. Audiometric and otoacoustic emission findings for two patients with acoustic neurinomas confirmed by magnetic resonance imaging. *A*, A 39-year-old man with an acoustic neurinoma on the left side who had normal tympanograms but absent reflexes bilaterally. Although speech testing was normal for the right ear (*open circles*), the speech discrimination score and speech reception threshold were 0% and 70 dB HL, respectively, for the involved left ear (*solid circles*). *B*, A 47-year-old woman with a right-sided acoustic neurinoma. Whereas her tympanograms were normal, the only measurable acoustic reflex threshold was for ipsilateral stimuli at 1 kHz. Speech testing was normal for the left ear, whereas the right ear was associated with a speech discrimination score of 60% and a speech reception threshold of 35 dB HL. Reproduced with permission from Balkany TJ et al.[187]

quency hearing was mirrored by a similar increase in DPOAEs and by the ability to now measure TEOAEs.

In addition to the clinical applications noted above for OAEs, a number of studies have also demonstrated the utility of using emitted responses to understand the fundamental basis of a number of cochlear lesions, including the effects of noise-induced hearing loss,[189] hereditary hearing impairment,[190] congenital hearing disorders,[188] ototoxicity,[191] bacterial meningitis,[192] and presbycusis[193] on both TEOAEs and DPOAEs. The interested reader is encouraged to refer to these comprehensive reviews concerning the strengths and limitations of the clinical application of OAEs.[194,195]

VESTIBULAR SYSTEM

General Principles

The nonauditory part of the inner ear (termed the vestibular apparatus) consists of two functional sub-

divisions: (1) the semicircular ducts consisting of two vertical and one horizontal and (2) the otolithic organs consisting of the saccule and the utricle. The semicircular ducts respond to head rotation (ie, angular acceleration), whereas the otolithic organs are stimulated by the effects of gravity and linear acceleration of the head. The primary function of the utricle is to signal head position relative to gravity. Ablation of the saccule produces a less significant deficit than ablation of the utricle; hence, the function of the saccule is less well defined than that of the utricle. In fact, it has been proposed that the saccule is a low-frequency auditory receptor.[196] However, a series of systematic studies using single-unit recordings have revealed that saccular nerve fibers respond only to linear acceleration.[197–199] These findings suggest that the saccular system provides the high-level vertical acceleration signals required to elicit the motor response necessary to land optimally from a fall.[200]

Conventionally, the vestibular system is regarded as one of three sensory systems that function to maintain body balance and equilibrium. The other two are the somatosensory (chiefly proprioceptive) and visual systems. Loss of proprioception (eg, as in tabes dorsalis) or vision causes more significant balance and equilibrium difficulty than does loss of vestibular function. With bilateral vestibular function loss, difficulties occur only when one of the other systems is disrupted (eg, when walking in the dark or on a soft surface) or when balance must be maintained under particularly difficult conditions (eg, walking on a narrow beam).[201] Thus, in humans, under physiologic conditions, the vestibular system is probably the least important of the three balance and equilibrium sensory systems. The most significant functional deficits occur when the vestibular system suffers acute, asymmetric damage and generates "false" head position or head rotation signals.

Vestibular Apparatus

The top portion of Figure 4–2 illustrates the general anatomic plan of the vestibular apparatus. There are many similarities with cochlear anatomy. For example, both end-organs are located in a tunnel that is hollowed out of the petrous portion of the temporal bone (embryologically, both organs come from the same tunnel). The tunnel is divided into an outer perilymph-filled bony labyrinth and an inner endolymph-filled membranous labyrinth. In addition, as in the cochlea, the receptor cells of the vestibular apparatus are ciliated, and these cilia extend into a gelatinous matrix.

The three semicircular canals are oriented orthogonally or at right angles to each other. They can be thought of as lying in a bottom corner of a box. The horizontal (or lateral or external) canal is in the plane of the bottom of the box, and the anterior-vertical (or superior) and posterior-vertical (or posterior) canals are in the planes of the two sides of the box. In humans, the entire canal complex is tilted upward about 30 degrees. In the physiologic position, the head is bent forward about 30 degrees from the earth's horizontal plane; therefore, the 30-degree upward tilt puts the horizontal canal in the horizontal position under everyday conditions.[202] The superior and posterior canals are oriented in vertical planes that form an angle of approximately 45 degrees with the sagittal head plane. Each semicircular canal lies parallel to one of the canals in the opposite vestibular labyrinth. Thus, the right horizontal canal is coplanar with the left horizontal canal, whereas the right superior and posterior canals are coplanar with the left posterior and superior canals, respectively.

The sensory epithelia or cristae of the semicircular ducts are located in enlarged areas, referred to as ampullae, at one end of each duct. The nonampullated ends of the vertically oriented superior and posterior ducts connect to form the common crus. The bony vestibular aqueduct originates in the vestibule (ie, the entrance to the inner ear) and connects to the cerebrospinal fluid space medially. Within the vestibular aqueduct is the fibrous endolymphatic duct that establishes a route between the membranous vestibule and the cranial meninges, where the duct ends in the endolymphatic sac. The utriculosaccular duct, connecting the utricle and saccule, permits communication between the saccule and endolymphatic duct. Although there is no agreement concerning the site of endolymph absorption, it is likely that the K^+-rich fluid is absorbed in the endolymphatic sac.[203] The ductus reuniens connects the saccule with the cochlear duct.

Figure 4–41, A, shows the anatomy of the crista. It is a saddle-shaped mound of tissue, attached to the ampullar wall at right angles to the long axis of the ampulla. The hair cells are on the surface of the crista. The ampullar nerve fibers travel

through the center of the crista to synapse at the hair cell bases. The hair cell cilia protrude from the surface of the crista into the fan-shaped cupula, a gelatinous structure consisting of mucopolysaccharides within a keratin framework.[204] The cupula partitions the semicircular duct by covering the top of the crista and extending to the opposite ampullar wall.

The sacculus and utriculus are two sacs in the membranous labyrinth, located in the vestibule. Their receptor organs, called maculae, can be seen in Figure 4–41, B, as patches of epithelia on the membranous labyrinth walls. The utricular macula lies on the floor of the utricle approximately in the plane of the horizontal semicircular canal. The saccular macula lies on the anteromedial wall of the saccule and is oriented principally in the vertical plane.

Figure 4–42 shows the structure of the macula. It consists of hair cells that are surrounded by supporting cells. The hair cell cilia are attached to a gelatinous otolithic membrane. On the top of the gelatinous membrane is a layer of calcium carbonate crystals called statoconia or otoliths. The otoliths are denser than the surrounding endolymph[205]—hence their ability to respond to gravity and inertial forces.

Figure 4–43 summarizes in a schematic form some TEM observations of vestibular hair cell morphology. First, as shown by Figure 4–43, A, there are two types of hair cells in both the macula and the crista.[206] Type I hair cells are flask-shaped and have "chalice"-type afferent nerve endings that surround all but the hair-bearing end. Efferent nerve terminals synapse with the afferent calyx near its base. Type II hair cells have a cylindrical or test tube shape and have several small bouton-type nerve endings that represent both afferent and efferent innervation only at the cell's base. Type I hair cells are concentrated in the central apex of the crista and the central part of the macula; type II hair cells are more numerous toward the peripheral region of the end-organ.[207] Although significant amounts of filamentous actin occur at the apical surfaces of both the sensory cells and the supporting cells in the hair cell–containing regions of the vestibular organs,[208] motile properties such as the fast rates of lengthening and shortening, which have been observed for cochlear OHCs, have not been shown for vestibular hair cells. However, there is some evidence that slow motility is exhibited by type II hair cells.[209]

The electron microscope has also demonstrated two types of cilia: stereocilia and kinocilia[210,211] (Figure 4–43, B). Each hair cell has only one motile kinocilium and a bundle of 60 to 100 stereocilia.

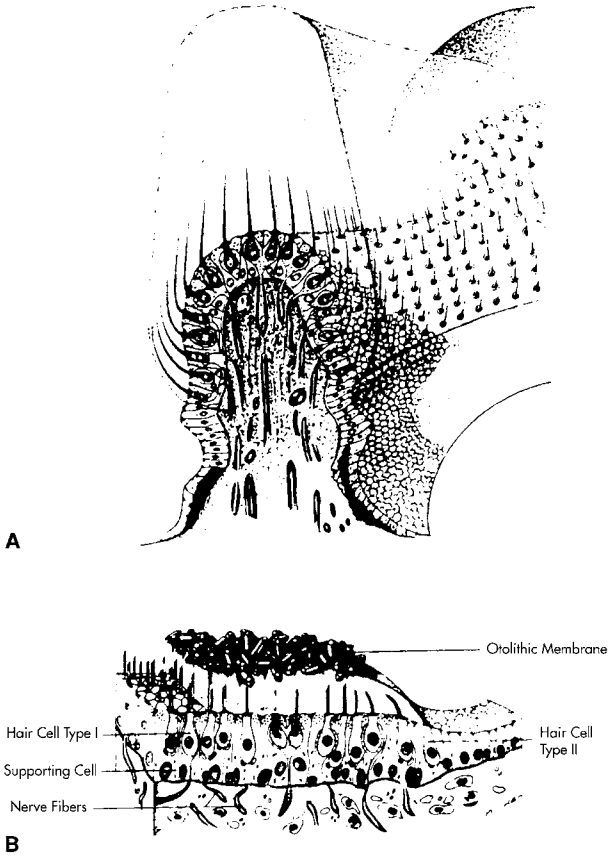


FIGURE 4–41. Sensory receptors of the vestibular system. *A*, Schematic drawing of the crista ampullaris illustrating the sensory cells, their hair bundles (cilia) that protrude into the cupula, and their innervating nerve fibers. *B*, Schematic drawing of a macula, showing how the cilia of the hair cells are embedded in a gelatinous membrane to which are attached calcium carbonate crystals (ie, otoconia). *A* reproduced with permission from Wersall J. Studies on the structure and innervation of the sensory epithelium of the cristae ampullaris in the guinea pig. Acta Otolaryngol Suppl (Stockh) 1956;126:1. *B* reproduced with permission from Iurato S. Submicroscopic structure of the inner ear. Oxford (England): Pergamon Press; 1967.

Stereocilia are relatively rigid, club-like rods, varying systematically in length in a "pipe organ" fashion, that extend, at the apical end of the hair cell, from a dense cuticular plate that consists of actin and myosin.[33] They are not homogeneous, as was originally thought. Rather, they contain longitudinally oriented microfilaments[212] composed of

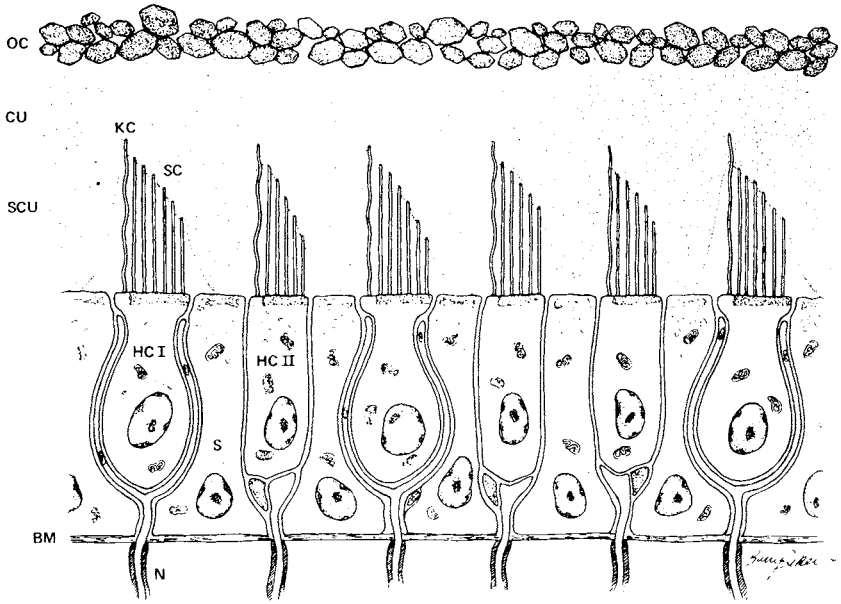


FIGURE 4–42. Schematic drawing of a cross-section of a macula. The gelatinous substance is divided into cupular (CU) and subcupular (SCU) layers. There are two types of cilia: kinocilia, labeled KC (one per hair cell), and stereocilia, labeled SC (many per hair cell). OC = otoconia; HC I = type I hair cell; HC II = type II hair cell; N = nerve fiber; BM = basement membrane; S = supporting cell.

actin.[32] The membrane enveloping the stereocilia is a thickened continuation of the hair cell cuticular membrane.[212] The stereocilia are constricted at their base and, when deflected, move like stiff rods pivoting around the basal constricted area.[36] Interconnections between stereocilia via fibrils or tip links have been demonstrated.[213]

Kinocilia end in basal bodies, located just beneath the hair cell membrane. Within each kinocilium, there are nine peripherally arranged double-tubular filaments, positioned regularly around two centrally located tubular filaments. This 9-plus-2 tubule pattern is found in many motile cilia (eg, respiration epithelia, oviduct epithelium, unicellular flagellates).[214] Stereocilia are coupled to the kinocilium and to each other so that during deflection, all cilia are stimulated as one bundle.[49]

On each hair cell, the single kinocilium is located to one side of the bundle of stereocilia. In both maculae and cristae, hair cells in the same area tend to have kinocilia on the same side of the stereocilia bundle. Thus, the vestibular sensory epithelium has a morphologic directional polarization[210] that is determined by the direction of the cilia alignment. Figure 4–43, C, illustrates the directional polarization of the utricular macula. The kinocilia all tend to point toward a line of demarcation called the striola, or linea alba, running across the approximate center of the macula. In the saccule, the kinocilia point away from the striola. In the horizontal duct crista, kinocilia point toward the utricle, whereas in the pair of vertical cristae, they point away from the utricle. Thus, vertical and horizontal cristae are morphologically polarized in opposite directions.[207]

Transduction and Coding

Mechanical Events As in the cochlea, the final mechanical event in vestibular transduction is the bending of the hair cell cilia. Figure 4–44 illustrates transduction by the macula. When the macular surface is tilted, the heavy otoliths tend to slide downward, carrying the gelatinous membrane and attached cilia with them.

The six semicircular ducts consist of a circular, narrow-bore tube in the temporal bone filled with fluid called endolymph. The tube originates from and returns to a reservoir (the utricle), but each duct may be treated, for practical purposes, like a fluid ring. When the head undergoes an angular acceleration, the fluid is left behind because of its inertia. This causes the endolymph to flow relative to the duct and to push on the cupula, which lies across the duct blocking it. The cupula is attached to the ampulla wall around its entire periphery and billows, like a sail, under endolymph pressure.[215,216] Normal deflections are tiny, in the range of 0.01 to 0.3 mm.[217] The subsequent strain on the hair cells by the bending of their cilia embedded in the cupula creates a generator potential that modulates the discharge rate of primary vestibular afferent fibers.

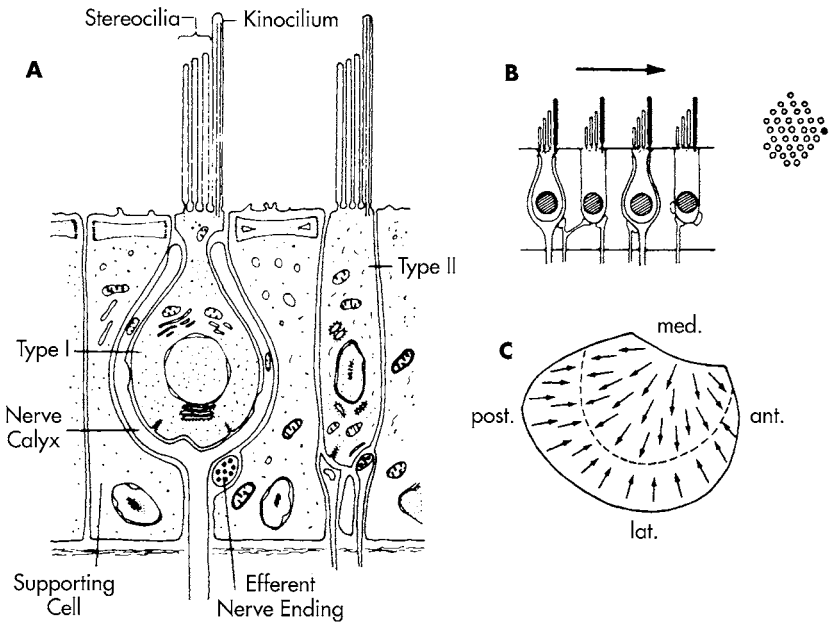


FIGURE 4–43. Microanatomy of the vestibular hair cells. *A*, Schematic drawing of the two vestibular hair cell types. *B*, Diagram showing morphologic polarization of the hair cells. *Arrow* shows the direction of cilia bending that produces excitation. At right is a cross-section through the hair-bearing end of the hair cell showing many stereocilia (*open circles*) and one kinocilium (*filled circle*). *C*, A diagrammatic surface view of the human utricular macula. *Arrows* indicate direction of polarization. *Dashed line* is the linea alba. *A* reproduced with permission from Brodal A.[223] *B* and *C* reproduced with permission from Lindeman HH, Ades HW, Bredberg G, Engstrom H. The sensory hairs and the tectorial membrane in the development of the cat's organ of Corti. Acta Otolaryngol (Stockh)1971;72:229–42 and Wersall J et al, respectively.[206]

The overall behavior of the cupula and duct was first described by Steinhausen.[218] If c is some measure of cupular (and thus endolymph) displacement, the force balance equation for the duct is:

$$m[c(d^2c/dt^2)] + r[c(dc/dt)] + kc = m[H(d^2H/dt^2)]$$

where m represents the net, equivalent moment of inertia of the endolymph, r is the lumped effect of viscous drag on fluid flow (c), and k represents the stiffness of the gelatinous cupula. The term $m[H(d^2H/dt^2)]$ is the driving force on the endolymph owing to head acceleration. During most natural head movements, the inertial reactance, $m[c(d^2c/dt^2)]$, and the cupular restoring force, kc, are very small compared to the viscous term, $r[c(dc/dt)]$, because the narrowness of the tube creates a high resistance to flow. Consequently, the above equation can be reduced to:

$$r[c(dc/dt)] = m[H(d^2H/dt^2)]$$

or, integrating both sides:

$$c = m/r\ [H(dH/dt)].$$

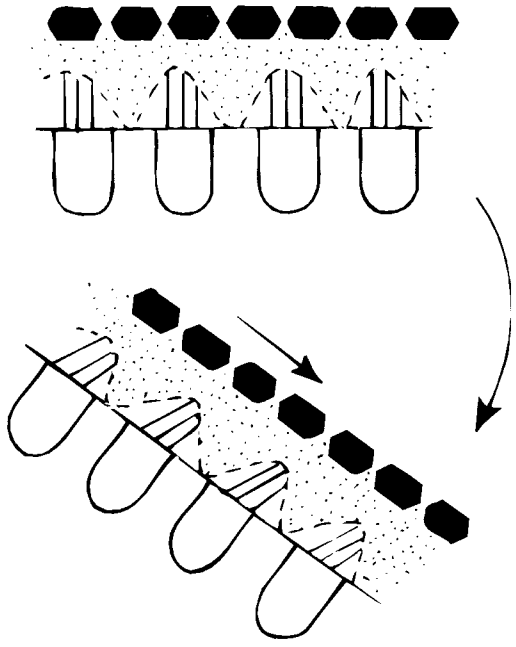


FIGURE 4–44. Diagrammatic representation of the conversion of a head tilt to bending of the hair cell cilia by the otoconia.

Thus, the discharge rate, which is proportional to cupular deflection, carries a signal into the central nervous system proportional to head velocity, not acceleration. In other words, a constant head acceleration applies a constant force on the endolymph, causing it to flow at a constant velocity. Consequently, cupula position, the integral of flow, must be proportional to head velocity (ie, the integral of head acceleration).

In the squirrel monkey, the typical resting discharge rate is 90 spikes/s.[219] The discharge rate increases for deflection of the cupula in one direction and decreases in the other. On the abscissa of Figure 4–45, B, B(dH/dt) may be substituted for c, in which case the slope of the line, Sv, for the average fiber is 0.4 (spikes/s)/(deg/s). Thus, the signal sent from the duct to the brainstem on the average fiber for most normal head movements is 90 + 0.4 [H(dH/dt)].

Response of Primary Vestibular Neurons Microelectrode studies in higher mammals have demonstrated the following characteristics of primary vestibular neuron activity. For example, most of the primary vestibular neurons called regular units exhibit a high (~100/s) and remarkably uniform spontaneous discharge.[220] When listening to such spikes on a loudspeaker, the neurons make a characteristic motorboat sound. There is also a small population of irregular neurons with a lower rate and less regular spontaneous discharge. The high-rate regular units have small-diameter, slowly conducting axons that predominantly innervate the periphery of the end-organ, that is, the type II hair cell region of the crista, whereas the low-rate irregular neurons come from the center of the crista, where most type I hair cells are located.[197] The high, regular spontaneous firing rate of the primary neurons permits the bidirectional sensitivity of the vestibular hair cell receptors.

When the head rotates, causing cupular displacement and a stimulatory deflection of hair cell cilia, the semicircular duct afferents change their discharge rate above, in response to hair cell depolarization, and below, in response to hair cell hyperpolarization, the resting rate depending on the direction of rotation.[221] When the head rotates sinusoidally at velocities encountered during normal function, ampullar nerve discharge rate varies sinusoidally, as shown in Figure 4–46. The phase relation-

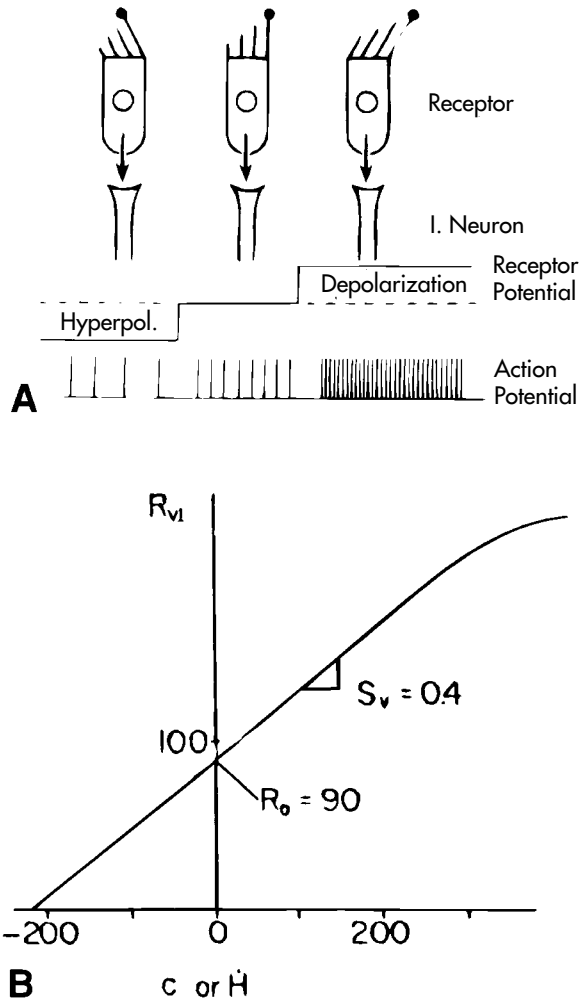


FIGURE 4–45. Motion transduction by vestibular hair cells. *A*, At rest there is a resting rate of action potential discharge in the primary vestibular afferents (*center*). Shearing forces on the hair cells cause either depolarization (*right*) or hyperpolarization (*left*), depending on whether the stereocilia are deflected toward or away from the kinocilium (indicated by longest cilium, with beaded end), respectively. This modulates the discharge rate in the vestibular nerve as shown below (action potential). *B*, Vestibular nerve discharge rate (R_{v1}) is 90 spikes/s at rest (R_o) in the squirrel monkey and changes approximately linearly with cupula deflection (c). When the latter is expressed in equivalent head velocity (H), the slope of the line S_v is about 0.4 (spikes/s)/(deg/s) for the average vestibular afferent fiber. *A* reproduced with permission from Baloh R, Honrubia V. Clinical neurophysiology of the vestibular system. 2nd ed. Philadelphia: FA Davis; 1990. *B* reproduced with permission from Zee DS, Leigh RJ. The neurology of eye movements. 2nd ed. Philadelphia: FA Davis; 1991.

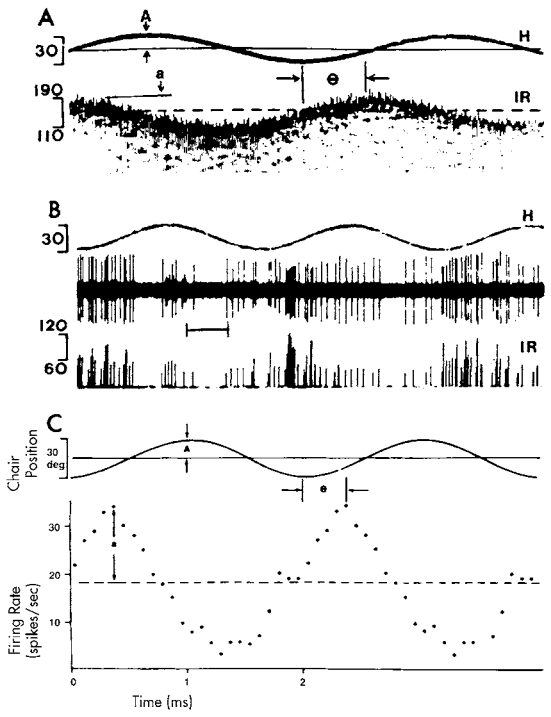


FIGURE 4–46. Peripheral vestibular neuron responses to sinusoidal rotation stimulation in the alert monkey. *A*, A regular (high-rate) neuron. *B*, An irregular (IR) neuron. Sinusoidal curve (H) plots head position. The records immediately below the sinusoidal traces show intervals between neuronal spikes. The height of vertical bars on these records is proportional to the interval between spikes. Because one of these bars occurs for every spike, their frequency reflects the frequency of the recorded spikes. *C*, Response of unit in *B* averaged over 10 cycles; 0 is the phase lag of firing rate (ie, the shortest interval between spikes), which occurs at about the 0 crossing of the head-position signal. Thus, firing rate at the output of the semicircular duct is related most closely to head velocity.

ship between head position and head velocity, which displays a phase lag of about 90 degrees, is such that the maximum discharge rate occurs at the head's zero crossing (ie, at the head's maximum velocity). Thus, the ampullar afferent discharge rate codes head angular velocity. The adequate stimulus is actually angular acceleration, but the hydrodynamics of the semicircular duct system are such that the head's acceleration is integrated to give a velocity output.[222]

Similarly, with the macular afferents, accurate coding by discharge rate of the head's position relative to gravitational vertical can be demonstrated in all species. The more complex orientation of the macular hair cells makes directional correlates more difficult to establish than for semicircular duct afferents. However, Fernandez and Goldberg[197] studied large populations of saccular and utricular neurons and were able to demonstrate neural "sensitivity vectors" among the neuronal populations that corresponded to the relative orientations of the saccular and utricular maculae.

Central Vestibular Pathway

Anatomy *Primary Afferent Connections.* The vestibular system functions primarily as an afferent reflex input to the motor system. In general, vestibular pathway–mediated reflexes involve three muscular systems, including the extrinsic oculomotor, cervical, and antigravity pathways.

As might be expected, the semicircular ducts, or the rotation sensors, connect primarily with the extrinsic oculomotor and cervical muscles (ie, the muscles that compensate for head rotation), whereas the otolithic organs, or position sensors, connect primarily with the antigravity muscles. Figure 4–47 outlines these major central vestibular connections.[223]

Scarpa's, or the vestibular, ganglia, in the internal auditory canal, within the vestibular portion of the eighth nerve, contain the bipolar ganglion cells of the first-order vestibular neurons. The vestibular nerve can be divided into superior and inferior divisions that innervate the sensory epithelia of the duct and otolithic end-organs. The superior portion innervates the cristae of the superior and horizontal semicircular ducts, the utricular macula, and a small region of the saccular macula. The inferior division of the vestibular nerve innervates the crista of the posterior semicircular duct and the remaining part of the saccular macula. Centrally, all first-order vestibular neurons synapse in the vestibular nuclear complex, which occupies a considerable area beneath the floor of the fourth ventricle and lies across the pontomedullary boundary. On entering the brainstem, the vestibular axons bifurcate into the ascending rostral and descending caudal divisions.

The vestibular nuclear complex, as shown in Figure 4–47, consists of four distinct subnuclei: (1) the superior (Bechterew's or angular), (2) lateral (Deiters'), (3) medial (Schwalbe's or principal or triangular), and (4) inferior (spinal or descending) nuclei. There is recent anatomic evidence that the

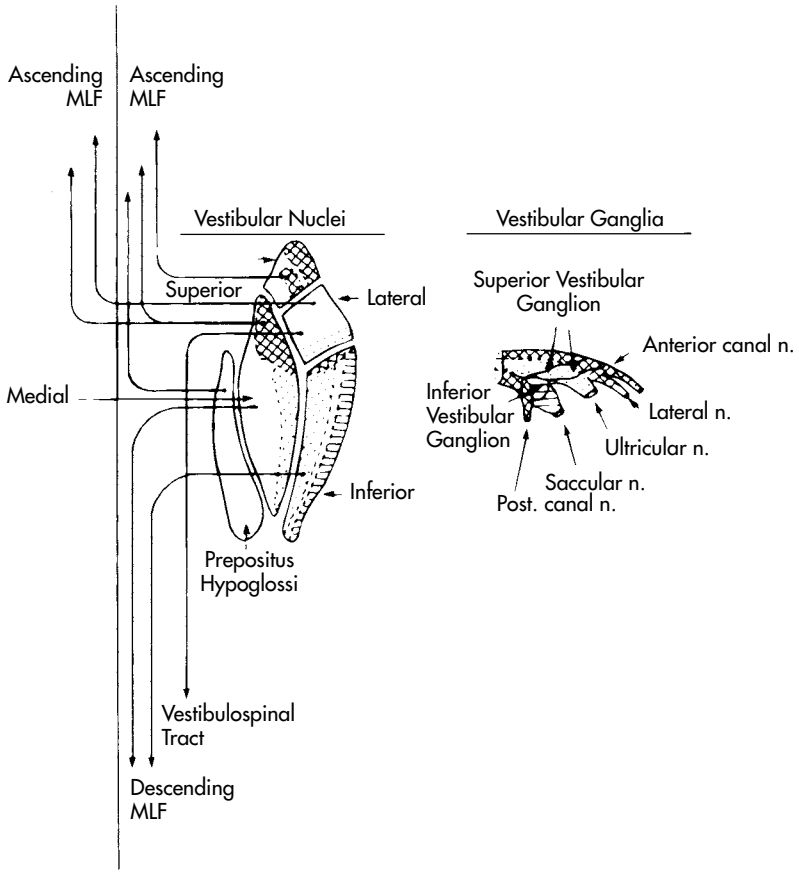


FIGURE 4–47. Diagram of major connections between primary vestibular neurons in the vestibular ganglia and the vestibular nuclei. Also shown are major outflows from the vestibular nuclei. Vestibulocerebellar connections are excluded. The primary afferent connections are modified from Carpenter MG. Human neuroanatomy. 7th ed. Baltimore: Williams and Wilkins; 1976. MLF = medial longitudinal fasciculus.

nucleus prepositus hypoglossi, which is classically thought to relate to taste sensation, has strong efferent and afferent vestibulo-oculomotor, through the medial longitudinal fasciculus (MLF), and cerebellar projections. In addition, microelectrode recordings from awake monkeys have demonstrated that prepositus neurons fire in relation to both vestibularly and visually induced eye movements.[224] Studies on the activity of ocular motoneurons in alert monkeys have shown that the neural commands for all conjugate eye movements (vestibular, optokinetic, saccadic, and pursuit) have both velocity and position commands.[225]

The vestibular input to prepositus is disynaptic and reciprocally organized (ie, ipsilateral input is excitatory, whereas contralateral input is inhibitory). In addition, if cells in the monkey's nucleus prepositus hypoglossi are lesioned with kainic acid, the monkey's eye movements only reflect the velocity command.[226] A great amount of neurophysiologic evidence indicates that the position command is obtained from the velocity command by the mathematical process of integration. That is, a neural network located in the nucleus prepositus hypoglossi integrates, in the mathematical sense, velocity-coded signals; thus, the prepositus is included in the diagram of the vestibular nuclear complex in Figure 4–47.

As indicated in Figure 4–47, each subnucleus has a unique set of connections with the periphery and with specific regions of the central nervous system including the spinal cord, cerebellum, and brainstem oculomotor nuclei (III, IV, VI). Most of the semicircular duct afferents terminate in the superior nucleus and rostral portion of the medial-vestibular nucleus. Both of these nuclei, in turn, project to the oculomotor nuclei of the extraocular muscles via the ascending MLF. The superior nucleus projects only to the ipsilateral MLF, whereas the medial nucleus projects bilaterally. The medial nucleus, via the medial vestibulospinal tracts in the descending MLF, also sends bilateral descending projections to spinal anterior horn cells that control the cervical musculature.[227] Thus, the input and output of the superior and medial vestibular nuclei provide a possible anatomic basis for the nystagmus and head-turning-reflex responses to semicircular duct stimulation (see below).

The otolith organs, particularly the utriculus, project primarily to the inferior nucleus and caudal part of the lateral nucleus. Outputs from these

nuclei, in turn, project downward to the ventral horn region throughout the length of the spinal cord via the lateral vestibulospinal tract. Thus, the afferent and efferent connections of the lateral vestibular nucleus provide a possible anatomic basis for antigravity muscle responses in the limbs (extensors of the legs and flexors of the arms) to postural change. The otolith organs have relatively sparse connections to the extraocular muscles. These may be the connections that produce the ocular counter-rolling response to head tilt.[228]

Efferent Innervation of the Vestibular System The vestibular organs in all vertebrates receive efferent innervation originating in brainstem nuclei. Although there are relatively few parent efferent neurons, their axons branch to innervate more than one end-organ and also ramify extensively in the neuroepithelium of the individual organs. By means of this divergent innervation pattern, efferent fibers provide a major source of synaptic input to type II hair cells, afferent calyces, and other unmyelinated afferent processes. In nonvestibular organs, efferent activation typically reduces afferent activity by hyperpolarizing the hair cell receptor. In contrast, efferent responses in vestibular organs are more heterogeneous. For example, in fish and mammals, there is an excitation of afferents, whereas in frogs and turtles, both excitation and inhibition are observed. In addition, there are variations in the efferent responses of vestibular afferents innervating different parts of the neuroepithelium, which differ in their responses to natural stimulation. Moreover, efferent neurons receive convergent inputs from several vestibular and nonvestibular receptors and also respond in association with active head movements. On the basis of the discharge properties of efferent neurons, one of their proposed functions is that they switch the vestibular organs from a postural to a volitional mode. For a more detailed review of the vestibular efferent system, see a recent summary by Goldberg.[229]

Vestibulocerebellar Connections *Primary Vestibular Fibers.* Primary vestibular neurons project not only to the vestibular nuclei but also to the cerebellum. Most of these fibers are distributed to the ipsilateral flocculus and nodulus and the medially located uvula.[230] Because of this innervation by primary vestibular fibers, these three cerebellar areas have been termed collectively the vestibulocerebellum. As discussed below, the primary vestibular input to the vestibulocerebellum appears to be important in controlling the vestibulo-oculomotor reflex (VOR). Moreover, the primary vestibular input to the vestibulocerebellum and the climbing fiber input from the inferior olive appear to be important in controlling the VOR.

Secondary Vestibular Fibers. The vestibulocerebellum receives secondary fibers primarily from the medial and inferior vestibular nuclei but also from the other divisions. In addition, the fastigial nucleus and the cortex of the vermis receive a strong, somatotopically organized projection from the lateral vestibular nucleus. Since this nucleus is the primary origin of the vestibulospinal tract, connections to it from the cerebellum are probably important in regulating antigravity reflexes that help to maintain an upright body posture.[231]

Projections to Cerebral Cortex Whether the vestibular system has a direct cortical projection has long been a controversial question. The functional corollary to this question is whether we can consciously appreciate a sensation owing to vestibular stimulation. This question also has generated debate because most of the subjective sensations produced by vestibular stimulation (eg, vertigo) are secondary to motor reflex or to autonomic responses, and it is difficult or impossible to separate a primary vestibular sensation from the secondary sensations.[200]

Experiments employing electrical stimulation of the vestibular nerve have demonstrated relatively short-latency localized cortical responses in both the cat and the monkey.[232] The cat's vestibular area is adjacent to both the auditory and somatosensory fields, but in the monkey, and possibly in the human, it is located near the face area of the somatosensory field on the postcentral gyrus. Deecke et al demonstrated that the ventroposteroinferior (VPI) nucleus is the thalamic relay for the vestibulocortical projection.[233] The VPI nucleus lies adjacent to the thalamosomatosensory representation of the face. Prior to this study, the function of the VPI nucleus was unknown.

Vestibular Influence on Postural Control The main unit for the control of tone in the trunk and extremity muscles is the myotactic reflex. These reflexes of the antigravity muscles are under the combined excitatory and inhibitory influence of multiple

supraspinal centers. Two of these supraspinal centers are facilitory, that is, the lateral vestibular nucleus and rostral reticular formation, and four are inhibitory centers including the pericruciate cortex, basal ganglia, cerebellum, and caudal reticular formation. The balance of input from these different centers determines the degree of tone in the antigravity muscles.

Physiology of Vestibular Nuclei *General Characteristics of Neuronal Responses.* Microelectrode studies of responses of vestibular nucleus neurons to electrical stimulation of individual ampullary nerves and to "natural" stimuli, such as those involving rotation and tilt, have established the following general characteristics: (1) neurons can be classified as tilt responders (otolith units) or rotation responders (duct units); (2) the locations of these two types correspond with that expected from the anatomic projections of otolithic and semicircular duct afferents to the vestibular nuclear complex. Thus, the tilt responders are found primarily in the areas innervated by the sacculus and utriculus, that is, the inferior nucleus and caudal part of lateral nucleus, whereas the rotation responders are found mainly in the areas innervated by the semicircular ducts, that is, the superior nucleus and the rostral part of medial nucleus; and (3) among the rotation responders, pathways from individual ducts seem to be preserved. Thus, for example, a neuron that responds to vertical rotation or to electrical stimulation of a vertical ampullary nerve does not respond to horizontal rotation or to stimulation of a horizontal ampullary nerve and vice versa.

Classification of Neuronal Response Types. Both duct and otolithic vestibular nucleus units preserve the basic properties of the primary afferent input, but some respond in the same direction (eg, ipsilateral horizontal rotation increases firing rate); others respond in the opposite direction. A few of the vestibular nucleus neurons respond with either inhibition or facilitation in both directions. Table 4–2 outlines the characteristics of the duct units. The type I units, exhibiting a primary response pattern, are innervated by the ipsilateral labyrinth, whereas the type II units, showing response patterns that are opposite to those of the primary units, are innervated by the contralateral labyrinth.

The analogous classification for otolith responders subdivides all types into (1) "a" units, which increase firing on ipsilateral tilt; (2) "b" units, which decrease firing on ipsilateral tilt; (3) "y" units, which increase firing in response to tilt in both directions; and (4) "o" units, which decrease firing by tilt in both directions. Analogous to the duct units, the relative populations of these units are a > b > y > o.[234]

Commissural Inhibition and Vestibular Compensation Many of the contralateral (type II) semicircular duct responses listed in Table 4–2 are abolished when a midline dorsal brainstem incision is made that interrupts the vestibular crossing fibers. Thus,

TABLE 4–2. Characteristics of Duct Responding Units in Vestibular Nuclei

	Input Side	*Response Directionality*	*Response Time Course*	*Latency*	*Gain*	*Spontaneous Activity*
Type I (67%)	Ipsilabyrinth	Same as 1 degree	—	—	—	—
Kinetic (56%)	Ipsilabyrinth	Same as 1 degree	Fast decay	Long (multisynaptic)	High	None
Tonic (11%)	Ipsilabyrinth	Same as 1 degree	Slow decay	Short (monosynaptic)	Low	High level
Type II (29%)	Contralabyrinth	Opposite 1 degree				
Type III (3%)	?	Facilitated both directions				
Type IV (< 1%)	?	Inhibited both directions				

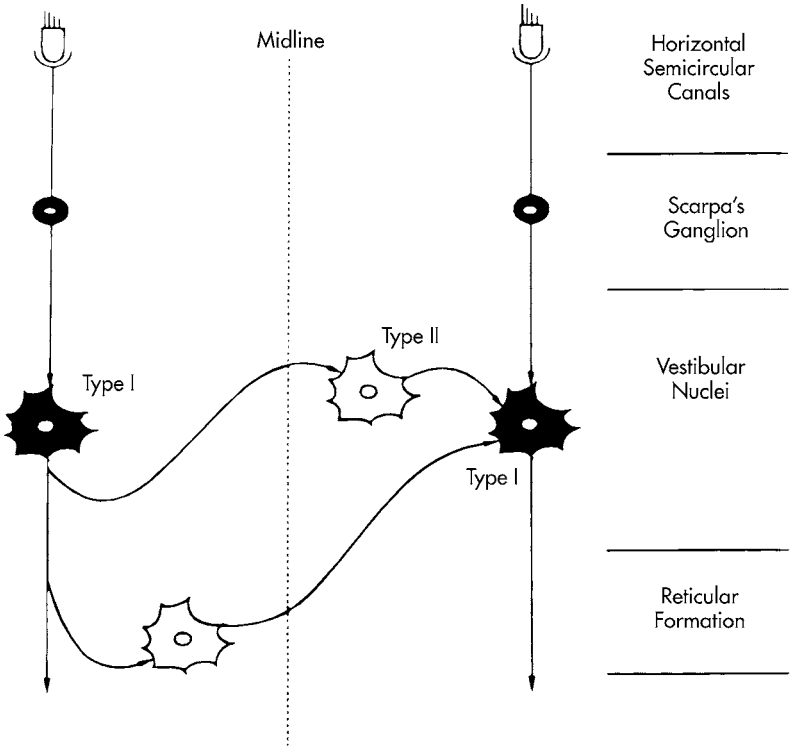


FIGURE 4–48. Interrelation of type I and type II secondary vestibular neurons. Dark neurons are excitatory, and light neurons are inhibitory. Reproduced with permission from Baloh R, Honrubia V. Clinical neurophysiology of the vestibular system. 2nd ed. Philadelphia: FA Davis; 1990.

the "opposite-primary" type II responses are likely elicited by inhibitory input from type I neurons of the opposite labyrinth. Figure 4–48 summarizes the simplest of the probable neuronal interconnections responsible for this contralateral inhibition.

The commissural inhibitory system is also important in the mechanism of compensation following labyrinthectomy. Thus, immediately after labyrinthectomy, the type I units on the labyrinthectomized side show no spontaneous activity and do not respond to rotation. However, within a few days, the deafferented type I units regain their spontaneous activity and, via inhibition from the contralateral pathway, also regain their normal response to rotation. The mechanisms by which this recovery process occurs are unknown, but a contributing factor is the lowered threshold of the deafferented type I neuron for contralateral input.[235] In animal studies, the course of compensation is affected by exercise,[236] visual experience,[237] and drugs.[238] Thus, as a rule, stimulants accelerate and sedatives slow compensation.[237] Fetter and Zee showed further that visual experience was not necessary for the acquisition or for the maintenance of this recovery process.[237]

If a second labyrinthectomy is performed after compensation for the first labyrinthectomy, the animal again develops signs of acute unilateral vestibular loss with nystagmus directed toward the previously operated ear, that is, a condition referred to as Bechterew's compensatory nystagmus.[239] That is, the overall effect is as if the first labyrinthectomy had not occurred. Compensation after the second labyrinthectomy is slightly faster than the first but still requires several days. For a recent review on the molecular mechanisms underlying vestibular compensation, see the summary by Darlington and Smith.[240]

Vestibulo-oculomotor Reflex Although there are neural connections between the maculae and the extraocular muscles, they are less important in humans, both functionally and clinically, than are the semicircular duct connections. This discussion of the VOR will therefore focus on the reflex connection between the semicircular ducts and the extraocular muscles.

An important function of the semicircular ducts is to provide afferent input to the VOR that generates eye movements that compensate for head movements by maintaining a stabilized visual image on the retina during rotation. Such eye adjustments are called compensatory eye movements. Figure 4–49 diagrams the neural connections that stabilize gaze during head movement.

Anatomic Connections. The VOR is subserved by a three-neuron pathway. That is, motions detected by the end-organ are transduced into neural impulses that are sent via the vestibular nerve to the vestibular nuclei and rostrally through the ascending MLF to the oculomotor nuclei of the extraocular muscles. Secondary vestibulo-ocular connections via the reticular formation have been described, but functionally these are less important than the "direct" three-neuron MLF vestibulo-ocular pathway.

Inhibitory crossed connections at the levels of both vestibular and oculomotor nuclei, which were described above, probably also participate in the VOR. The crossed inhibitory oculomotor connections subserve antagonistic extraocular muscles.[241] These inhibitory interconnections are important in the formation of conjugate eye movements.

Functional Connections. A number of investigations have uncovered many details of the facilitatory and inhibitory interactions between neurons of the vestibular nuclear complex and motoneurons of the extraocular eye muscles. For a review, the reader is referred to Wilson and Melvill-Jones.[234] However, in older experiments, the effect of mass stimulation of individual ampullary nerves on eye movements provided the clinician with the most useful integrative concepts of the functional organization of the VOR. The results of all of these ampullary nerve-stimulating experiments agree on the principle, illustrated by Figure 4–49, that stimulation of an ampullary nerve generates conjugate eye movements away from the side stimulated and in the plane of the duct stimulated.

This principle allows a straightforward description of the events leading from endolymph movement in the horizontal duct to the compensatory eye movement depicted in Figure 4–50: (1) head movement to the left generates endolymph movement to the right in the left duct. This is toward the ampulla and therefore increases the ampullary nerve discharge rate; (2) the increased ampullary nerve output, according to the aforementioned rule, causes conjugate eye deviation away from the side stimulated (ie, to the right, which is opposite to the direction of head movement); (3) similarly, head movement to the left generates ampullofugal movement in the right duct, which decreases neural output; and (4) decreased neural output inhibits the extraocular muscles, thus causing the eyes to deviate conjugately to the left, which provides a neural output that is synergistic to the conjugate right-eye deviation, generated by the output of the opposite duct. Figure 4–49, upper right, depicts this sequence of events.

Functionally, the VOR provides more than a simple one-to-one transfer of information from the semicircular duct to the extraocular muscles. It should be recalled that the response of the ampullary nerves reflects head velocity, yet the ocular rotation, at the vestibular reflex output, reflects head position. Therefore, at some point in the transfer from vestibular input to ocular output, some form of neural integration must occur that converts the velocity signal at the input into the displacement signal at the output. The location of this integrator, as noted above, is the nucleus prepositus hypoglossi.[226]

Control of the Gain of the Vestibulo-oculomotor Reflex. In studies of the VOR, the concept of gain is important. Gain of the VOR is simply the amplitude of eye rotation divided by amplitude of the head rotation. Generally, amplitudes of eye and head rotation are expressed as angles.

Because eye rotation is supposed to be equal and opposite to head rotation (ie, to compensate fully and leave gaze steady), it might seem at first that the ideal VOR gain would be –1. However, there are many real-life situations for which the ideal VOR gain is not –1. Functionally, such situations can be divided into long-term (ie, slow) and short-term (ie, fast) adjustments. The short-term or fast VOR gain adjustments occur in the course of visual tracking tasks, in which both the head and the eyes follow the fixation target at varying relative velocities. For example, in a tracking task in which the head and eyes must move in the same direction, the VOR must be completely suppressed. In the clinical test situation, this visual suppression of the VOR is observed as suppression of caloric or rotational nystagmus during visual fixation.

An example of functionally beneficial long-term VOR gain control is the situation in which a prescription for new glasses has been received. The change in magnification of the visual image changes the speed with which the visual image moves across the retina (eg, increasing magnification would increase the speed of movement of retinal images). Therefore, to maintain a stable retinal image after new glasses have been obtained, the VOR gain must be adjusted for the new magnification levels. Such

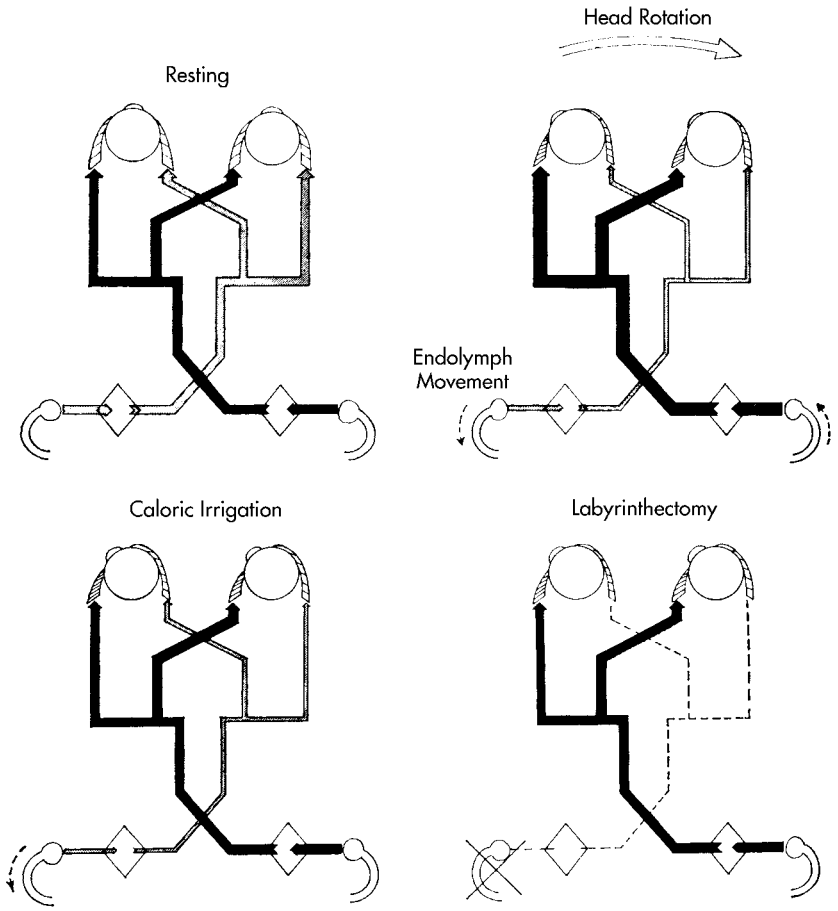


FIGURE 4–49. Generation of the vestibular (slow) phase of different kinds of nystagmus. The thickness of the lines connecting the semicircular duct to the extraocular muscles is proportional to the intensity of neural discharge along the nerve pathways.

adjustments have been demonstrated experimentally in humans by testing the VOR rotational response in the dark after long-term periods of wearing highly corrective lenses.[242]

An extreme example of the plasticity of the VOR gain is provided by adjusting to reversing prisms. Reversing prism glasses cause a normal visual image to be reversed right to left; hence, movement of the visual image, when the head or eyes are rotated horizontally, is exactly opposite of normal. Even in this extreme example, humans and other animals have been shown to function normally after a few days of wearing the prisms,[243,244] even in situations requiring visually controlled movements during head movements (eg, mountain climbing).[245] Testing the VOR in humans after a few days of wearing reversing prisms has demonstrated actual reversal (ie, conversion of a negative to a positive gain) of the VOR.[246]

There is strong evidence that the vestibulocerebellum plays an important role in both short-[247]

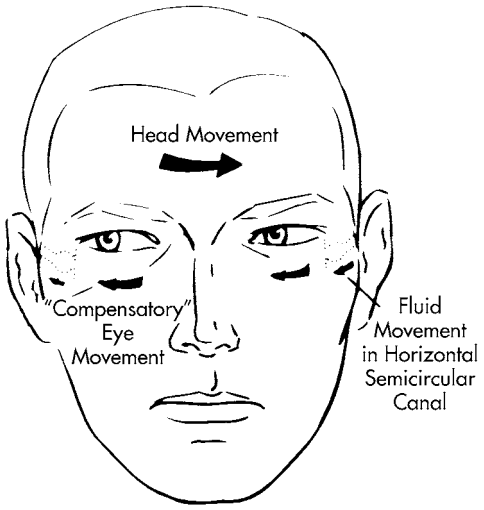


FIGURE 4–50. A compensatory eye movement initiated by the horizontal semicircular ducts. The eyes move in the direction of endolymph movement. If rotation of the head is continued at an ever-increasing speed, the compensatory eye movement becomes the slow phase of a rotational nystagmus.

and long-term[248] control of the VOR gain. The strong primary vestibular afferent projection to the vestibulocerebellum and equally strong projection from the vestibulocerebellum to the oculomotor portion of the vestibular nuclear complex have already been described. There is also a strong input to the vestibulocerebellum from the retina by way of the inferior olivary climbing fibers. Most of these visual units detected in the vestibulocerebellum are "movement detectors."[249] They are thus ideally suited to provide information on "retinal slip" of the visual image, which is needed to modify VOR gain control. The visual and vestibular inputs to the cerebellum converge on the Purkinje cells of the flocculus cortex. Furthermore, experiments on alert monkeys, in which rotation of the animal's visual environment and its head were controlled independently, showed that visual and vestibular inputs modulate Purkinje cell outputs in ways exactly appropriate for the functionally useful control of the VOR gain.

Ablation experiments further support the role of the flocculonodular lobes in control of VOR gain. It has been demonstrated, for example, that removal of the vestibulocerebellum eliminates suppression of caloric nystagmus by visual fixation.[250] An analogous failure of fixation suppression (FFS) of nystagmus is also observed in humans with cerebellar lesions.[251] However, clinical data suggest that FFS can be produced by lesions in the brainstem and, less frequently, cerebral hemispheres, as well as the cerebellum.[252]

Robinson reported the results of a notable ablation experiment that demonstrates the role of the flocculus in causing long-term or plastic changes in VOR gain.[244,248] Specifically, cats wore reversing prisms for several days, and VOR gain was measured each day by rotating the cats in the dark. After several days, as expected, VOR gain dropped from 0.9 to 0.1. When the vestibulocerebellum was removed, the VOR gain promptly increased to about 1.2. Furthermore, it proved impossible to reinstate the VOR gain decrease by continued wearing of prisms after removal of the vestibulocerebellum. In addition, Miles et al uncovered results in the monkey that implicated synaptic changes at other sites, especially in the brainstem.[252] More recently, Lisberger provided further evidence that the modifiable synapses were located between vestibular primary afferents and those second-order and/or possibly third-order vestibular neurons that also received synapses from Purkinje cells.[253] Finally, Luebke and Robinson reversibly silenced the flocculus of alert cats adapted to either high- or low-gain VOR.[254] This reversible floccular shutdown did not alter the adapted VOR gain, yet the animals were subsequently unable to modify their VOR gain, while the flocculus was silenced.

Clinical Considerations of Vestibular Reflexes

General Pattern. It has been stated that the vestibular system functions primarily as an afferent input for motor reflexes and that, at rest, most primary vestibular neurons have high and remarkably regular spontaneous discharge rates. For the purposes of the following clinically oriented discussion, the spontaneous vestibular neuronal activity will be referred to as the resting discharge.

Normally, with the head at rest in the neutral position, the resting discharges in the two vestibular nerves are equal. Vestibulomotor reflexes are elicited when inputs from the two vestibular end-organs or their central projections are made unequal (ie, they are unbalanced). Such unbalancing can occur by an abnormality involving one side to a greater degree than the other. In this case, we term the resultant vestibular reflex spontaneous. Vestibular input from the two sides can also be unbalanced by stimulating one or both of the vestibular end-organs. In this case, the resultant vestibular reflex is termed "induced."

When a vestibular input imbalance is unusually large or prolonged, a constellation of stereotyped responses occurs.[255] These responses are head turning, falling (or swaying), past pointing, vertigo representing a sensation of rotation, and nystagmus involving a rhythmic back-and-forth eye movement, with alternate slow phases in one direction and fast phases in the opposite direction. These slow and fast eye movements are termed the slow and fast phases of vestibular nystagmus. The head turning, falling, past pointing, and nystagmus slow phase are all in the same direction, that is, away from the ear with the greater output. In this discussion, the reference is to the nystagmus slow or vestibular phase. The clinical convention is to designate nystagmus direction by the fast of central phase, presumably because these movements can be easily observed. Thus, the indicated reflex directions may be the opposite of those to which the reader is accustomed. Fast-phase nystagmus is in the direction of the angular acceleration. The rotation sensation, or vertigo, may be

referred either to the subject's body (ie, subjective vertigo) or to the environment (ie, objective vertigo). If the body rotates, it is in the same direction as the nystagmus slow phase; if the environment rotates, it is in the opposite direction.

Methods of Eliciting. Vestibular responses can be generated by rotational, caloric, or galvanic stimulation of the vestibular periphery. The caloric and rotational responses are discussed in the following section. The galvanic vestibular response occurs when an electrical current is applied to the head in the vicinity of the ear.[251] Electrical current applied in this manner elicits the entire constellation of vestibular responses except vertigo. Interestingly, however, some subjects may report a slight sense of disorientation. If the current is positive, the nystagmus slow-phase, past-pointing, and body-sway responses are toward the stimulated ear; in contrast, if the current is negative, these responses are away from the stimulated ear.

Because the galvanic stimulus probably acts at a retrolabyrinthine location (ie, either vestibular nerve, Scarpa's ganglia, or a more central location), it has been suggested as a means of differentiating vestibular end-organ from vestibular nerve lesions. However, this procedure has not found widespread clinical acceptance, possibly because of uncertainty over the exact locus of action. For example, patients with their vestibular nerves sectioned intracranially still have a recognizable body-sway galvanic response,[256] suggesting that at least part of the action of the galvanic stimulus is central to the vestibular nerve (ie, at the level of the brainstem or possibly cerebellum).

Vestibular Nystagmus. Vestibular nystagmus occurs when the semicircular duct system is overstimulated. For example, if the head is continuously rotated in one direction at an ever-increasing speed, the semicircular duct–initiated compensatory eye movement becomes repeatedly interrupted by rapid, snap-back movements and hence becomes a rotational nystagmus. Because the relatively slow compensatory eye movement phase of the nystagmus comes from the semicircular duct system, it is called the vestibular or slow phase of the nystagmus. Conversely, because the fast, snap-back eye movement comes from the brain, it is called the central or fast phase of the nystagmus.

Caloric nystagmus is produced by a temperature change in the region of the vestibular apparatus. In the clinical laboratory, this temperature change is usually produced by running cool or warm water into the external auditory canal. The temperature change is conducted through bone to the semicircular canals. Since the horizontal semicircular canal is closest to the irrigating water, it is the most affected; hence, caloric nystagmus is almost entirely in the horizontal plane.

Figure 4–51 illustrates the generation of caloric nystagmus. The subject is in the supine position, with his head elevated 30 degrees to bring the horizontal canal into the vertical position. The cold caloric irrigation cools the endolymph in part of the horizontal duct. This cooled endolymph becomes slightly denser than the surrounding endolymph and thus falls, causing the eyes to rotate toward the irrigated ear. This vestibular eye movement is repeatedly interrupted by a compensatory rapid central eye movement in the opposite direction. A cold caloric nystagmus is thus created, with its slow phase toward and its fast phase away from the irrigated ear.

Although the slow phase is the vestibular phase of vestibular nystagmus, it is the clinical convention to designate the direction of the vestibular nystagmus by the direction of the fast phase. This convention, as noted above, originated because the fast phase is easier to observe than the slow phase.

Neural Mechanisms of Nystagmus. Figure 4–49 diagrams the neural mechanisms of different kinds

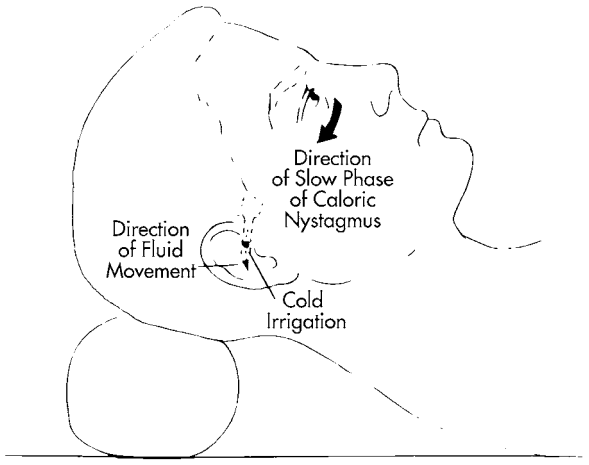


FIGURE 4–51. The mechanism of cold caloric nystagmus. The density of the cooled portion of the endolymph is increased, causing it to fall, thereby producing endolymph movement. The slow phase of the caloric nystagmus is in the direction of endolymph movement.

of vestibular nystagmus.[257] This figure shows the physiologic rather than the anatomic form of the horizontal VOR system. It is drawn to reflect the fact that increased neural activity from a horizontal semicircular duct rotates the eyes horizontally away from the duct.[258]

At rest, both ducts generate equal neural activity, which represents the resting discharge described above. Unbalancing the vestibular inputs from the two ears generates vestibular nystagmus if the imbalance is prolonged. Caloric and rotational stimuli unbalance the vestibular input via the normal transducer action of the semicircular ducts. Lesioning one labyrinth by surgery or disease causes a pathologic imbalance by eliminating or reducing the resting discharge on the involved side.

THE FUTURE

Regeneration of Hair Cells

A great amount of study has focused over the past 15 years on the ability of vertebrate hair cells to regenerate. One popular model has been the bird, in which it has been shown that the sound- or drug-damaged sensory epithelium, that is, the basilar papilla, is capable of undergoing significant structural repair.[258–260] Detailed morphologic examination of damaged cochleae showed that immediately after damage, there is a substantial loss of hair cells, an expansion of the surrounding supporting-cell surface area,[261] and a degeneration of the tectorial membrane overlying the region of the lost hair cells.[262] In the days following acoustic overstimulation, the regenerative processes repair many of the damaged structures so that the hair cells are almost completely replaced by regenerated hair cells,[263] cochlear ganglion cells appear to form new synapses with the regenerated hair cells,[264] and a new tectorial membrane is partially regenerated over the lesioned area.[262] In parallel to the structural repair is a considerable recovery of auditory function as measured with physiologic techniques.[265]

Other investigators have documented the capacity of the avian cochlear and vestibular sensory epithelia to repair following aminoglycoside toxicity[266] and even to restore function.[267,268] Using cell proliferation marker procedures (ie, tritiated-thymidine autoradiography or proliferating cell nuclear-antigen immunochemistry), it has been well documented that the supporting cells are the precursors or progenitors of the regenerated cochlear[269] and vestibular[270] hair cells.

The knowledge that noise- or drug-damaged hair cells of birds, reptiles, and even guinea pigs[271] can regenerate has motivated researchers to learn how to stimulate the regrowth of the sensory cells of the human inner ear. In fact, recently it has been shown that brief treatments with forskolin enhance proliferation and regeneration by increasing the amounts of growth factor receptors at the membrane.[272] It is quite likely that, in the future, it will be possible to correct hearing and balance impairments by stimulating the growth of new hair cells.

Molecular Basis of Hearing and Balance and Clinical Implications A current view is that many of the 60% of cases of hearing impairment that have no obvious environmental origins, that is those other than conditions associated with meningitis, perinatal complications, maternal-fetal infections, acoustic trauma, and ototoxic drugs, have a genetic basis.[273] Moreover, it has been estimated that 30% of prelingual deafness causes are syndromic. Several hundred such syndromes, consisting of hearing loss in association with a variety of anatomic anomalies involving, for example, eye, musculoskeletal, renal, nervous system, and pigmentary disorders, have been described.[274] Syndromic hearing loss can have many modes of transmission including maternal inheritance, owing to a mitochondrial mutation. The forms may be conductive, sensorineural, or mixed defects.

The nonsyndromic forms of hearing loss are collectively referred to as DFN for the X-linked forms (1 to 3% incidence), DFNA for the autosomal dominant forms (15% incidence), and DFNB for the autosomal recessive forms (85% incidence). The autosomal recessive forms of hearing loss are often the most severe and account for the vast majority of cases of congenital profound deafness. To date, 20 DFNB, 14 DFNA, and four DFN loci have been mapped to specific chromosomal positions.[275] In addition, within the past 10 years, 10 genes have been cloned and shown to underlie certain nonsyndromic hearing losses. Two of these genes frequently are the cause of nonsyndromic forms of deafness. First, the connexin 26 gene, which contains a single coding exon, was shown to account for up to 50% of all causes of prelingual, autosomal, and recessive hearing loss.[276] Second, a mutation in the mitochon-

drial 12s ribosomal ribonucleic acid gene was shown to underlie an isolated form of sensorineural deafness and deafness induced by aminoglycoside treatment.[277] In addition, the sequencing of the human genome is expected to be complete in 2003, which will allow more deafness-related genes to be discovered. For certain, new technologies, such as the deoxyribonucleic acid microarrays used to study differential gene expression under various normal and pathologic conditions, promise to rapidly advance our knowledge of the diseases of the ear.[278] For a recent review of this active field of research, see Kalatzis and Petit.[275]

So far, however, this research has yet to result in the development of new treatments for deafness. Currently, the treatments available are the amplification of sound, using analog or digital hearing aids, or the artificial, electrical stimulation of the cochlear nerve via a cochlear implant. Efforts to induce regeneration of inner ear sensory cells and methods to introduce replacement genes via viral vectors[279] will soon become the most exciting treatment options leading to a possible reversal of hearing impairment.

REFERENCES

1. Shaw EA, Teranishi R. Sound pressure generated in an external-ear replica and real human ears by a nearby point source. J Acoust Soc Am 1968;44:240–9.
2. Shaw EAG. The external ear. In: Keidel WD, Neff WD, editors. Handbook of sensory physiology. Vol V/1: auditory system, anatomy physiology (ear). New York: Springer-Verlag; 1974. p. 455–90.
3. Pickles JO. An introduction to the physiology of hearing. New York: Academic Press; 1982.
4. von Békésy G. Experiments in hearing. Translated and edited by EG Wever. New York: McGraw-Hill; 1960.
5. Møller AR. An experimental study of the acoustic impedance of the middle ear and its transmisssion properties. Acta Otolaryngol (Stockh) 1965;60:129–49.
6. Tonndorf J, Khanna SM. The role of the tympanic membrane in middle ear transmission. Ann Otol Rhinol Laryngol 1970;79:743–53.
7. Khanna SM, Tonndorf J. Tympanic membrane vibration in cats studied by time-averaged holography. J Acoust Soc Am 1972;51:1904–20.
8. Salomon G, Starr A. Electromyography of middle ear muscles in man during motor activities. Acta Neurol Scand 1963;Suppl (39):161–8.
9. Borg E. On the neuronal organization of the acoustic middle ear reflex. A physiological and anatomical study. Brain Res 1973;49:101–23.
10. Møller AR. The sensitivity of contraction of the tympanic muscles in man. Ann Otol Rhinol Laryngol 1962;71:86–95.
11. Møller A: Acoustic reflex in man. J Acoust Soc Am 1962;34:1524.
12. Djupesland G, Zwislocki JJ. Effect of temporal summation on the human stapedius reflex. Acta Otolaryngol (Stockh) 1971;71:262–5.
13. Jerger J. Clinical experience with impedance audiometry. Arch Otolaryngol 1970;92:311–24.
14. Klockhoff IH. Middle ear muscle reflexes in man. Acta Otolaryngol Suppl (Stockh) 1961;164:1.
15. Reger SN. Effect of middle ear muscle action on certain psychophysical measurements. Ann Otol Rhinol Laryngol 1960;69:1179–98.
16. Smith HD. Audiometric effects of voluntary contraction of the tensor tympani muscles. Arch Otolaryngol 1943;38:369–72.
17. Holst HE, et al. Ear drum movements following stimulation of the middle ear muscles. Acta Otolaryngol Suppl (Stockh) 1963;182:73–84.
18. Keefe DH, Bulen JC, Arehart KH, Burns EM. Ear-canal impedance and reflection coefficient in human infants and adults. J Acoust Soc Am 1993;94:2617–38.
19. Goode RL, Ball G, Nishihara S, Nakamura K. Laser Doppler vibrometer (LDV)—a new clinical tool for the otologist. Am J Otol 1996;17:813–22.
20. Zakrisson JE, Borg E, Diamant H, Miller AR. Auditory fatigue in patients with stapedius muscle paralysis. Acta Otolaryngol (Stockh) 1975;79:228–32.
21. Simmons FB. Perceptual theories of middle ear muscle function. Ann Otol Rhinol Laryngol 1964; 73:724–40.
22. Harford ER, Jerger JF. Effect of loudness recruitment on delayed speech feedback. J Speech Hear Res 1959; 2:361–8.
23. Borg E, Zakrisson JE. The activity of the stapedius muscle in man during vocalization. Acta Otolaryngol (Stockh) 1973;79:325–33.
24. Jahnke K. The fine structure of freeze-fractured intercellular junctions in the guinea pig inner ear. Acta Otolaryngol Suppl (Stockh) 1975;336:1–40.
25. Engstrom H. The cortilymph, the third lymph of the inner ear. Acta Morphol Neer Scand 1960;3:195–204.
26. Nadol JB. Intercellular fluid pathways in the organ of Corti of cat and man. Ann Otol Rhinol Laryngol 1979;88:2–11.

27. Steele KP. Tectorial membrane. In: Altschuler RA, Hoffman DW, Bobbin RP, editors. Neurobiology of hearing: the cochlea. New York: Raven Press; 1986.
28. Smith CA. Structure of the cochlear duct. In: Naunton RF, Fernandez C, editors. Evoked electrical activity in the auditory nervous system. New York: Academic Press; 1978. p. 3–19.
29. Lim DJ. Fine morphology of the tectorial membrane: its relationship to the organ. Corti Arch Otolaryngol 1972;96:199–215.
30. Flock A. Physiological properties of sensory hairs in the ear. In: Evans EF, Wilson JP, editors. Psychophysics and physiology of hearing. New York: Academic Press; 1977. p. 15–25
31. Nielsen DW, Slepecky N. Stereocilia. In: Altschuler RA, Hoffman DW, Bobbin RP, editors. Neurobiology of hearing: the cochlea. New York: Raven Press; 1986. p. 23–46.
32. Flock A, Cheung HC. Actin filaments in sensory hairs of inner ear receptor cells. J Cell Biol 1977;75: 339–43.
33. Flock A, Bretscher A, Weber K. Immunohistochemical localization of several cytoskeletal proteins in inner ear sensory and supporting cells. Hear Res 1982;7:75–89.
34. Hudspeth AJ, Choe Y, Mehta AD, Martin P. Putting ion channels to work: mechanoelectrical transduction, adaptation, and amplification by hair cells. Proc Natl Acad Sci U S A 2000;97:11765–72.
35. Slepecky N, Chamberlain SC. Distribution and polarity of actin in the sensory hair cells of the chinchilla cochlea. Cell Tissue Res 1982;224:15–24.
36. Flock A, Flock B, Murray E. Studies on the sensory hairs of receptor cells in the inner ear. Acta Otolaryngol (Stockh) 1977;83:85–91.
37. Steele CR. A possibility for sub-tectorial membrane fluid motion. In: Møller AR, editor. Basic mechanisms in hearing. New York: Academic Press; 1973.
38. Dallos P, Billone MC, Durrant JD, et al. Cochlear inner and outer hair cells: functional differences. Science 1972;177:356–8.
39. Russell IJ, Sellick PM. The responses of hair cells to low frequency tones and their relationship to the extracellular receptor potentials and sound pressure level in the guinea pig cochlea. In: Syka J, Aitkin L, editors. Neuronal mechanisms of hearing. New York: Plenum Press; 1981.
40. von Békésy G. Shearing microphonics produced by vibrations near the inner and outer hair cells. J Acoust Soc Am 1953;25:786.
41. Tasaki I, Davis H, Eldredge DH. Exploration of cochlear potentials in guinea pig with a microelectrode. J Acoust Soc Am 1954;26:765–73.
42. Tasaki I, Spyropoulos CS. Stria vascularis as source of endocochlear potential. J Neurophysiol 1959;22: 149–55.
43. Konishi T, Kelsey E. Effect of cyanide on cochlear potentials. Acta Otolaryngol (Stockh) 1968;65:381–90.
44. Tasaki I, Fernandez C. Modification of cochlear microphonics and action potentials by KCl solution and by direct currents. J Neurophysiol 1952;15: 497–512.
45. Davis H, Derbyshire AJ, Lurie JH, Saul LJ. The electric response of the cochlea. Am J Physiol 1934;107:311–32.
46. Tasaki I, Davis H, Legouix JP. The space-time pattern of the cochlear microphonic (guinea pig) as recorded by differential electrodes. J Acoust Soc Am 1952;24:502–19.
47. Davis H, Deatheridge BH, Eldredge DH, Smith CA. Summating potentials of the cochlea. Am J Physiol 1958;195:251–61.
48. Konishi T, Yasuno T. Summating potential of the cochlea in the guinea pig. J Acoust Soc Am 1963; 35:1448–52.
49. Hudspeth AJ. Extracellular current flow and the site of transduction by vertebrate hair cells. J Neurosci 1982;2:1–10.
50. Dallos P, Cheatham MA. Production of cochlear potentials by inner and outer hair cells. J Acoust Soc Am 1976;60:510–2.
51. Durrant JD, Wang J, Ding DL, Salvi RJ. Are inner or outer hair cells the source of summating potentials recorded from the round window? J Acoust Soc Am 1998;104:370–7.
52. Davis H. A model for transducer action in the cochlea. Cold Spring Harb Symp Quant Biol 1965; 30:181–90.
53. Wever EG, Bray CW. Action currents in the auditory nerve in response to acoustical stimulation. Proc Natl Acad Sci U S A 1930;16:344–500.
54. Davis H, Tasaki I, Goldstein R. The peripheral origin of activity with reference to the ear. Cold Spring Harb Symp Quant Biol 1952;17:143–54.
55. Brownell WE. Cochlear transduction: an integrative model and review. Hear Res 1982;6:335–60.
56. Hudspeth AJ. The hair cells of the inner ear. Sci Am 1983;248:54–64.
57. LeBlanc CS, Fallon M, Parker MS, et al. Phosphorothioate oligodeoxynucleotides can selectively alter neuronal activity in the cochlea. Hear Res 1999;135:105–12.

58. Puel J-L. Chemical synaptic transmission in the cochlea. Prog Neurobiol 1995;47:449–76.
59. Werman R. Criteria for identification of a central nervous system transmitter. Comp Biochem Physiol 1966;18:745–66.
60. Kuriyama H, Jenkins O, Altschuler RA. Immunocytochemical localization of AMPA selective glutamate receptor subunits in the rat cochlea. Hear Res 1994; 80:233–40.
61. Jager W, Goiny M, Herrera-Marschitz M, et al. Sound-evoked efflux of excitatory amino acids in the guinea-pig cochlear in vitro. Exp Brain Res 1998; 121:425–32.
62. Bobbin RP, Bledsoe SC, Jenison GL. Neurotransmitters of the cochlea and lateral line organ. In: Berlin CI, editor. Hearing science: recent advances. San Diego: College-Hill Press; 1984.
63. Schwartz IR, Ryan AF. Uptake of amino acids in the gerbil cochlea. In: Altschuler RA, Hoffman DW, Bobbin RP, editors. Neurobiology of hearing: the cochlea. New York: Raven Press; 1986.
64. Fex J, Altschuler RA. Neurotransmitter-related immunocytochemistry of the organ of Corti. Hear Res 1986;22:249–63.
65. Plinkert PK, Mohler H, Zenner HP. A subpopulation of outer hair cells possessing GABA receptors with tonotopic organization. Arch Otorhinolaryngol 1989; 246:417–22.
66. Gitter AH, Zenner HP. Gamma-aminobutyric acid receptor activation of outer hair cells in the guinea pig cochlea. Eur Arch Otorhinolaryngol 1992;249:62–5.
67. Eybalin M. Neurotransmitters and neuromodulators of the mammalian cochlea. Physiol Rev 1993;73: 309–73.
68. Frisina RD. Subcortical neural coding mechanisms for auditory temporal processing. Hear Res 2001; 158:1–27.
69. LePage EL, Johnstone BM. Nonlinear mechanical behaviour of the basilar membrane in the basal turn of the guinea pig cochlea. Hear Res 1980;2:183–9.
70. Liberman MC, Kiang NYS. Acoustic trauma in cats. Cochlear pathology and auditory-nerve activity. Acta Otolaryngol Suppl (Stockh) 1978;358:1–63.
71. Ruggero MA, Rich NC. Application of a commercially manufactured Doppler-shift laser velocimeter to the measurement of basilar-membrane vibration. Hear Res 1991;51:215–30.
72. Møller AR. Studies of the damped oscillatory response of the auditory frequency analyzer. Acta Physiol Scand 1970;78:299–314.
73. Evans EF. Auditory frequency selectivity and the cochlear nerve. In: Zwicker E, Terbardt E, editors. Facts and models in hearing. New York: Springer-Verlag; 1974. p. 118–29.
74. Kiang NYS, Moxon EC, Levine RA. Auditory nerve activity in cats with normal and abnormal cochleas. In: Wolstenhome GEW, Knight J, editors. Ciba Symposium on Sensorineural Hearing Loss. London: Churchill Press; 1970. p. 241–68.
75. Robertson D. Cochlear neurons: frequency selectivity altered by perilymph removal. Science 1974;186: 153–5.
76. Cody AR, Johnstone BM. Single auditory neuron response during acute acoustic trauma. Hear Res 1980;3:3–16.
77. Johnstone BM, Boyle AJ. Basilar membrane vibration examined with the Mössbauer technique. Science 1967;158:389–90.
78. Rhode WS. Measurement of vibration of the basilar membrane in the squirrel monkey. Ann Otol Rhinol Laryngol 1974;83:619–25.
79. Kim DO, Neely ST, Molnar CE, Matthews JW. An active cochlear model with negative damping in the partition: comparison with Rhode's ante- and postmortem observations. In: van den Brink G, Bilsen FA, editors. Psychophysical, physiological, and behavioral studies in hearing. Delft, Netherlands: Delft University Press; 1980. p. 7–14.
80. Kemp DT. Stimulated acoustic emissions from within the human auditory system. J Acoust Soc Am 1978;64:1386–91.
81. Wit HP, Ritsma RJ. Evoked acoustical responses from the human ear: some experimental results. Hear Res 1980;2:253–61.
82. Mountain DC. Changes in endolymphatic potential and crossed olivocochlear bundle stimulation alter cochlear mechanics. Science 1980;210:71–2.
83. Engdahl B, Kemp DT. The effect of noise exposure on the details of the distortion-product otoacoustic emissions in humans. J Acoust Soc Am 1996;99:1573–87.
84. Allen GC, Tiu C, Koike K, et al. Transient-evoked otoacoustic emissions in children after cisplatin chemotherapy. Otolaryngol Head Neck Surg 1998; 118:584–8.
85. Brownell WE, Bader CR, Bertrand D, de Ribaupierre Y. Evoked mechanical responses of isolated cochlear outer hair cells. Science 1985;227:194–6.
86. Zenner HP, Zimmermann U, Schmitt U. Reversible contraction of isolated mammalian cochlear hair cells. Hear Res 1985;18:127–33.

87. Flock A. Hair cells, receptors with a motor capacity? In: Klinke R, Hartman R, editors. Hearing, physiological bases and psychophysics. New York: Springer-Verlag; 1983. p. 2–7.
88. Dallos P, Evans BN, Hallworth R. Nature of the motor element in electrokinetic shape changes of cochlear outer hair cells. Nature 1991;350:155–7.
89. Brownell WE. Outer hair cell electromotility and otoacoustic emissions. Ear Hear 1990;11:82–92.
90. Zheng J, Shen W, He DZ, et al. Prestin is the motor protein of cochlear outer hair cells. Nature 2000; 405:149–55.
91. Oliver D, He DZ, Klocker N, et al. Intracellular anions as the voltage sensor of prestin, the outer hair cell motor protein. Science 2001;292:2340–3.
92. Rubel EW. Ontogeny of structures and function in the vertebrate auditory system. In: Jacobson M, editor. Handbook of sensory physiology. Vol IX: development of sensory systems. New York: Springer-Verlag; 1978. p. 135–237.
93. Kim DO. Functional roles of the inner- and outer-hair-cell subsystems in the cochlea and brainstem. In: Berlin CI, editor. Hearing science: recent advances. San Diego: College-Hill Press; 1984. p. 241–62.
94. Wever EG. Theory of hearing. New York: John Wiley & Sons Press; 1949.
95. Warr WB, Guinan JJ, White JS. Organization of the efferent fibers: the lateral and medial olivocochlear systems. In: Altschuler RA, Hoffman DW, Bobbin RP, editors. Neurobiology of hearing: the cochlea. New York: Raven Press; 1986. p. 333–48.
96. Spangler KM, Warr WB. The descending auditory system. In: Altschuler RA, Bobbin RP, Clopton BM, Hoffman DW, editors. Neurobiology of hearing: the central auditory system. New York: Raven Press; 1991.
97. Evans EF, Harrison RV. Correlation between cochlear outer hair cell damage and deterioration of cochlear nerve tuning properties in the guinea pig. J Physiol 1976;256:43–4.
98. Brown MC, Nuttal AL, Masta RI. Intracellular recordings from cochlear inner hair cells: effects of stimulation of the crossed olivocochlear efferents. Science 1983;222:69–72.
99. Gifford ML, Guinan JJ. Effects of crossed-olivocochlear-bundle stimulation on cat auditory nerve fiber responses to tones. J Acoust Soc Am 1983; 74:115–23.
100. Winslow RL, Sachs MB. Effect of electrical stimulation of the crossed olivocochlear bundle on auditory nerve response to tones in noise. J Neurophysiol 1987;57:1002–21.
101. Collet L. Use of otoacoustic emissions to explore the medial olivocochlear system in humans. Br J Audiol 1993;27:155–9.
102. Berlin CI, Hood LJ, Hurley A, Wen H. Contralateral suppression of otoacoustic emissions: an index of the function of the medial olivocochlear system. Otolaryngol Head Neck Surg 1994;110:3–21.
103. Penner MJ. An estimate of the prevalence of tinnitus caused by spontaneous otoacoustic emissions. Arch Otolaryngol Head Neck Surg 1990;116:418–23.
104. Fetterman BL. Distortion-product otoacoustic emissions and cochlear microphonics: relationships in patients with and without endolymphatic hydrops. Laryngoscope 2001;111:946–54.
105. Schweinfurth JM, Cacace AT, Parnes SM. Clinical applications of otoacoustic emissions in sudden hearing loss. Laryngoscope 1997;107:1457–63.
106. Evans EF. The sharpening of cochlear frequency selectivity in the normal and abnormal cochlea. Audiology 1975;14:419–42.
107. Evans EF. The frequency response and other properties of single fibers in the guinea-pig cochlear nerve. J Physiol 1972;226:263–87.
108. Brugge JF, Anderson DJ, Hind JE, Rose JE. Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds. J Neurophysiol 1969;32:386–401.
109. Rose JE, Brugge JF, Anderson DJ, Hind JE. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. J Neurophysiol 1967;30:769–93.
110. Kiang NY. Processing of speech by the auditory nervous system. J Acoust Soc Am 1980;68:830–5.
111. Sachs MB, Young ED. Encoding of steady state vowels in the auditory nerve: representation in terms of discharge rate. J Acoust Soc Am 1979;66: 470–9.
112. Kiang NYS, Watanabe T, Thomas EC, Clark LF. Discharge patterns of single fibers in the cat's auditory nerve. Cambridge (MA): Massachusetts Institute of Technology Press; 1965. MIT research monograph no. 35.
113. Kiang NYS. A survey of recent developments in the study of auditory physiology. Ann Otol Rhinol Laryngol 1968;77:656–75.
114. Evans EF. The dynamic range problem: place and time coding at the level of the cochlear nerve and nucleus. In: Syka J, Aitkin L, editors. Neuronal

mechanisms of hearing. New York: Plenum Press; 1981. p. 69–85.
115. Evans EF, Palmer AR. Responses of units in the cochlear nerve and nucleus of the cat to signals in the presence of bandstop noise. J Physiol 1975; 252:60P–62P.
116. Palmer AR, Evans EF. On the peripheral coding of the level of individual frequency components of complex sound levels. Exp Brain Res 1979; Suppl 2:19–20.
117. Young ED, Sachs MB. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. J Acoust Soc Am 1979;66:1381–403.
118. Liberman MC. Auditory-nerve response from cats raised in low-noise chamber. J Acoust Soc Am 1978;63:442–55.
119. Kawase T, Liberman MC. Spatial organization of the auditory nerve according to spontaneous discharge rate. J Comp Neurol 1992;319:312–8.
120. Viemeister NF. Intensity coding and the dynamic range problem. Hear Res 1988;34:267–74.
121. Rasmussen GL. Anatomic relationships of the ascending and descending auditory systems. In: Fields WS, Alford BR, editors. Neurological aspects of auditory and vestibular disorders. Springfield (MA): Charles C. Thomas Press; 1964. p. 1–19.
122. Warr WB. The olivocochlear bundle: its origins and terminations in the cat. In: Naunton RF, Fernandez C, editors. Evoked electrical activity in the auditory nervous system. New York: Academic Press; 1978. p. 43–63.
123. Iurato S, Smith CA, Eldredge DH, et al. Distribution of the crossed olivocochlear bundle in the chinchilla's cochlea. J Comp Neurol 1978;182:57–76.
124. Spoendlin HH, Gacek RR. Electronmicroscopic study of the efferent and afferent innervation of the organ of Corti in the cat. Ann Otol Rhinol Laryngol 1963;72:660.
125. Liberman MC, Brown MC. Physiology and anatomy of single olivocochlear neurons in the cat. Hear Res 1986;24:17–36.
126. Liberman MC. Response properties of cochlear efferent neurons: monaural vs. binaural stimulation and the effects of noise. J Neurophysiol 1988; 60:1779–98.
127. Le Prell CG, Bledsoe SC, Bobbin RP, Puel J-L. Neurotransmission in the inner ear: functional and molecular analyses. In: Jahn AF, Santos-Sacchi J, editors. Physiology of the ear. 2nd ed. San Diego: Singular/Thompson Learning Press; 2001. p. 575–611.
128. Elgoyhen AB, Johson DS, Boulter J, et al. Alpha 9: an acetylcholine receptor with novel pharmacological properties expressed in rat cochlear hair cells. Cell 1994;79:705–15.
129. Desmedt JE. Physiological studies of the efferent recurrent auditory system. In: Keidel WD, Neff WD, editors. Handbook of sensory physiology. Vol V/2: physiological (CNS), behavioral studies, psychoacoustics. New York: Springer-Verlag; 1974. p. 219–46.
130. Vetter DE, Liberman MC, Mann J, et al. Role of alpha 9 nicotinic ACh receptor subunits in the development and function of cochlear efferent innervation. Neuron 1999;23:93–103.
131. Fuchs PA. Synaptic transmission at vertebrate hair cells. Curr Opin Neurobiol 1996;6:514–9.
132. Ishii D, Balogh K. Distribution of efferent nerve endings in the organ of Corti. Their graphic reconstruction in cochleae by localization of acetylcholinesterase activity. Acta Otolaryngol (Stockh) 1968;66:282–8.
133. Igarashi M, Cranford JL, Nakai Y, Alford BR. Behavioral auditory function after transection of crossed olivo-cochlear bundle in the cat. IV. Study of pure-tone frequency discrimination. Acta Otolaryngol (Stockh) 1979;87:79–83.
134. Moulin A, Collet L, Duclaux R. Contralateral auditory stimulation alters acoustic distortion products in humans. Hear Res 1993;65:193–210.
135. Davis H. An active process in cochlear mechanics. Hear Res 1983;9:79–90.
136. Geisler CD. Hypothesis on the function of the crossed olivocochlear bundle. J Acoust Soc Am 1974;56:1908–9.
137. Cody AR, Johnstone BM. Temporary threshold shift modified by binaural acoustic stimulation. Hear Res 1982;6:199–205.
138. Canlon B. The effect of acoustic trauma on the tectorial membrane, stereocilia, and hearing sensitivity: possible mechanisms underlying damage, recovery, and protection. Scand Audiol Suppl 1988;27:1–45.
139. Franklin DJ, Lonsbury-Martin BL, Stagner BB, Martin GK. Altered susceptibility of $2f_1$-f_2 acoustic-distortion products to the effects of repeated noise exposure in rabbits. Hear Res 1991;53:185–208.
140. Dagli S, Canlon B. The effect of repeated daily noise exposure on sound-conditioned and unconditioned guinea pigs. Hear Res 1997;104:39–46.

141. Brown MC, Kujawa SG, Liberman MC. Single olivocochlear neurons in the guinea pig. II. Response plasticity due to noise conditioning. J Neurophysiol 1998;79:3088–97.
142. Canlon B, Fransson A, Viberg A. Medial olivocochlear efferent terminals are protected by sound conditioning. Brain Res 1999;850:253–60.
143. LePage EL. Frequency-dependent self-induced bias of the basilar membrane and its potential for controlling sensitivity and tuning in the mammalian cochlea. J Acoust Soc Am 1987;82:139–54.
144. Heimer L, Robards MJ. Neuroanatomical tract-tracing methods. New York: Plenum Press; 1981.
145. Helfert RH, Snead CR, Alschuler RA. The ascending auditory pathways. In: Altschuler RA, Bobbin RP, Clopton BM, Hoffman DW, editors. Neurobiology of hearing: the central auditory system. New York: Raven Press; 1991. p. 1–25.
146. Moore JK. Cochlear nuclei: relationship to the auditory nerve. In: Altschuler RA, Hoffman DW, Bobbin RP, editors. Neurobiology of hearing: the cochlea. New York: Raven Press; 1986. p. 283–301.
147. Phillips DP. Introduction to the central auditory nervous system. In: Jahn AF, Santos-Sacchi J, editors. Physiology of the ear. 2nd ed. San Diego: Singular/Thompson Learning Press; 2001. p. 613–38.
148. Lorenté de No R. Anatomy of the eighth nerve: I. The central projection of the nerve endings of the internal ear. Laryngoscope 1933;43:1–38.
149. Brawer JR, Morest DK, Kane EC. The neuronal architecture of the cochlear nucleus of the cat. J Comp Neurol 1974;155:251–300.
150. Young ED. Response characteristics of neurons of the cochlear nucleus of the cat. In: Berlin CI, editor. Hearing science: recent advances. San Diego: College-Hill Press; 1984. p. 423–60.
151. Goldberg JM, Brown PB. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. J Neurophysiol 1969;32:613–36.
152. Thompson GC. Structure and function of the central auditory system. Semin Hear 1983;4:81.
153. Brugge JF, Anderson DJ, Aitkin LM. Responses of neurons in the dorsal nucleus of the lateral lemniscus of cat to binaural tonal stimulation. J Neurophysiol 1970;33:441–58.
154. Schofield BR, Cant NB. Ventral nucleus of the lateral lemniscus in guinea pigs: cytoarchitecture and inputs from the cochlear nucleus. J Comp Neurol 1997;379:363–85.
155. Morest DK. The neuronal architecture of the medial geniculate body of the cat. J Anat 1964; 98:611–30.
156. David E, et al. Decoding processes in the auditory system and human speech analysis. In: Evans EF, Wilson JP, editors. Psychophysics and physiology of hearing. New York: Academic Press; 1977. p. 509–16.
157. Keidel WD. Information processing in the higher parts of the auditory pathway. In: Zwicker E, Terhardt E, editors. Facts and models in hearing. New York: Springer-Verlag; 1974. p. 216–26.
158. Rose JE. The cellular structure of the auditory region of the cat. J Comp Neurol 1949;91:409–39.
159. Winer JA, Diamond IT, Raczkowski D. Subdivisions of the auditory cortex of the cat: the retrograde transport of horseradish peroxidase to the medial geniculate body and posterior thalamic nuclei. J Comp Neurol 1977;176:387–417.
160. Coats AC. On electrocochleographic electrode design. J Acoust Soc Am 1974;56:708–11.
161. Picton TW, Hillyard SA, Krausz HI, Galabos R. Human auditory evoked potentials. I. Evaluation of components. Electroencephalogr Clin Neurophysiol 1974;36:179–90.
162. Haapaniemi J, Laurikainen E, Johansson R, Karjalainen S. Transtympanic versus tympanic membrane electrocochleography in examining cochleovestibular disorders. Acta Otolaryngol Suppl (Stockh) 2000;543:127–9.
163. Elberling C. Simulation of cochlear action potentials recorded from the ear canal in man. In: Ruben RJ, Elberling C, Salomon G, editors. Electrocochleography. Baltimore: University Park Press; 1976.
164. Coats AC. Electrocochleography: recording techniques and clinical applications. Semin Hear 1986; 7:247.
165. Ferraro JA, Tibbils RP. SP/AP area ratio in the diagnosis of Meniere's disease. Am J Audiol 1999;8:21–8.
166. Wuyts FL, Van de Heyning PH, Van Spaendonck M, et al. Rate influences on tone burst summating potential amplitude in electrocochleography: clinical (a) and experimental (b) data. Hear Res 2001;152:1–9.
167. Moffat DA, Gibson WP, Ramsden RT, et al. Transtympanic electrocochleography during glycerol dehydration. Acta Otolaryngol (Stockh) 1978; 85:158–66.
168. Møller AR. Physiology of the ascending auditory pathway with special reference to the auditory brainstem response (ABR). In: Pinheiro ML,

Musiek FE, editors. Assessment of central auditory dysfunction: foundations and clinical correlates. Baltimore: Williams and Wilkins; 1985.

169. Wada S, Starr A. Generation of auditory brainstem responses. III. Effects of lesions of the superior olive, lateral lemniscus, and inferior colliculus on the ABR in guinea pig. Electroencephalogr Clin Neurophysiol 1983;56:352–66.
170. Musiek FE, Gollegly KM, Kibbe KS, Verkest SB. Current concepts on the use of ABR and auditory psychophysical tests in the evaluation of brain stem lesions. Am J Otol 1988;Suppl 9:25–35.
171. Hosford-Dunn H. Auditory function tests. In: Cummings CW, et al, editors. Otolaryngology—head and neck surgery. St. Louis: CV Mosby; 1986.
172. Kaga K, Hink RF, Shimada Y, Suzuki J. Evidence for a primary cortical origin of a middle latency auditory evoked potential in cats. Electroencephalogr Clin Neurophysiol 1980;50:254–66.
173. McGee TJ, Özdamar Ö, Kraus N. Auditory middle latency responses in the guinea pig. Am J Otolaryngol 1983;4:116–22.
174. Scherg M, Von Cramer D. Topographical analysis of auditory evoked potentials: derivation of components. In: Nodar R, Barber C, editors. Evoked potentials. Stoneham (MA): Butterworth Press; 1984.
175. Fifer RC, Sierra-Irizzary B. Clinical applications of the auditory middle latency response. Am J Otol 1988;Suppl 9:47–56.
176. Kileny PR, Kemink JL. Electrically evoked middle-latency auditory potentials in cochlear implant candidates. Arch Otolaryngol Head Neck Surg 1987; 113:1072–7.
177. Duffy FH. Auditory brain mapping. In: Jahn AF, Santos-Sacchi J, editors. Physiology of the ear. New York: Raven Press; 1988.
178. Baran JA, Long RR, Musiek FE, Ommaya A. Topographic mapping of brain electrical activity in the assessment of central auditory nervous system pathology. Am J Otol 1988;Suppl 9:72–6.
179. Hari R, Lounasmaa OV. Recording and interpretation of cerebral magnetic fields. Science 1989;244: 432–6.
180. Hari R, Pelizzone M, Makela JP, et al. Neuromagnetic responses from a deaf subject to stimuli presented through a multi-channel cochlear prosthesis. Ear Hear 1988;9:148–52.
181. Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. Cereb Cortex 2001; 11:946–53.
182. Shaywitz BA, Shaywitz SE, Pugh KR, et al. Sex differences in the functional organization of the brain for language. Nature 1995;373:607–9.
183. Melcher JR, Sigalovsky IS, Guinan JJ, Levine RA. Lateralized tinnitus studied with functional magnetic resonance imaging: abnormal inferior colliculus activation. J Neurophysiol 2000;83:1058–72.
184. Cacace AT, Tasciyan T, Cousins JP. Principles of functional magnetic resonance imaging: application to auditory neuroscience. J Am Acad Audiol 2000;11:239–72.
185. Filipo R, Delfini R, Fabiani M, et al. Role of transient-evoked otoacoustic emissions for hearing preservation in acoustic neuroma surgery. Am J Audiol 1997;18:746–9.
186. Kemp DT, Ryan S. The use of transient evoked otoacoustic emissions in neonatal hearing screening programs. Semin Hear 1993;14:30–45.
187. Balkany TJ, Telischi FF, Lonsbury-Martin BL, Martin GK. Otoacoustic emissions in clinical practice. Am J Otol 1994;Suppl 15:29–38.
188. Martin GK, Ohlms LA, Franklin DJ, et al. Distortion-product otoacoustic emissions in humans: III. Influence of sensorineural hearing loss. Ann Otol Rhinol Laryngol 1990;Suppl 147:30–42.
189. Hotz MA, Probst R, Harris FP, Hauser R. Monitoring the effects of noise exposure using transiently evoked otoacoustic emissions. Acta Otolaryngol (Stockh) 1993;113:478–82.
190. Fiore C, Cagini C, Menduno P, et al. Evoked otoacoustic emissions behaviour in retinitis pigmentosa. Doc Ophthalmol 1994;87:167–76.
191. Zorowka PG, Schmitt HJ, Gutjahr P. Evoked otoacoustic emissions and pure tone threshold audiometry in patients receiving cisplatinum therapy. Int J Pediatr Otorhinolaryngol 1993;25:73–80.
192. Fortnum H, Farnsworth A, Davis A. The feasibility of evoked otoacoustic emissions as an in-patient hearing check after meningitis. Br J Audiol 1993; 27:227–31.
193. Stover L, Norton SJ. The effects of aging on otoacoustic emissions. J Acoust Soc Am 1993;94: 2670–81.
194. Hall JW, Baer JE, Chase PA, Schwaber MK. Clinical application of otoacoustic emissions: what do we know about factors influencing measurement and analysis? Otolaryngol Head Neck Surg 1994;110: 22–38.
195. Lonsbury-Martin BL, Martin GK, Balkany T. Clinical applications of otoacoustic emissions. In: Lucente

FE, editor. Highlights of the instructional courses. Alexandria (VA): American Academy of Otolaryngology—Head and Neck Surgery; 1994. p. 343–55.
196. Igarashi M, Miyata L, Alford BR. Utricular ablation and dysequilibrium in squirrel monkeys. Acta Otolaryngol (Stockh)1972;74:66–72.
197. Fernandez C, Goldberg JM. Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. I. Response to static tilts and to long-duration centrifugal force. J Neurophysiol 1976;39:970–84.
198. Fernandez C, Goldberg JM. Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. II. Directional selectivity and force-response relations. J Neurophysiol 1976;39:985–95.
199. Fernandez C, Goldberg JM. Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. III. Response dynamics. J Neurophysiol 1976;39:996–1008.
200. Wendt GR. Vestibular functions. In: Stevens SS, editor. Handbook of experimental psychology. New York: John Wiley & Sons; 1951.
201. Cogan DG. Some objective and subjective observations on the vestibulo-ocular system. Am J Ophthalmol 1958;45:74.
202. de Beer GR. Presidential address: how animals hold their heads. Proc Linn Soc Lond 1947;159:125.
203. Juhn SK. Biochemistry of the labyrinth: a manual. Rochester (MN): American Academy of Ophthalmology and Otolaryngology; 1973.
204. Dohlman GF. Critical review of the concept of cupula function. Acta Otolaryngol Suppl (Stockh) 1980;376:1–30.
205. Carlstrom D, et al. Electron microscopic and x-ray diffraction studies of statoconia. Laryngoscope 1953;63:1052–7.
206. Wersall J, et al. Ultrastructure of the vestibular end organs. In: De Reuck AVS, Knight J, editors. Myotatic, kinesthetic and vestibular mechanisms. Boston (MA): Little, Brown Press; 1968. p. 105–16.
207. Spoendlin H. Ultrastructure of the vestibular sense organ. In: Wolfson RJ, editor. The vestibular system and its diseases. Philadelphia (PA): University of Pennsylvania Press; 1966.
208. Anniko M, Thornell LE, Virtanen I. Cytoskeletal organization of the human inner ear. IV. Expression of actin in vestibular organs. Acta Otolaryngol Suppl (Stockh) 1987;437:5–76.
209. Zenner HP, Zimmermann U. Motile responses of vestibular hair cells following caloric, electrical or chemical stimuli. Acta Otolaryngol (Stockh) 1991; 111:291–7.
210. Wersall J, Flock A, Lundquist PG. Structural basis for directional sensitivity in cochlear and vestibular sensory receptors. Cold Spring Harb Symp Quant Biol 1965;30:115–32.
211. Engstrom H, Bergstrom B, Ades HW. Macula utriculi and macula saculi in the squirrel monkey. Acta Otolaryngol Suppl (Stockh) 1972;301:75–126.
212. Tilney LG, DeRosier DJ, Mulroy MJ. The organization of actin filaments in the stereocilia of cochlear hair cells. J Cell Biol 1980;86:244–59.
213. Assad JA, Shepherd GM, Corey DP. Tip link integrity and mechanical transduction in vertebrate hair cells. Neuron 1991;7:985–94.
214. Spoendlin H. Receptor ultrastructure. In: De Reuck AVS, Knight J, editors. Symposium on hearing mechanisms in vertebrates. Boston (MA): Little Brown Press; 1968. p. 89–119.
215. Hillman DE, McLaren JW. Displacement configuration of semiciruclar canal cupulae. Neuroscience 1979;4:1989–2000.
216. McLaren JW, Hillman DE. Displacement of semicircular canal cupula during sinusoidal rotation. Neuroscience 1979;4:2001-8.
217. Oman CM, Young LR. Physiological range of pressure difference and cupula deflections in the human semicircular canal: theoretical considerations. Prog Brain Res 1972;37:529–39.
218. Steinhausen W. The cupula. Z Hals Nas Ohrenh 1931;29:211.
219. Goldberg JM, Ferenadez C. Physiology of peripheral neurons innervating semicircular canal of the squirrel monkey. I. Resting discharge and response to constant angular accelerations. J Neurophysiol 1971;34:635–60.
220. Goldberg JM, Fernandez C. Vestibular mechanisms. Ann Rev Physiol 1975;37:129–62.
221. Keller EL. Behavior of horizontal semicircular canal afferents in alert monkey during vestibular and optokinetic stimulation. Exp Brain Res 1976; 24:459–71.
222. Fernandez C, Goldberg JM. Physiology of peripheral neurons in innervating semicircular canals of the squirrel monkey. II. Response to sinusoidal stimulation and dynamics of peripheral vestibular system. J Neurophysiol 1971;34:661–75.
223. Brodal A. Neurological anatomy in relation to clinical medicine. New York: Oxford University Press; 1969.

224. Baker R. The nucleus prepositus hypoglossi. In: Brooks BA, Bajandas FJ, editors. Eye movements. New York: Plenum Press; 1977.
225. Keller EL. The behavior of eye movement motoneurons in the alert monkey. Bibl Ophthalmol 1972;82: 7–16.
226. Cannon SC, Robinson DA. Loss of the neural integrator of the oculomotor system from brainstem lesions in monkey. J Neurophysiol 1987;57:1383–409.
227. Wilson VJ, Wylie RM, Marco LA. Organization of the medial vestibular nucleus. J Neurophysiol 1968; 31:166–75.
228. Miller EF, Graybiel A. A comparison of ocular counter rolling movements between normal persons and deaf subjects with bilateral labyrinthine defects. Ann Otol Rhinol Laryngol 1963;72:885–93.
229. Goldberg JM. Afferent diversity and the organization of central vestibular pathways. Exp Brain Res 2000;130:277–97.
230. Kotchabhakdi N, Walberg F. Primary vestibular afferent projections to the cerebellum as demonstrated by retrograde axonal transport of horseradish peroxidase. Brain Res 1978;142:142–6.
231. Kotchabhakdi N, Walberg F. Cerebellar afferent projections from the vestibular nuclei in the cat: an experimental study with the method of retrograde axonal transport of horseradish peroxidase. Exp Brain Res 1978;31:591–604.
232. Fredrickson JM, Frigge U, Scheid P, Kornhuber HH. Vestibular nerve projection to the cerebral cortex of the rhesus monkey. Exp Brain Res 1966;2:318–27.
233. Deecke L, Schwartz DW, Fredrickson JM. The vestibular thalamus in the rhesus monkey. Adv Otorhinolaryngol 1973;19:210–9.
234. Wilson VJ, Melvill-Jones G. Mammalian vestibular physiology. New York: Plenum Press; 1979.
235. Precht W, Shimazu H, Markham CH. A mechanism of central compensation of vestibular function following hemilabyrinthectomy. J Neurophysiol 1966; 29:996–1010.
236. Igarashi M, Levy JK, O-Uchi T, Reschke MF. Further study of physical exercise and locomotor balance compensation after unilateral labyrinthectomy in squirrel monkeys. Acta Otolaryngol (Stockh) 1981;92:101–5.
237. Fetter M, Zee DS. Recovery from unilateral labyrinthectomy in rhesus monkey. J Neurophysiol 1988;59:370–93.
238. Brandt T. Management of vestibular disorders. J Neurol 2000;247:491–9.
239. Bechterew W. Ergebnisse der Durchschneidung des N. acusticus, nebst Eroerterung der Bedeutung der semicirculaeren Canaele fuer das Koerpergleichgewicht. Pfluegers Arch Ges Physiol 1883;30:312–47.
240. Darlington CL, Smith PF. Molecular mechanisms of recovery from vestibular damage in mammals: recent advances. Prog Neurobiol 2000;62:313–25.
241. Highstein SM. Abducens to medical rectus pathway in the MLF: a possible cellular basis for the syndrome of internuclear ophthalmoplegia. In: Brooks BA, Bajandas FJ, editors. Eye movements. New York: Plenum Press; 1977.
242. Gauthier GM, Robinson DA. Adaptation of the human vestibulo-ocular reflex to magnifying lenses. Brain Res 1975;92:331–5.
243. Miles FA, Fuller JH. Adaptive plasticity in the vestibulo-ocular responses of the rhesus monkey. Brain Res 1974;80:512–6.
244. Robinson DA. Adaptive gain control of the vestibulo-ocular reflex by the cerebellum. J Neurophysiol 1976;39:954–69.
245. Kohler I. Experiments with goggles. Sci Am 1962; 206:62–84.
246. Melvill-Jones G, et al. Long-term effects of maintained vision reversal: is vestibulo-ocular adaptation either necessary or sufficient? In: Baker R, Berthoz A, editors. Control of gaze by brainstem neurons. Vol 1: developments in neuroscience. Amsterdam: Elsevier/North-Holland Biomedical; 1977.
247. Miles FA. The primate flocculus and eye-head coordination. In: Brooks BA, Bajandas FJ, editors. Eye movements. New York: Plenum Press; 1977.
248. Robinson DA. Is the cerebellum too old to learn? In: Brooks BA, Bajandas FJ, editors. Eye movements. New York: Plenum Press; 1977.
249. Simpson J, Alley KE. Visual climbing fiber input to rabbit vestibulo-cerebellum: a source of direction-specific information. Brain Res 1974;82:301–8.
250. Honrubia V, Koehn WW, Jenkins HA, Fenton WH. Effect of bilateral ablation of the vestibular cerebellum on visual-vestibular interaction. Exp Neurol 1982;75:616–26.
251. Sato Y, Kato I, Kawasaki T, et al. Failure of fixation suppression of caloric nystagmus and ocular motor abnormalities. Arch Neurol 1980;37:35–8.
252. Miles FA, Braitman DJ, Dow BM. Long-term adaptive changes in primate vestibuloocular reflex. IV. Electrophysiological observations in flocculus of adapted monkeys. J Neurophysiol 1980;43: 1477–93.

253. Lisberger SG. The neural basis for learning of simple motor skills. Science 1988;242:728–35.
254. Luebke AE, Robinson DA. Gain changes of the cat's vestibulo-ocular reflex after flocculus deactivation. Exp Brain Res 1994;98:379–90.
255. McNally WJ, Stuart WA. Physiology of the labyrinth. Chicago: American Academy of Ophthalmology and Otolaryngology; 1967.
256. Benson AJ, Jobson PH. Body sway induced by a low frequency alternating current. Int J Equilib Res 1973;3:55–61.
257. Cohen B. Vestibulo-ocular relations. In: Bach-y-Rita P, Collins CC, Hyde JE, editors. The control of eye movements. New York: Academic Press; 1971. p. 105–48.
258. Corwin JT, Jones JE, Katayama A, et al. Hair cell regeneration: the identities of progenitor cells, potential triggers, and instructive cues. In: Regeneration of vertebrate sensory receptor cells (Ciba Foundation Symposium 160). Chichester (England): Wiley Press; 1991. p. 103–19.
259. Rubel EW. Regeneration of hair cells in the avian inner ear. In: Dancer AL, Henderson D, Salvi RJ, Hamernik RP, editors. Noise-induced hearing loss. St. Louis: Mosby Yearbook; 1992.
260. Cotanche DA, Lee KH, Stone JS, Picard DA. Hair cell regeneration in the bird cochlea following noise damage or ototoxic drug damage. Anat Embryol (Berl) 1994;189:1–18.
261. Cotanche DA. Regeneration of hair cell stereociliary bundles in the chick cochlea following severe acoustic trauma. Hear Res 1987;30:181–95.
262. Cotanche DA. Video-enhanced DIC images of the noise-damaged and regenerated chick tectorial membrane. Exp Neurol 1992;115:23–6.
263. Ryals BM, Rubel EW. Hair cell regeneration after acoustic trauma in adult *Coturnix* quail. Science 1988;240:1774–6.
264. Ryals BM, Westbrook EW. TEM analysis of neural terminals on autoradiographically identified regenerated hair cells. Hear Res 1994;72:81–8.
265. Cohen YE, Saunders JC. The effects of sound overexposure on the spectral response patterns of nucleus magnocellularis in the neonatal chick. Exp Brain Res 1993;95:202–12.
266. Weisleder P, Rubel EW. Hair cell regeneration after streptomycin toxicity in the avian vestibular epithelium. J Comp Neurol 1993;331:97–110.
267. Goode CT, Carey JP, Fuchs AF, Rubel EW. Recovery of the vestibulocolic reflex after aminoglycoside ototoxicity in domestic chickens. J Neurophysiol 1999;81:1025–35.
268. Woolley SM, Wissman AM, Rubel EW. Hair cell regeneration and recovery of auditory thresholds following aminoglycoside ototoxicity in Bengalese finches. Hear Res 2001;153:181–95.
269. Girod DA, Duckert LG, Rubel EW. Possible precursors of regenerated hair cells in the avian cochlea following acoustic trauma. Hear Res 1989;42:175–94.
270. Tsue TT, Watling DL, Weisleder P, et al. Identification of hair cell progenitors and intermitotic migration of their nuclei in the normal and regenerating avian inner ear. J Neurosci 1994;14:140–52.
271. Forge A, Li L, Nevill G. Hair cell recovery in the vestibular sensory epithelia of mature guinea pigs. J Comp Neurol 1998;397:69–88.
272. Montcouquiol M, Corwin JT. Brief treatments with forskolin enhance s-phase entry in balance epithelia from the ears of rats. J Neurosci 2001;21:974–82.
273. Marazita ML, Ploughman LM, Rawlings B, et al. Genetic epidemiological studies of early-onset deafness in the U.S. school-age population. Am J Med Genet 1993;46:486–91.
274. Tseng CJ, Lalwani AK. Cracking the auditory genetic code: Part II. Syndromic hereditary hearing impairment. Am J Otol 2000;21:437–51.
275. Kalatzis V, Petit C. The fundamental and medical impacts of recent progress in research on hereditary hearing loss. Hum Mol Genet 1998;7:1589–97.
276. Smith RJ, Van Camp G. Cloning genes for non-syndromal hearing impairment. Br J Audiol 1999; 33:271–8.
277. Fischel-Ghodsian N. Genetic factors in aminoglycoside toxicity. Ann N Y Acad Sci 1999;84:99–109.
278. Lomax MI, Huang L, Cho Y, et al. Differential display and gene arrays to examine auditory plasticity. Hear Res 2000;147:293–302.
279. Luebke AE, Steiger JD, Hodges BL, Amalfitano A. A modified adenovirus can transfect cochlear hair cells in vivo wthout compromising cochlear function. Gene Ther 2001;8:789–94.

USEFUL WEBSITES

www.bme.jhu.edu/labs/chb/oto.wustl.edu/cochleafluids
www.medschool.lsumc.edu/otor/kresge.htm
www.parmly.luc.edu/parmly/
www.khri.med.umich.edu/index.htm-depts.washington.edu/hearing
www.oae.it
www.neurophys.wisc.edu/www/aud/ctl.augie.edu/perry/ear/tours.htm

SUGGESTED READING

Axelsson A. The vascular anatomy of the cochlea in the guinea pig and man. Acta Otolaryngol (Stockh) 1968;243:1.

Binder JR, Rao SM, Hammeke TA, et al. Functional magnetic resonance imaging of human auditory cortex. Ann Neurol 1994;35:662–72.

Brugge JF, Reale RA, Hind JE. The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. J Neurosci 1996;16:4420–37.

Clopton BM, Winfield JA, Flammino FJ. Tonotopic organization: review and analysis. Brain Res 1974;76:1–20.

Dallos P, He DZ, Lin X, et al. Acetylcholine, outer hair cell electromotility, and the cochlear amplifier. J Neurosci 1997;17:2212–26.

Fitzpatrick DC, Batra R, Stanford TR, Kuwada S. A neural population code of sound localization. Nature 1997;388:871–4.

Guth PS, Perin P, Norris CH, Valli P. The vestibular hair cells: post-transductional signal processing. Prog Neurobiol 1998;54:193–247.

Hackney C, Furness D. Mechanotransduction in vertebrate hair cells: structure and function of the stereociliary bundle. Am J Physiol 1995;268:C1–13.

Housley GD, Ashmore JF. Direct measurement of the action of acetylcholine on isolated outer hair cells of the guinea pig cochlea. Proc R Soc Lond B Biol Sci 1991;244:161–7.

Kimura R. The ultrastructure of the organ of Corti. Int Rev Cytol 1975;42:173–222.

Kurokawa H, Goode RL. Sound pressure gain produced by the human middle ear. Am J Otolaryngol Head Neck Surg 1995;113:349–55.

Langner G, Sams M, Heil P, Schulze H. Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. A J Comp Physiol 1997;181:665–76.

Lim DJ. Cochlear anatomy related to cochlear micromechanics: a review. J Acoust Soc Am 1980;67:1686–95.

Lim DJ. Human tympanic membrane. An ultrastructural observation. Acta Otolaryngol (Stockh) 1970;70: 176–86.

Markovitz NS, Pollak GD. Binaural processing in the dorsal nucleus of the lateral lemniscus. Hear Res 1994; 73:121–40.

Martin P, Hudspeth AJ. Active hair-bundle movements can amplify a hair cell's response to oscillatory mechanical stimuli. Proc Natl Acad Sci U S A 1999; 96:14306–11.

Minor LB. Physiological principles of vestibular function on earth and in space. Otolaryngol Head Neck Surg 1998;118:S5–15.

Niedzielski AS, Safieddine S, Wenthold RJ. Molecular analysis of excitatory amino acid receptor expression in the cochlea. Audiol Neurootol 1997;2: 79–91.

Palmer AR, Russell IJ. Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. Hear Res 1986;24: 1–15.

Pirvola U, Hallbrook F, Xing-Qun L, et al. Expression of neutrophins and Trk receptors in the developing, adult and regenerating avian cochlea. J Neurobiol 1997;33:1019–33.

Probst R, Lonsbury-Martin BL, Martin GK. A review of otoacoustic emissions. J Acoust Soc Am 1991;89: 2027–67.

Rabinowitz WM. Measurement of the acoustic input admittance of the human ear. J Acoust Soc Am 1981; 70:1025–35.

Ross M. Fluorescence and electron microscopic observations of the general visceral, efferent innervation of the inner ear. Acta Otolaryngol Suppl (Stockh) 1971; 286:1–18.

Ruggero MA. Responses to sound of the basilar membrane of the mammalian cochlea. Curr Opin Neurobiol 1992;2:449–56.

Ryan AF, Goodwin P, Woolf NK, Sharp F. Auditory stimulation alters the pattern of 2-deoxyglucose uptake in the inner ear. Brain Res 1982;234: 213–25.

Shanks JE, Lilly DJ, Margolis RH, et al. Tympanometry. J Speech Hear Disord 1988;53:354–77.

Stone JS, Oesterle EC, Rubel EW. Recent insights into regeneration of auditory and vestibular hair cells. Curr Opin Neurol 1998;11:17–24.

Warr WB, Boche JB, Neely ST. Efferent innervation of the inner hair cell region: origins and terminations of two lateral olivocochlear systems. Hear Res 1997; 108:89–111.

Wright CG. Development of the human external ear. J Am Acad Audiol 1997;8:379–82.

Yost WA, Sheft S. Auditory processing. In: Yost WA, Popper AN, Fay RF, editors. Handbook of auditory research. New York: Springer-Verlag; 1993.

Zwislocki J. The role of the external and middle ear in sound transmission. In: Tower DB, editor. The nervous system. Vol 3. Human communication and its disorders. New York: Raven Press; 1975. p. 45–55.

CHAPTER 5

Diagnostic Audiology, Hearing Aids, and Habilitation Options

James W. Hall III, PhD, M. Samantha Lewis, MA

Within the past 25 years, dramatic advances in technology and techniques have contributed to more powerful audiologic test batteries and more effective management options for pediatric and adult populations. With auditory brainstem response (ABR) and otoacoustic emissions (OAEs), newborn infants can be screened for hearing impairment within days after birth and managed audiologically within the next critical 6 months. New techniques and strategies for the assessment of auditory function in adults have also been introduced in recent years. Pure-tone audiometry, immittance measurements (tympanometry and acoustic reflexes), and calculation of word recognition scores continue to be important for hearing assessment, and the traditional audiogram remains very useful in summarizing the results of basic audiologic assessment. Clinical audiology, however, now also includes other behavioral and electrophysiologic test procedures. For example, electrocochleography (ECochG) can contribute to the diagnosis of Meniere's disease. Auditory brainstem response offers a readily accessible and relatively inexpensive means for identification of retrocochlear auditory dysfunction. A variety of speech and nonspeech behavioral measures and several cortical auditory evoked responses are available for clinical assessment of central auditory nervous system dysfunction and associated auditory processing disorders. Finally, OAEs, because of their unique sensitivity and specificity to cochlear dysfunction, have become the latest addition to the clinical audiologic test battery.

The otolaryngologist is in a pivotal position to identify children and adults at risk for hearing loss, to work closely with audiologists in diagnostic hearing assessment, and to contribute to timely and appropriate medical or surgical intervention. In this chapter, we summarize current techniques and strategies for hearing assessment of adults, with an emphasis on the application of a test battery approach that maximizes diagnostic accuracy and efficiency while minimizing test time and costs. The chapter includes a review of current hearing aid technology for nonmedical management of hearing impairment. Following the review is a summary of pediatric audiologic habilitation approaches. At the end of the chapter, we define in a glossary common audiologic terms and abbreviations.

BASIC AUDIOLOGIC TEST BATTERY

Pure-Tone Audiometry

Pure-tone audiometry, the most common hearing test, is a measure of hearing sensitivity using sinusoid stimuli at octave frequencies from 250 Hz up to 8,000 Hz and usually at two interoctave frequencies (3,000 and 6,000 Hz). The normal-hearing young (under 20 years) ear responds to frequencies from 20 to 20,000 Hz. Test results are graphed on an audiogram. Two common audiogram versions are illustrated in Figure 5–1, A and B. All audiograms include, minimally, a graph for plotting hearing threshold levels (HTLs) as a function of the frequency of pure-tone signals, although the exact format and symbols may vary.

The unit of stimulus intensity is the decibel (dB), a logarithmic unit. The intensity of any sound is defined by a ratio of its sound pressure (or sound intensity) compared to a reference sound pressure (or sound intensity). The reference sound pressure is the amount of pressure against the eardrum, caused by air molecules when a sound is present, that vibrates the eardrum and can just be detected by a normal human ear. Briefly, the relationship for sound *intensity* is described as dB = 10 log 10 (sound

intensity/reference intensity) or for sound pressure as dB = 20 log 10 (sound pressure/reference pressure). The reference sound pressure is defined as dB sound pressure level (SPL) and is derived from one of two physical quantities (0.0002 dynes/cm^2, 20 micropascals RMS [root-mean-square], or 2×10^{-5} newtons/m^2 RMS).

Clinically, the intensity of sound is not usually described in dB SPL but, rather, in *dB hearing level* (HL), with a biologic reference level. On audiograms

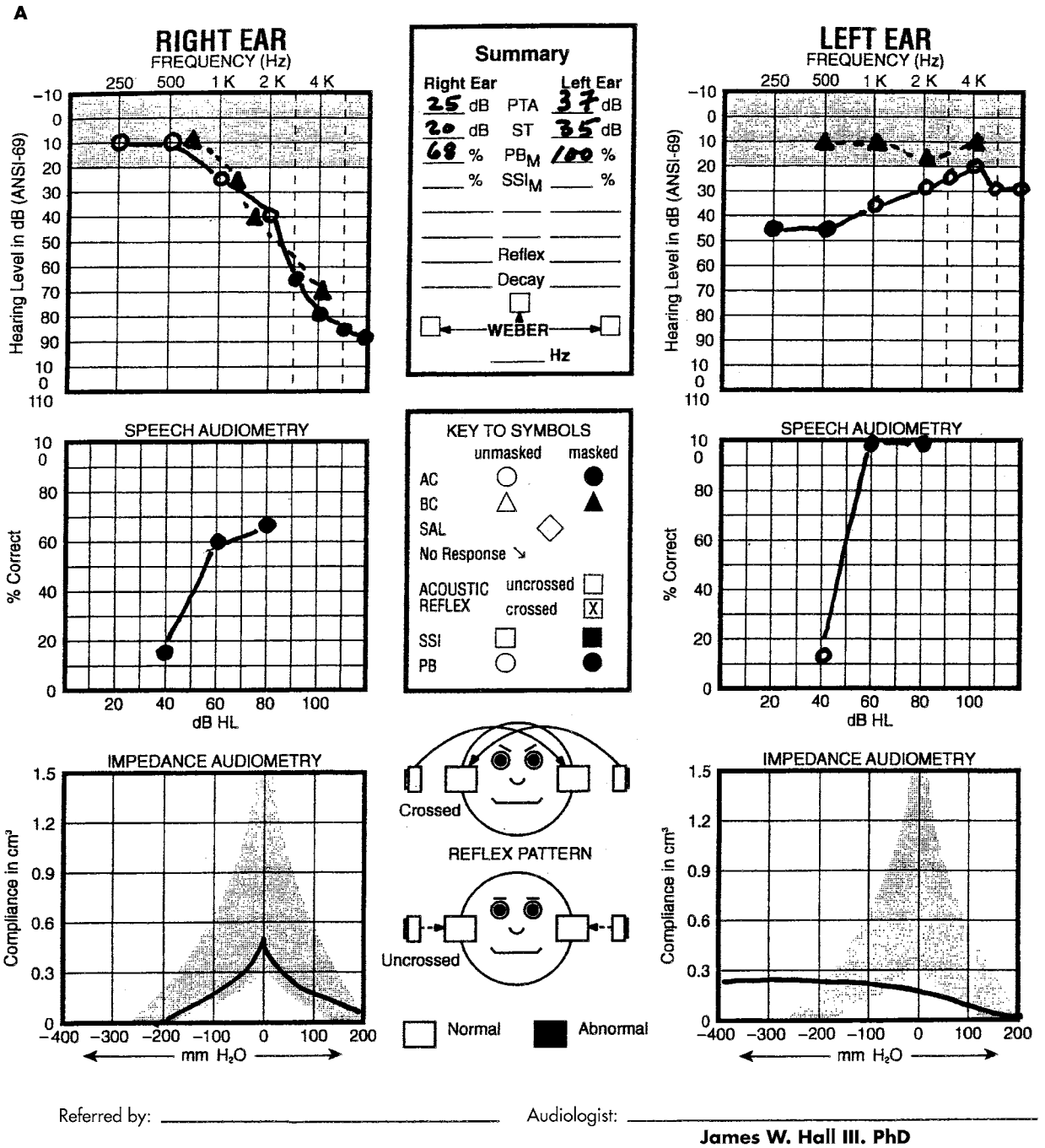


FIGURE 5–1. Two examples of audiogram formats. *A* includes sections for graphically and numerically reporting results for pure-tone audiometry (*top portion*), speech audiometry (*middle portion*), and aural immittance measurement (*bottom portion*). Masking is indicated by *filled symbols*. Findings for the right ear represent a typical sensorineural hearing loss audiometric pattern, whereas left ear findings typify a conductive hearing loss. *continued*

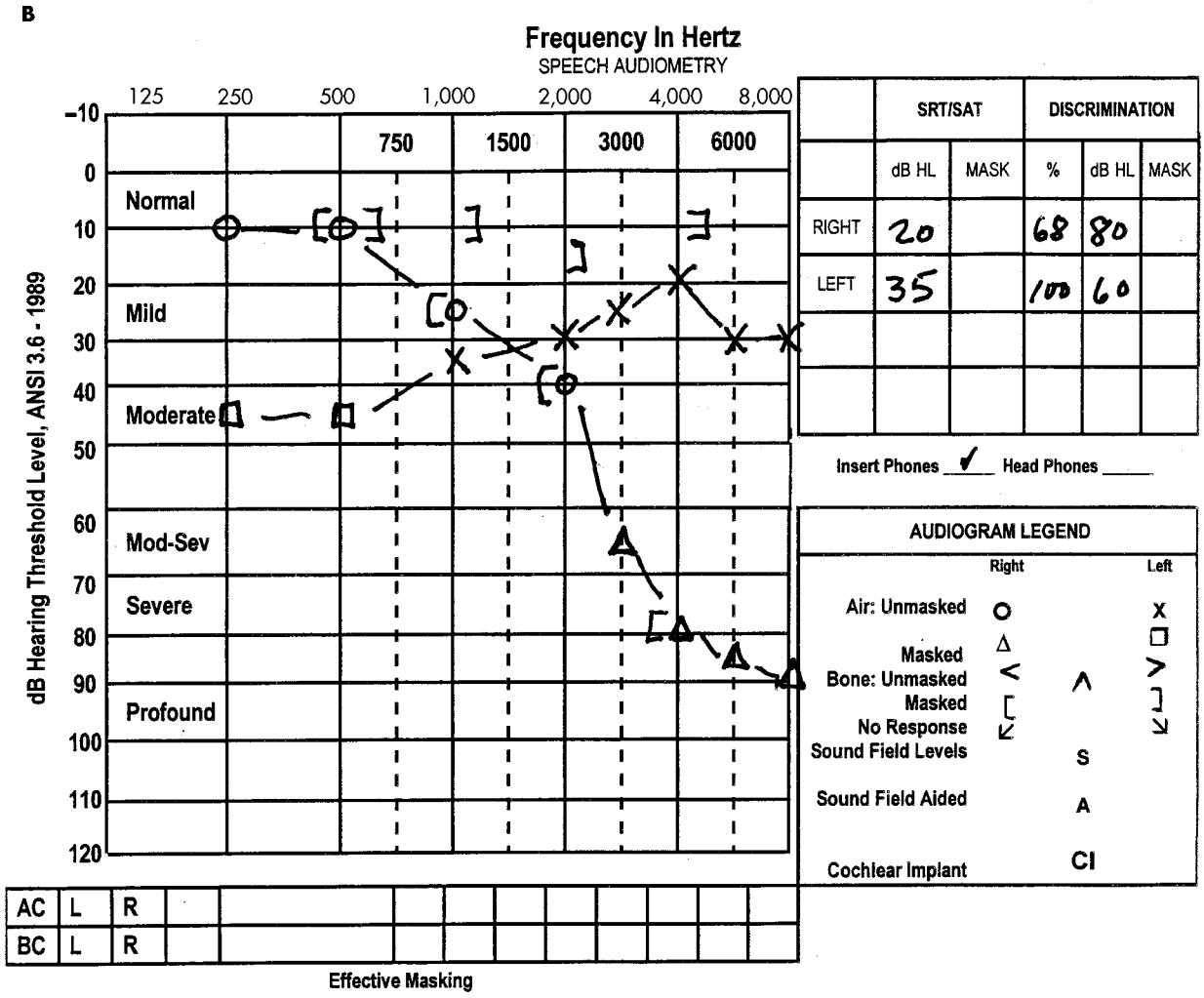


FIGURE 5–1 Continued. With *B*, the results for both ears are plotted on the same graph, requiring a separate symbol system for each ear. The same audiometric findings shown in *A* are also plotted in *B*, producing less clear analysis.

(see Figure 5–1), the dB scale has as its reference 0 dB, which is described as "audiometric 0 (zero)." This is the standard for the intensity level that corresponds to the average normal HTL, the minimal detectable intensity for each test frequency for normal hearers. Another common unit for expressing sound intensity is *dB sensation level* (SL) which is the intensity of the stimulus in dB above an individual's hearing threshold. For example, a word recognition test may be administered at an intensity level of 40 dB SL (40 dB above the patient's pure-tone average).

In adult audiologic assessment, hearing thresholds for tonal or speech signals are measured separately for each ear with earphones (air-conduction stimulation). Insert earphones (ER-3A) are now the transducer of choice for routine audiologic assessment. They offer distinct advantages over the traditional supra-aural earphones, including increased comfort, reduced likelihood of ear canal collapse, greater interaural attenuation, disposability (aural hygiene), and greater acceptance by young children.[1] Pure-tone audiometry can also be performed with stimuli presented by a bone-conduction oscillator or vibrator placed on the mastoid bone. During pure-tone audiometry, all equipment meets ANSI (American National Standards Institute) specifications, and testing is carried out according to clinical adaptations of psychoacoustic methods.[2] Patients are instructed to listen carefully for the tones and to respond (usually by pushing a button that activates a response light on the audiometer or raising their hand) every time they think they hear a tone. To

minimize interference by ambient background acoustic noise, pure-tone audiometry should always be carried out with the patient in a double-walled, sound-treated room meeting ANSI specifications.[1,3]

For adults, the clinically normal region on the audiogram is from 0 to 20 dB HL. The normal region for children is more limited because even very mild hearing loss can interfere with speech and language acquisition. Pediatric hearing threshold levels exceeding 15 dB may be considered abnormal. Thresholds in the 20 to 40 dB HL region constitute a mild hearing loss, 40 to 60 dB HL thresholds define a moderate loss, and threshold levels greater than 60 dB HL are considered a severe hearing loss.[1] As a reference, the intensity level of whispered speech close to the ear is less than 25 dB HL, conversational speech is in the 40 to 50 dB HL region, and a shouted voice within a foot of the ear is at a level of about 80 dB HL. The essential frequencies for understanding speech are in the 500 through 4,000 Hz region, although higher frequencies also contribute to discrimination between certain speech sounds. Hearing sensitivity within the "speech frequency" region is traditionally summarized by calculating the *pure-tone average* (PTA) (hearing thresholds for 500, 1,000, and 2,000 Hz divided by three and reported in dB).

The validity of audiometric results depends on whether the patient's responses result from stimulation of the test ear. If a sound of greater than 40 dB HL is presented to one ear via standard earphones with supra-aural (resting on the outer ear) cushions, it is possible that the acoustic energy will cross over from one side of the head to the other and stimulate the nontest ear. The mechanism for the crossover is presumably bone-conduction stimulation caused by vibration of the earphone cushion against the skull at high stimulus intensity levels. The amount of sound intensity needed before the crossover occurs is a reflection of *interaural attenuation*, that is, the sound insulation between the two ears provided by the head. Interaural attenuation is usually about 50 dB for lower test frequencies and 60 dB for higher test frequencies (such as those contributing to the ABR). Interaural attenuation is considerably higher for insert earphones.[1] With bone-conduction stimulation, interaural attenuation is very limited (at most 10 dB). Clinically, one must assume that interaural attenuation for bone-conducted signals is 0 dB. That is, any sound presented to the mastoid bone of one ear by a bone-conduction vibrator may be transmitted through the skull to either or both inner ears. Actual perception of this bone-conducted signal will, of course, depend on the patient's sensorineural hearing sensitivity in each ear.

Masking is the audiometric technique used to eliminate participation of the nontest ear whenever air- and bone-conduction stimulation exceeds interaural attenuation. An appropriate noise (narrowband noise for pure-tone signals and speech noise for speech signals) is presented to the nontest ear when the stimulus is presented to the test ear. With adequate masking, any signal crossing over to the nontest ear is masked by the noise. Selection of appropriate masking is sometimes difficult, especially when there is bilateral hearing impairment. One should always attempt to verify that appropriate masking was used in interpreting audiologic results when the tester is not known and, particularly, if testing was not performed by an audiologist.

Comparison of the hearing thresholds for air- versus bone-conduction signals is useful in classifying *type of hearing loss*, that is, whether a hearing loss is sensorineural (no air–bone gap), conductive (normal bone conduction and a loss by air conduction), or mixed (loss by bone conduction with a superimposed air- versus bone-conduction gap). *Configuration* refers to hearing loss as a function of the test frequency. With the sloping configuration, hearing is better for low frequencies and then becomes poorer for higher frequencies. High-frequency deficit hearing loss is the most common pattern associated with a sensorineural hearing impairment. A rising configuration is typified by relatively poor hearing for lower-frequency stimuli and better hearing for the high frequencies. The rising configuration can result from varied types of middle ear pathology. One exception to the typical association of conductive hearing loss with rising configuration is Meniere's disease (see Chapter 20). Meniere's disease is a manifestation of a cochlear lesion that produces a rising configuration. A flat audiometric configuration is often recorded from patients with mixed hearing loss, that is, when both sensorineural and conductive components are present.

Speech Audiometry

Speech audiometry measures how well a person hears and understands speech signals. Speech audiometry procedures are used routinely to measure hearing

sensitivity (thresholds in dB) for words or to estimate word recognition (ie, speech discrimination) ability. Spondee reception threshold (SRT), also referred to as the speech reception threshold (SRT) or speech threshold (ST), is the softest intensity level at which a patient can correctly repeat words approximately 50% of the time. Spondee words, two-syllable words with equal stress on each syllable (eg, airplane, baseball, cowboy), are presented to the patient monaurally via earphones. The technique is equivalent to the method for determining pure-tone thresholds described previously.

Because the PTA indicates HTLs in the speech frequency region and ST or SRT is measured with a speech signal, close agreement between the PTA and the ST is expected. If the difference between PTA and ST exceeds ±7 dB, there is reason to suspect that one or both of the measures is invalid. An unusually good ST relative to PTA (eg, ST of 5 dB and PTA of 45 dB) should immediately alert the tester to the possibility of a nonorganic hearing loss, as in malingering. With cooperative adult patients, particularly if pure-tone hearing thresholds are within the normal region from 500 to 4,000 Hz, there is probably little or no clinical benefit in measuring speech thresholds. Test time can be saved, with no loss of diagnostic information, by excluding speech threshold measurement from the test battery for such patients.

The common clinical approach for estimating a person's ability to hear and understand speech is speech recognition for phonetically balanced (PB) words.[1] Usually, a list of 25 or 50 single-syllable words is presented to the patient via earphones at one or more fixed intensity levels, and the percentage of words correctly repeated by the patient is calculated by the tester. One ear is tested at a time. Within the list of words, specific speech sounds (phonemes) occur approximately as often as they would in everyday conversation (they are "phonetically balanced"). Traditionally, these words were spoken into a microphone by the tester, while the level was monitored with a VU (volume unit) meter. Then the words were routed to the patient through the audiometer after selection of the test ear and desired intensity level. This is, however, an outdated and poor clinical practice since it lacks standardization and consistency and increases the variability of test outcome. With adult patients, it is almost always possible and always preferable to use professionally produced (and commercially available) speech materials presented via a compact disc player and an audiometer.[1] Diagnostic speech audiometry using more sophisticated materials (eg, spectrally degraded or temporally distorted speech, or speech in noise materials) is feasible for assessment of the central auditory system.[1,4]

Immittance Measurement

Aural immittance (impedance) measures are an important part of the basic audiometry test battery. Immittance is a term derived from the terms for two related techniques for assessing middle ear function (*im*pedance and ad*mittance*), techniques that have been applied clinically since 1970.[5] Briefly, the external ear canal is sealed with a soft rubber probe tip.[1] Connected to the probe tip is a device producing a tone that is delivered toward the eardrum. Middle ear impedance or admittance is calculated from the intensity and other physical properties (eg, phase) of the tone in the ear canal. A middle ear (tympanic membrane and ossicles) system with low impedance (high admittance) more readily accepts the acoustic energy of the probe tone, whereas a middle ear with abnormally high impedance (low admittance) caused, for example, by fluid within the middle ear space tends to reject energy flow. Thus, impedance (admittance) characteristics of the middle ear system can be inferred objectively with this quick and noninvasive technique and then related to well-known patterns of findings for various types of middle ear lesions.

Tympanometry Tympanometry is the continuous recording of middle ear impedance as air pressure in the ear canal is systematically increased or decreased. The technique is a sensitive measure of tympanic membrane integrity and middle ear function (Figure 5–2). Compliance (the reciprocal of stiffness) of the middle ear, the dominant component of immittance, is the vertical dimension of a tympanogram. Tympanometry is popular clinically because it requires minimal technical skill and less than a minute to perform. Because immittance measurement is an electrophysiologic (versus behavioral) method, it does not depend on cooperation of the patient. Importantly, it is a sensitive measure of middle ear function. Tympanometric patterns, in combination with audiogram patterns, permit differentiation among and classification of middle ear disorders.

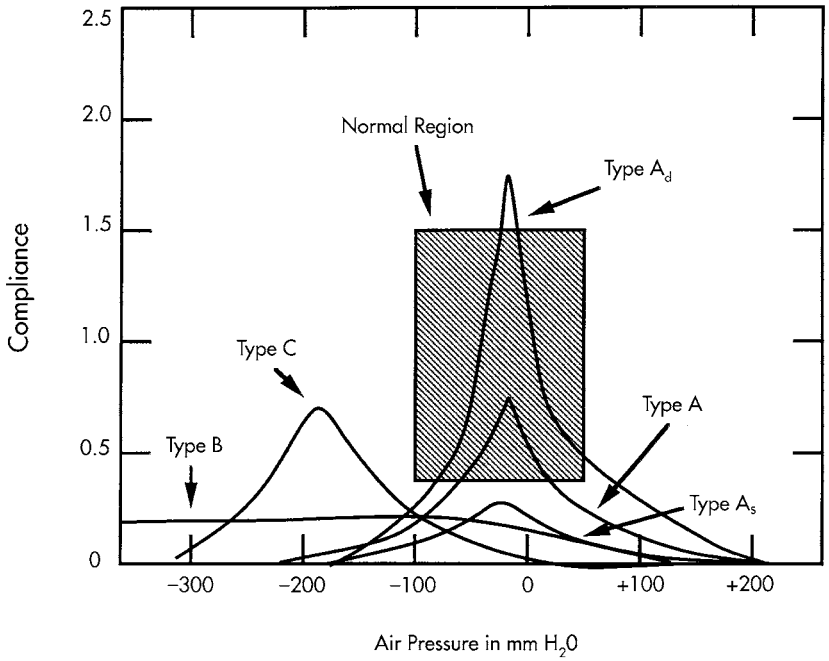


FIGURE 5–2. Classification system for tympanograms. Adapted from Jerger J.[5]

In 1970, James Jerger described what has become the most clinically widespread approach for describing tympanograms. There are three general tympanogram types: *A*, *B*, and *C*. The normal, or type A, tympanogram has a distinct peak in compliance within 0 to –100 mm of water pressure (decapascals) in the ear canal (see Figure 5–2). To be classified as normal, the location of the compliance peak on the pressure dimension and the height of the peak must be within the normal range, indicated in the figure by the stippled area. With a type B tympanogram, there is no peak in compliance but, rather, a flat pattern with little or even no apparent change in compliance as a function of pressure in the ear canal. This pattern is most often associated with fluid within the middle ear space (eg, otitis media), although other middle ear lesions may give rise to a type B tympanogram as well. Type C tympanograms also have a distinct peak in compliance (as with the type A), but the peak is within the negative pressure region beyond –150 mm H_2O (or dkPa). This pattern is usually found in patients with eustachian tube dysfunction and inadequate ventilation of the middle ear space and often precedes the type B tympanogram in the development of otitis media.

The type A_s (see Figure 5–2) is a variation of the type A tympanogram. The "s"stands for shallow. Peak compliance is below the lower normal limits of compliance. That is, middle ear impedance is abnormally high. The type A_s pattern is typically recorded from patients with fixation of the ossicular chain, including those with the diagnosis of otosclerosis. In contrast, with an usually steep and high-compliance tympanogram (type A_d for deep), the peak may actually exceed the upper compliance limits of the equipment. The A_d tympanogram type is recorded from patients with disruption of the ossicular chain, which leaves the middle ear extremely mobile and hypercompliant (ie, very little impedance). In the absence of serious hearing loss, however, this tympanogram pattern is usually associated with minor tympanic membrane abnormality, such as scarring. During the initial portion of the tympanometry procedure, high positive or negative pressure is introduced into the ear canal. This essentially decouples the middle ear system from the measurement. If the immittance device records an abnormally large equivalent volume of air (eg, 2 cm^3 or more in an adult or twice the volume that was recorded for the other ear) between the probe tip and presumably the eardrum at this stage in the procedure, the integrity of the eardrum should be questioned. That is, the immittance device is recording not just ear canal volume but also volume of the middle ear space. This test finding is consistent with either a perforation of the tympanic membrane or an open (patent) middle ear ventilation tube.

Acoustic Stapedial Reflex Measurement Measurement of contractions of the middle ear stapedial mus-

cle to high sound intensity levels (usually 80 dB or greater) is the basis of the acoustic reflex test. The stapedial muscle within the middle ear is the smallest muscle in the body. Acoustic reflex measurement is clinically useful for estimating hearing sensitivity and for differentiating among sites of auditory disorders, including the middle ear, inner ear, eighth cranial nerve, and auditory brainstem. The afferent portion of the acoustic reflex arc is the eighth cranial nerve. There are complex brainstem pathways leading from the cochlear nucleus on the stimulated side to the region of the motor nucleus of the seventh (facial) nerve on both sides (ipsilateral and contralateral to the stimulus) of the brainstem. The efferent portion of the arc is the seventh nerve innervating the stapedius muscle. When the muscle contracts, the result is increased stiffness (decreased compliance) of the middle ear system. This small change in compliance following stapedius muscle contraction (within 10 ms) is detected by the probe and immittance device.

Measurement of the acoustic reflex is useful clinically because it can quickly provide objective information on the status of the auditory system from the middle ear to the brainstem. Distinctive acoustic reflex patterns for ipsilateral and contralateral stimulation and measurement conditions characterize middle ear, cochlear, eighth nerve, brainstem, and even facial nerve dysfunction (see "faces" in lower portion of Figure 5–1, A). Comparison of acoustic reflex threshold levels, the lowest stimulus intensity level that activates the reflex, for tonal versus noise signals permits estimation of the degree of cochlear hearing impairment.[1] This technique is especially valuable in children and difficult-to-test patients.

AUDITORY EVOKED RESPONSES

Auditory evoked responses are electrophysiologic recordings of responses to sounds. With proper test protocols, the responses can be recorded clinically from activation of all levels of the auditory system, from the cochlea to the cortex. More than a dozen subtypes of auditory evoked responses can be recorded beyond the brainstem from auditory regions of the thalamus, hippocampus, internal capsule, and cortex. Prominant among the cortical evoked responses measured for clinical purposes are the auditory middle latency response (AMLR), the auditory late response (ALR), the P300 response, and the mismatch negativity (MMN) response.[1,6] In fact, cortical auditory evoked responses were reported as early as the 1930s, and, with the exception of the MMN, all of the above responses were well described before the ABR was even discovered. Cortical auditory evoked responses are characterized by longer latencies (100 to 300 ms) than ECochG and ABR because they arise from more rostral regions of the auditory central nervous system and are dependent on multisynaptic pathways. Amplitudes of the cortical responses are considerably larger (2 to 20 times larger) than those of the earlier responses because they reflect activity evoked from a greater number of neurons. Measurement parameters are distinctly different for the cortical versus cochlear or brainstem responses. For example, stimulus rate must be slower and physiologic filter settings lower. As a rule, stimulus intensities are moderate rather than high. Cortical evoked responses are best elicited with longer-duration, and therefore frequency-specific, tonal stimuli rather than the click stimuli that are optimal for evoking ECochG and ABR. The analysis time must, of course, extend beyond the expected latency of the response (> 300 ms) for the cortical responses. Recording electrode sites also are different for the cortical responses, with more emphasis on scalp sites over the hemispheres and less concern about electrode sites near the ears.

AUDITORY BRAINSTEM RESPONSE

Among the many auditory evoked responses, the ABR (often referred to by neurologists as the brainstem auditory evoked response or BAER) and ECochG are applied most often clinically. An ABR recording is shown schematically in Figure 5–3. The ABR is generated with transient acoustic stimuli (clicks or tone bursts) and detected with surface electrodes (disks) placed on the forehead and near the ears (earlobe or within the external ear canal). Using a computer-based device, it is possible to present rapidly (eg, at rates of 20 to 30 per second) thousands of sound stimuli and to average reliable ABR waveforms in a matter of minutes. Automated devices are now available for special clinical applications, for example, newborn hearing screening. The anatomic generators of the ABR components have been studied extensively.[6] Waves I through V arise from the eighth cranial nerve and auditory regions

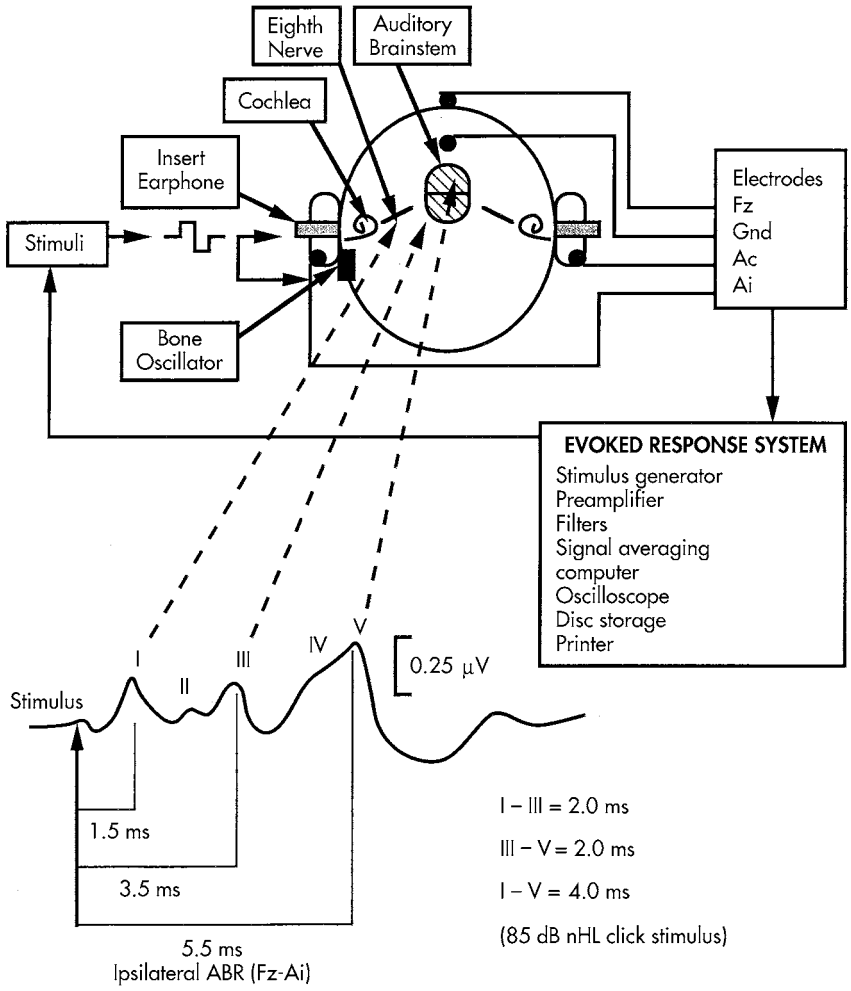


FIGURE 5–3. Schematic of the instrumentation used for recording the auditory brainstem response (ABR) and major relations among auditory anatomy and waveform components. A simple strategy for analysis of ABR waveform in neurodiagnosis is also shown.

in the caudal and rostral brainstem (see Figure 5–3). Wave I clearly represents the synchronously stimulated compound action potentials from the distal (cochlear) end of the eighth cranial nerve. Wave II may also arise from the eighth nerve, but near the brainstem (the proximal end). Waves I and II are generated by structures ipsilateral to the ear stimulated. All later ABR waves have multiple generators within the auditory brainstem. Wave III, which is usually prominent, is generated within the caudal pons, with likely contributions from the cochlear nuclei, the trapezoid body, and the superior olivary complex. The largest and most rostral component of the ABR, wave V, is generated in the region of the lateral lemniscus as it approaches the inferior colliculus, presumably mostly on the side contralateral to the ear stimulated.

The first objective in ABR waveform analysis is to be sure that the response is reliably recorded. Minimally, two replicated waveforms should be averaged. If the response is not highly replicable, modifications in the test protocol should be made, and then potential technical problems must be considered and systematically ruled out. When a replicable response is confirmed, absolute latencies for each replicable wave component and relative (interwave) latencies between components are calculated in milliseconds. These latency data for each ear are assessed for symmetry (wave V should be within 0.4 ms between ears) and also compared with appropriate normative data. Common ABR waveform patterns are illustrated in Figure 5–3. A well-formed and clear wave I at a delayed latency value for the maximum stimulus intensity level is characteristic of a conductive or mixed hearing loss. When wave I is small and poorly formed but interwave latency values are within normal limits (the wave I to V latency value is less than 4.60 ms), a high-frequency sensory (cochlear) hearing loss is suspected. Delayed interwave latency values are the signature of retrocochlear auditory

dysfunction. Abnormal delays between the early wave components (eg, I to III) are consistent with posterior fossa lesions involving the eighth nerve and/or lower brainstem, whereas a prolonged III to V latency suggests intra-axial auditory brainstem dysfunction.

A primary goal in any neurodiagnostic ABR is to record a clear and reliable wave I component. Wave I serves as the benchmark for peripheral auditory function. Subsequent interwave latencies offer indices of retrocochlear (eighth nerve and brainstem) function that are relatively unaffected by conductive or sensory hearing loss. The likelihood that a wave I will be recorded is enhanced by the use of either ear canal (TIPtrode™) or tympanic membrane electrode designs, along with alterations of the test protocol, such as a slower stimulus rate, rarefaction stimulus polarity, and maximum stimulus intensity level. Reports on ABR dating back to the late 1970s confirmed that waveforms evoked by high intensity yielded neurodiagnostic information on cochlear and retrocochlear auditory function and could be applied in the identification of retrocochlear disorders (eg, acoustic neuromas), with a hit rate exceeding 95%. With the development of sophisticated neuroradiologic techniques, such as computed tomography (CT) or magnetic resonance imaging (MRI) with enhancement, reports of normal ABR findings among patients with very small posterior fossa tumors have appeared in the literature. These are clearly false-negative ABR outcomes in patients at risk for retrocochlear auditory dysfunction usually caused by very small intracanalicular vestibular schwannomas. On the other hand, false-positive outcomes for CT and MRI have also been reported for patients with normal ABRs and no surgical evidence of tumor.[1] Auditory brainstem response continues to be a readily available, relatively inexpensive, and reasonably sensitive procedure for initial diagnostic assessment of eighth nerve and auditory brainstem status in patients with retrocochlear signs and symptoms. Auditory brainstem response is also valuable in electrophysiologic monitoring of the eighth nerve and auditory brainstem function during certain neurotologic surgeries (eg, vestibular nerve section, posterior fossa tumor removal).

Electrocochleography

For over 30 years, ECochG has been applied in the assessment of peripheral auditory function. Currently, ECochG is performed most often for intraoperative monitoring of cochlear and eighth nerve status and in the diagnosis of Meniere's disease. Optimal ECochG waveforms are recorded from a small needle electrode placed through the tympanic membrane to the promontory, although tympanic membrane and, to a lesser extent, ear canal electrode locations are also clinically useful. Stimulus and acquisition parameters for recording ECochG have been well defined for decades.[6] The three major components of the ECochG are the cochlear microphonic (CM), the summating potential (SP), and the action potential (AP). The CM and SP reflect cochlear bioelectric activity, whereas the AP is generated by synchronous firing of distal afferent eighth nerve fibers and is equivalent to ABR wave I (Figure 5–4). The typical ECochG analysis technique in neurotology requires determination of the amplitude of the SP and the AP from a common baseline. Then the ratio of the SP to the AP is calculated and reported in percent. Normal SP to AP ratio ranges and cutoffs in percent have been reported for each electrode type. Abnormal SP to AP ratio values are defined as > 50% for the ear canal electrode type (the TIPtrode), > 40% for the tympanic membrane electrode, and > 30% for the transtympanic needle electrode type.[1,6]

The characteristic ECochG finding for patients with Meniere's disease is an abnormal enlargement of the SP amplitude relative to the AP amplitude. With the tympanic membrane electrode technique, sensitivity of ECochG in the diagnosis of endolymphatic hydrops is reported as 57%, whereas specificity is 94% in a series of 100 patients.[7] Only 3 of 30 patients yielded false-positive findings. Thus, an abnormally high SP to AP ratio is highly suggestive of endolymphatic hydrops according to these data. In this study, the likelihood of an abnormal ECochG SP to AP ratio was statistically higher as hearing loss increased and when the hearing loss fluctuated.

OTOACOUSTIC EMISSIONS

Otoacoustic emissions are low-intensity sounds produced by the cochlea in response to an acoustic stimulus.[8] A moderate-intensity click, or an appropriate combination of two tones, can evoke outer hair cell movement or motility. Outer hair cell motility affects basilar membrane biomechanics, resulting in a form of intracochlear energy amplification, as well as cochlear tuning for more precise frequency

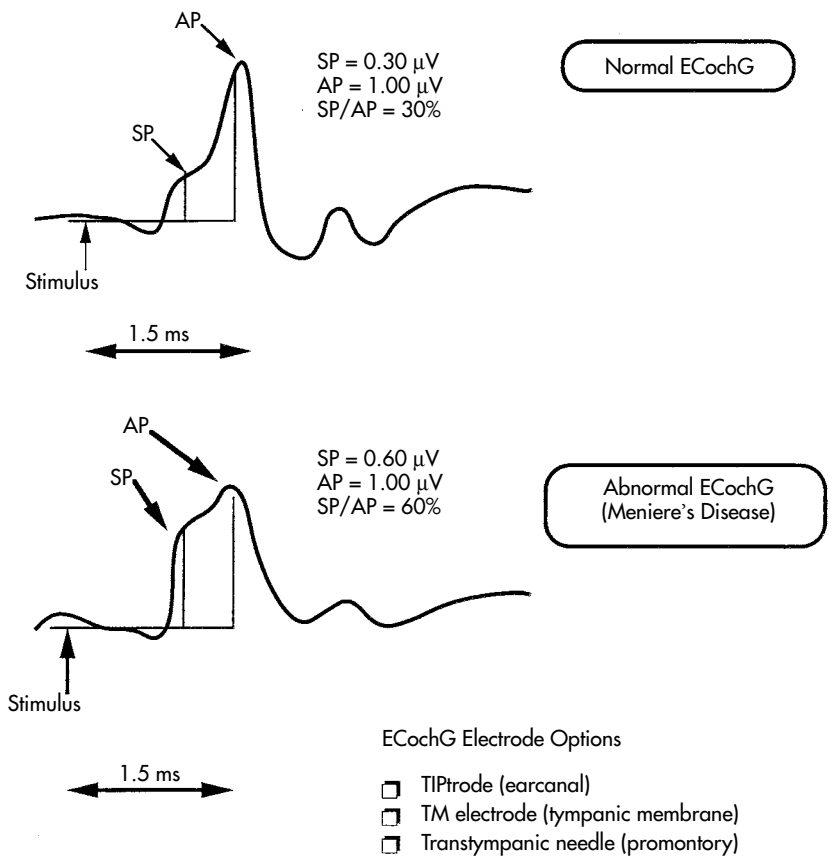


FIGURE 5–4. Electrocochleography (ECochG) waveforms illustrating normal relation of summating potential (SP) and action potential (AP) and abnormally enlarged SP/AP relation in a patient with Meniere's disease. Absolute and relative amplitude values for the SP and AP components and the criteria for definition of a normal response vary significantly for different electrode sites (ear canal versus tympanic membrane versus promontory). For a discussion of ECochG measurement and analysis, see Schwaber MK, Hall JW. A simplified technique for transtympanic electrocochleography. Am J Otol 1990;11:260–5.

resolution. The outer hair cell motility generates mechanical energy within the cochlea that is propagated outward via the middle ear system and the tympanic membrane to the ear canal. Vibration of the tympanic membrane then produces an acoustic signal (the OAE), which can be measured by a sensitive microphone. There are two broad classes of OAEs: spontaneous and evoked. *Spontaneous* otoacoustic emissions (SOAEs), present in only about 70% of persons with normal hearing, are measured in the external ear canal when there is no external sound stimulation. A significant gender effect for SOAE has been confirmed, with females demonstrating SOAEs at twice the rate of males.

Evoked OAEs, elicited by moderate levels (50 to 80 dB SPL) of acoustic stimulation in the external ear canal, are generally classified according to characteristics of the stimuli used to elicit them or characteristics of the cochlear events that generate them. *Stimulus-frequency* otoacoustic emissions (SFOAEs), which are technically difficult to record, are the least studied of the evoked OAEs. *Distortion-product* otoacoustic emissions (DPOAEs) are produced when two pure-tone stimuli at frequencies f_1 and f_2 are presented to the ear simultaneously (Figure 5–5). The most robust DPOAE occurs at the frequency determined by the equation $2f_1$-f_2, whereas the actual cochlear frequency region that is assessed with DPOAE is between these two frequencies and probably close to the f_2 stimulus for recommended test protocols. With all of the five US Food and Drug Administration (FDA)-approved instruments that are commercially available for recording DPOAE, amplitude as detected in the ear canal and described in dB SPL is plotted as a function of the frequencies of the stimuli in a DPOAEgram (see Figure 5–5). *Transient evoked* otoacoustic emissions (TEOAEs) are elicited by brief acoustic stimuli such as clicks or tone bursts. Although there are distinct differences in the methodology for recording DPOAEs versus TEOAEs, and the exact cochlear mechanisms responsible for their generation are also different, each type of evoked OAE is now being incorporated into routine auditory assessment of children and adults.[8] Clinical applications of OAEs, along with their rationale, are summarized in Table 5–1.

When outer hair cells are structurally damaged or at least nonfunctional, OAEs cannot be evoked by

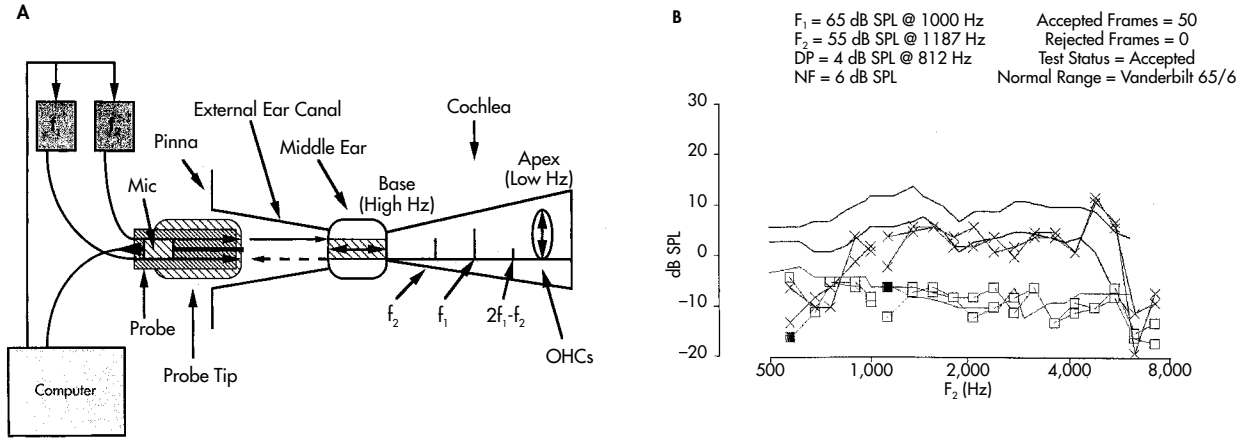


FIGURE 5–5. *A*, Schematic illustration of the generation of distortion-product otoacoustic emissions (DPOAEs) from outer hair cells (OHCs) within the cochlea showing the two stimulus-frequency generators, the probe for delivery of the stimuli and housing of a microphone (Mic) for detection of the OAE in the canal, and the inward propagation of the stimuli to the cochlea and the outward propagation of the DPOAEs from the cochlea. *B*, DPgram recorded from the left ear of a 5-year-old girl with a ventilation tube in the tympanic membrane, yet hearing sensitivity is within normal limits. Replicated plots of DPOAE amplitudes in dB SPL (*the X symbols*) are shown as a function of the f_2 stimuli in a "DPgram" format. Noise levels as a function of frequency are indicated in the lower portion of the DPgram by the *square symbols*. The *faint solid line* represents the upper limits (95th percentile) for normal noise levels. Distortion-product amplitudes are reduced in the low-frequency region (500 to 1,000 Hz), consistent with a slight air–bone gap in pure-tone hearing thresholds. Distortion-product amplitudes for higher frequencies are generally within normal limits (the two *solid lines* representing the 5th to 95th percentile region for normal adult subjects). NF = noise floor.

acoustic stimuli. Among patients with mild cochlear dysfunction, OAEs may be recorded, but amplitudes are below normal limits for some or all stimulus frequencies. Importantly, some patients with abnormal OAEs, consistent with cochlear dysfunction, will have normal pure-tone audiograms. An example is a patient with the symptom of tinnitus yet a normal audiogram. Abnormal OAEs are expected in the frequency region represented by the tinnitus. Up to 30% of a population of outer hair cells may be damaged without substantially affecting the simple audiogram. In such cases, however, abnormal OAE findings are invariably recorded. The noninvasive nature of OAE recording, coupled with their accuracy and objectivity in assessing cochlear, in particular outer hair cell, function suggests diverse potential clinical applications, ranging from auditory screening to sensorineural diagnosis.

INDICATIONS FOR DIAGNOSTIC AUDIOLOGIC ASSESSMENT

CHILDREN

Hearing loss influences speech and language development of infants and young children, and these effects begin within the first 6 months of life.[9,10] Even mild peripheral hearing deficits among preschool- and school-age children interfere with educational development. Therefore, early identification of hearing loss, coupled with timely and appropriate intervention and management, is nec-

TABLE 5–1. Clinical Applications of Otoacoustic Emissions (OAEs) and Their Rationale

Application	*Rationale*
Newborn hearing screening	OAEs can be recorded reliably from newborn infants
	OAE recording can be performed in nursery setting (test performance may be affected by noise)
	Normal OAEs are recorded in persons with normal sensory (cochlear) function
	OAEs are abnormal in persons with even mild degrees of sensory hearing loss; the main objective of screening is to detect sensory hearing impairment
	OAE recording may require a relatively brief test time
	OAE measurement may be performed by nonaudiologic personnel (ie, at reduced cost)
Pediatric audiometry	OAE recording is electrophysiologic and not dependent on patient behavioral response
	OAEs assess cochlear function specifically (behavioral audiometry and ABR are dependent also on the status of the central auditory nervous system)
	OAEs can be recorded from sleeping or sedated children
	OAE recording requires a relatively short test time
	OAEs provide ear-specific audiologic information
	OAEs provide frequency-specific audiologic information
	OAEs are a valuable contribution to the "crosscheck principle"*
Diagnosis of central auditory processing disorders	OAE recording is electrophysiologic and not dependent on patient behavioral response OAEs are a sensitive means of ruling out cochlear (outer hair cell) dysfunction
Assessment in suspected functional hearing loss	OAE recording is electrophysiologic and not dependent on patient behavioral response
	Normal OAEs invariably imply normal sensory function
	OAEs provide frequency-specific audiologic information
Differentiation of cochlear versus retrocochlear auditory dysfunction	OAEs are site specific for cochlear (sensory) auditory dysfunction In combination with ABR, OAEs can clearly distinguish sensory versus neural auditory disorders
Monitoring ototoxicity	OAEs are site specific for cochlear (sensory) auditory dysfunction
	Ototoxic drugs exert their effect on outer hair cell function; OAEs are dependent on outer hair cell integrity
	OAE recording is electrophysiologic and not dependent on patient behavioral response; can be recorded from patients who, owing to their medical condition, are unable to perform behavioral audiometry tasks or from infants and young children
	OAEs can detect cochlear dysfunction before it is evident by pure-tone audiometry
	OAEs provide frequency-specific audiologic information
Tinnitus	OAEs are site specific for cochlear (sensory) auditory dysfunction
	OAEs can provide objective confirmation of cochlear dysfunction in patients with tinnitus and normal audiograms

Continued

TABLE 5–1 *Continued*. Clinical Applications of Otoacoustic Emissions (OAEs) and Their Rationale

Application	*Rationale*
	OAEs provide frequency-specific audiologic information that may be associated with the frequency region of tinnitus
Noise/music exposure	OAEs are site specific for cochlear (sensory) auditory dysfunction
	Excessive noise/music intensity levels affect outer hair cell function; OAEs are dependent on outer hair cell integrity
	OAEs can provide objective confirmation of cochlear dysfunction in patients with normal audiograms
	OAE findings are associated with cochlear frequency specificity, ie, "tuning"; musician complaints of auditory dysfunction can be confirmed by OAE findings, even with a normal audiogram
	OAEs can provide an early and reliable "warning sign" of cochlear dysfunction owing to noise/music exposure before any problem is evident in the audiogram

*The crosscheck principle in pediatric audiology is that the results of any single audiologic test should not be relied on without independent corroboration or crosscheck from another audiologic test.
Adapted from Hall JW.[8] ABR = auditory brainstem response.

essary for children to reach their full communicative and educational potential.[11] A number of factors put children at risk for hearing impairment. Table 5–2 summarizes risk factors for hearing impairment identified by the 2000 Joint Committee on Infant Hearing.[12] Otolaryngologists and audiologists must coordinate efforts in properly and promptly assessing and managing pediatric hearing impairment.

In addition to peripheral hearing impairment secondary to common causes, such as conductive hearing loss with otitis media or sensorineural hearing loss in ototoxicity, one must also consider the possibility of auditory processing disorders (APDs) in children. According to a recent consensus conference report,[4] auditory processing disorders are defined as "deficits in the processing of information that is specific to the auditory modality." Often APDs are related to central auditory nervous system dysfunction. Common risk factors for APDs in school-age children include the following:

- teacher concern about hearing despite a normal audiogram
- academic underachievement
- recurrent middle ear disease
- suggestion of APD on language evaluation
- poor response to language treatment
- attention-deficit disorder (ADD) with or without hyperactivity
- reading delay or disorder
- learning disabilities
- history of central nervous system insult (eg, head injury, neonatal asphyxia)

There are now a variety of behavioral and electrophysiologic measures for diagnosis of APDs. Electrophysiologic procedures were outlined above. Among them, cortical evoked responses are especially appropriate for evaluation of APD. The behavioral procedures include measures of detection (eg, auditory integration tasks), suprathreshold discrimination (eg, temporal ordering tasks), and identification (eg, recognition of speech signals from phonemes to sentences), presented monotically, diotically, and dichotically.[4] Speech and nonspeech materials for assessment of APDs are available in compact disc format from multiple commercial sources.[1,13]

Adults

In adults, risk factors for central auditory nervous system dysfunction include, but are not limited to, advanced age and history or clinical evidence of stroke, head injury, brain neoplasms, Alzheimer's disease, and other disorders affecting the central nerv-

TABLE 5–2. Joint Committee on Infant Hearing Screening 2000 Indicators Associated with Sensorineural Hearing Loss

Birth through age 28 d (neonate)
Family history of congenital or delayed-onset childhood hereditary sensorineural hearing loss
Congenital infection, such as toxoplasmosis, syphilis, rubella, cytomegalovirus, and herpes
Craniofacial anomalies including abnormalities of the pinna and ear canal, absent philtrum, and low hairline.
Birth weight less than 1,500 g (3.3 lb); hyperbilirubinemia at level requiring exchange transfusion
Ototoxic medications including but not limited to the aminoglycosides (eg, gentamicin, tobramycin, kanamycin, streptomycin) used in multiple courses or in combination with loop diuretics
Bacterial meningitis
Apgar scores of 0 to 4 at 1 min or 0 to 6 at 5 min
Mechanical ventilation lasting 5 days or longer
Stigmata or other findings associated with a syndrome known to include sensorineural and/or conductive hearing loss (eg, Waardenburg's or Usher's syndrome).
Age 29 d through 2 y (infant)
Parent caregiver concern regarding hearing, speech, language, and/or developmental delay
Bacterial meningitis and other infections associated with sensorineural hearing loss
Ototoxic medications including but not limited to chemotherapeutic agents or aminoglycosides used in multiple courses or in combination with loop diuretics
Recurrent or persistent otitis media with effusion for at least 3 mo
For use with infants (age 29 d through 3 y) requiring monitoring of hearing
Indicators associated with delayed-onset sensorineural hearing loss
Family history of hereditary childhood hearing loss
In utero infections (eg, cytomegalovirus, rubella, syphilis, herpes, or toxoplasmosis)
Neurofibromatosis 2 and neurodegenerative disorders
Indicators associated with conductive hearing loss
Recurrent or persistent otitis media with effusion
Anatomic deformities and other disorders that affect eustachian tube function
Neurodegenerative disorders

ous system. In general, it is a good clinical policy to always consider the possibility of central auditory dysfunction when a patient's hearing complaints exceed expectations based on the audiogram.

HEARING AIDS

A hearing aid is a medical device that delivers an amplified acoustic signal into the ear canal. The use of some form of amplification to overcome the deleterious effects of a hearing impairment on communication has been used for centuries. The most simplistic "hearing aid" is easily made by cupping one's hand behind the ear. This technique is effective because it helps channel the acoustic signal into the ear canal. Another simple means for improving the reception of an acoustic signal is to move closer to the sound source. When the distance between the listener and the sound source is doubled, the acoustic intensity is reduced by 6 dB at the listener. Similarly, a reduction in the distance by one half increases the intensity by 6 dB. When the acoustic signal is from an electronic device, such as a television or radio, reception can easily be enhanced by adjusting the volume control to make a more intense signal. All of these methods enhance the

reception of a particular acoustic signal and demonstrate that amplification is often an effective means of improving one's ability to hear. Although there are a few exceptions, amplification can be used to improve hearing in virtually all forms of hearing impairment as well.

Electronic hearing aids enhance the signal strength by amplifying the acoustic signals that reach its microphone. Hearing aid design has evolved rapidly with advances in electronics, with instruments becoming smaller, more reliable, and more efficient. Today's hearing aids are small enough to fit completely within the ear canal, and some have advanced technological features that can significantly improve listening in noise. Perhaps the most amazing aspect of hearing aid design is that only a single disposable button-type battery of less than 1.5 V powers most units.[14] Average battery life ranges between 10 and 14 16-hour days.

Electronic hearing aids comprise three basic components: a microphone, an amplifier, and a receiver, which are housed in a plastic case designed to fit behind or in the ear. The *microphone* is located externally and is activated by miniscule fluctuations in air pressure caused by the sound source. These vibrations are converted into an electrical impulse and delivered to the *amplifier*. At this stage, the signal is amplified and filtered. The altered signal is delivered to the *receiver*, which converts the electronic impulse into an acoustic signal. The amplified acoustic signal is then delivered to the ear canal for reception of the signal by the wearer.[15]

Types of Hearing Aids

There are basically two ways in which a hearing aid may be classified. The first is by the style of the hearing aid, which usually refers to the size and placement of the aid. The second is by the technological features of the device.

Hearing Aid Styles The most commonly used hearing aid styles (although other types do exist) are the behind-the-ear (BTE), in-the-ear (ITE), in-the-canal (ITC), and completely-in-the-canal (CIC) hearing instruments (Figure 5–6).

Behind-the-ear hearing aids are aptly named because the body of the hearing aid lies on the posterosuperior aspect of the auricle. The receiver is coupled to the ear canal by a piece of acoustic tubing attached to a custom-made earmold. Behind-the-ear hearing aids are extremely flexible because of their design and large size. Any degree of hearing loss, ranging from mild to profound, and almost any configuration can be managed with a BTE aid. A defective unit can be replaced by removing the tubing/earmold assembly and attaching it to a new unit. Changes in the size of the auricle and the ear canal are inexpensively accommodated by making a new earmold. It is not necessary to modify the body of the hearing aid. For this reason, BTE units have been a popular choice for children. In addition, the earmold can be made of material varying in density from hard (lucite) to very soft (silicone). The choice of earmold material can significantly add to the comfort and flexibility of the instrument. A disadvantage of the BTE instrument, however, is the location of the microphone. The ideal location of a microphone would be deep in the ear canal using the natural resonance of the auricle and ear canal. The microphones of BTE instruments are located above the auricle and are oriented anteriorly. The loss of natural cues must be compensated for by the hearing aid. Currently, BTE hearing aids account for less than 20% of annual hearing aid sales in the United States, and the market share continues to decrease.[14,16]

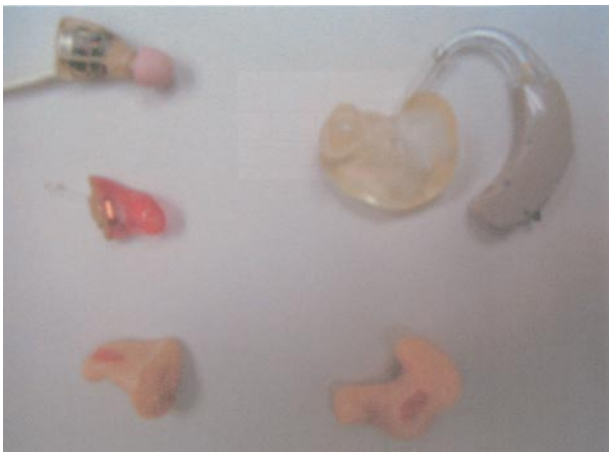


FIGURE 5–6. Examples of major hearing aid styles. The four types of hearing aids are, in a clockwise direction beginning with the top right corner, behind-the-ear (BTE), in-the-ear (ITE), in-the-canal (ITC), completely-in-the-canal (CIC), and disposable ITC.

In-the-ear hearing aids are housed completely within a custom-made earmold designed to fit in the auricle and ear canal. The body of the hearing aid

completely fills the concha and contains most of the electronic components, the battery, and a volume control. The anterosuperior arm of the ITE hearing aid fits into the helix of the auricle inferior to the triangular fossa and functions primarily to keep the instrument in place. The microphone is sometimes located within this portion of the instrument. The medial portion fits partially in the ear canal and contains the receiver. In-the-ear hearing aids are completely self-contained. There is no tubing or earmold assembly. The hearing aid is essentially built into the earmold. Properly fitted ITE hearing aids are appropriate for mild to severe losses of hearing. The location of the microphone, which is close to the level of the ear canal, is an advantage of the ITE style. However, the size of the instrument within the auricle minimizes the use of its natural resonance. Despite these advantages and its cosmetic appeal, it does have a few disadvantages. First, when the instrument must be repaired or replaced, the patient is usually without a hearing aid for a number of days. A BTE loaner may be provided when coupled with a temporary earmold, but this is usually not an acceptable alternative. Additionally, ITE hearing aids may be difficult for some patients to insert and remove, especially for individuals with manual dexterity problems.

In-the-canal hearing aids fit primarily into the ear canal. Only the faceplate of the hearing aid is visible in the concha area. A properly fitted ITC hearing aid may accommodate up to a moderately severe hearing loss. The location of the microphone is at the opening of the ear canal, taking advantage of the acoustic properties of most of the auricle.[16] Also, these devices are inserted deeper than ITE hearing aids, which creates an increase in gain of about 5 dB. Although the size of the ITC instrument is an acoustic advantage, there are a few disadvantages. First, there is less structure holding the hearing aid in place, making it easier to become dislodged and lost. Second, their small size makes them more difficult to insert and remove, hindering use by those with manual dexterity limitations.

Completely-in-the-canal hearing aids are the smallest of the hearing aid styles and therefore the most cosmetically appealing and most popular. These hearing aids are inserted deeply into the ear canal by the patient and are essentially invisible to the casual observer. These instruments are designed to be placed past the second bend of the ear canal, with the medial end of the hearing aid within 5 mm of the tympanic membrane.[17] They have a nylon cord attached to its lateral surface to aid in retrieval of the instruments from the ear canal. Because of its size and microphone location, these devices have a number of advantages.[17,18] These advantages include a reduction in the occlusion effect, ease of use on the telephone, no problems with wind noise on the microphone, an increase in gain by as much as 15 dB, and an enhancement of the acoustic effects of the auricle and ear canal. Unfortunately, these devices also have a number of disadvantages. These disadvantages include problems getting a good acoustic seal, which may lead to feedback problems, increased repair problems owing to cerumen, difficulty inserting and removing the device from the ear canal, increased expense, reduction in the number of advanced features, and an inability to appropriately fit more severe hearing losses.

Hearing Aid Technology

In the past few years, hearing aid technology has grown by leaps and bounds. Hearing aids today have many advanced components and features. We will now review briefly some of the more important developments in hearing aid technology. Unfortunately, a detailed review of all of the available features is not appropriate here. The reader is referred to a text by Sandlin for a more in-depth review.[19]

Disposable and *instant fit* hearing aids are the newest development to emerge on the hearing aid market. These devices are receiving increased attention because of their ability to be fit immediately. The products are available in a wide range of sizes and technology, from BTE to CIC instruments.[20] An advantage of these devices is their significantly reduced price in comparison with custom-made products. Disposable hearing aids have an added benefit in that there is never a need to replace the battery since the entire hearing aid is replaced after 40 days. However, these devices may only be fit on individuals with milder degrees of hearing impairment, rendering those with severe and profound hearing losses ineligible for this technology.

Conventional hearing aids are the least technologically sophisticated of the custom-made hearing aids. These devices are analog hearing aids. An analog hearing aid receives the acoustic signal from the microphone, converts it into an electrical signal, amplifies and filters it according to the individual's

hearing loss, and reconverts it into an acoustic signal that is delivered to the ear canal.[19] In this type of technology, the audiologist selects a matrix for the hearing aid, which determines its overall gain, slope, and maximum output. Relatively few adjustments can be made directly by the hearing aid dispenser or audiologist. These minor changes may be made by screwdriver controls. Typically, these devices provide linear amplification, although some do provide some forms of compression. Linear amplification refers to the types of hearing aids that provide the same amount of gain regardless of the input signal. As the input signal increases in loudness, so will the resulting output of the hearing aid until the hearing aid's maximum output limit is reached. At this point, signal distortion occurs. Compression, on the other hand, helps eliminate the problem of distortion created by linear hearing aids. In this type of hearing aid, the amount of gain provided by the hearing aid depends on the intensity of the incoming signal. Signals with greater intensity receive less amplification than those with lower intensities. This reduction in gain helps decrease distortion as the signal becomes closer to the maximum output provided by the hearing aid. Typically, these instruments have a volume control that may be adjusted by the user.

Programmable hearing aids provide increased flexibility and options for the wearer. Just as in conventional hearing aids, these instruments have analog amplifiers. However, unlike the conventional hearing instruments, these devices can be modified in the audiologist's office via a computer or handheld programmer. This programming allows for increased flexibility in shaping the frequency response, output limitations, compression characteristics, and enhanced features provided by the hearing aid. Additionally, many of these devices offer multiple memories or programs. The use of these multiple programs allows the patient access to different settings that may be designed for various listening environments. These multiple programs may be accessed via a remote control or a switch on the hearing instrument.

Digital hearing aids are the most popular hearing aid technology available in today's market.[19,21] A digital hearing aid processes the acoustic signal differently than an analog instrument. In this arrangement, the acoustic signal is converted to a digital or binary code, which minimizes the possibility for distortion of the signal. The signal is then amplified and converted back to an acoustic signal that is delivered to the ear canal through the receiver. These devices offer a number of advantages, which include increased flexibility of shaping the frequency response of the instrument, feedback suppression capabilities, improved sound quality, decreased battery drain, and less internal circuit noise. As was noted in programmable hearing instruments, these devices may have multiple memories and may be adjusted via a computer or handheld programmer. The memories may also be accessed via a remote control or a switch on the device. Recently, one hearing aid company produced an instrument that changes the program automatically based on the noise levels in the environment.

The major complaint of individuals with sensorineural hearing loss is the inability to hear in background noise. Currently, *directional microphones* provide the individual with hearing impairment the best opportunity for understanding in noise. Simply put, this microphone arrangement allows the wearer to hear the sound originating from the front, which is where the sound source is usually located, and reduces sounds from the rear, which is usually where noise is located.[19] Recent studies have noted 5 to 8 dB of improvement when listening in noise in comparison with standard, omnidirectional microphones.[19,22] This improvement may result in improvement in speech recognition by as much as 60%. Currently, this microphone technology is available in conventional, programmable, and digital hearing aids. However, the performance of the directional microphone is better in the advanced devices. Additionally, this microphone technology is available in BTE instruments through half-shell ITE instruments.

Indications for Amplification

Hearing aids are indicated whenever it can be demonstrated that the patient's ability to communicate will be significantly improved through the use of amplification. A hearing aid is not recommended under the following conditions:

- effective medical treatment can be implemented to restore normal hearing
- hearing aid use would exacerbate the disease or interfere with treatment
- a hearing aid fails to improve the ability of the patient to communicate

A common myth is that amplification is of no benefit in the case of sensorineural hearing loss. Although there is no evidence correlating the use of a hearing aid with restoration or treatment of the impaired auditory system, hearing aids are routinely used to offset the effects of sensory hearing impairment. Sensory hearing loss accounts for 90% of all types of hearing loss. In excess of 95% of all hearing aids are purchased by patients with sensorineural hearing loss. Amplification is an effective means of improving the communicative abilities of individuals with sensory hearing impairment, as attested by hundreds of research articles published in the last 50 years.

Amplification is an extremely effective means of improving the hearing of patients with conductive hearing loss. A hearing aid is clearly not an acceptable alternative to effective medical treatment of a pathologic condition, but it may be used after the course of treatment to offset residual hearing impairment. Residual conductive hearing loss is commonly present following placement of an ossicular prosthesis, for example. In this case, a hearing aid is often an effective means for further improving hearing sensitivity. In contrast, a hearing aid may be of little benefit for patients with retrocochlear (neural or central) auditory disorders. In some cases, amplification may actually exacerbate the effects of the disorder. The use of amplification must be carefully evaluated in all cases of central impairment. However, there are alternatives for these patients. Perhaps the most common approach is the recommendation of an assistive listening device, like an FM system, for improvement of speech understanding by significantly enhancing the speech signal in relation to background noise.

Benefits of Amplification

It is readily apparent that the use of amplification significantly improves the communicative ability of an individual with hearing impairment. However, recent research has revealed additional benefits in psychosocial and functional health measures.[23] These studies have shown that individuals using amplification reported less depressive feelings, richer social relationships, and less anxiety and paranoia. Additionally, and perhaps as a result of improved communicative psychosocial functioning, these individuals noted improvements in their physical health status. Hence, the use of amplification has the potential to improve significantly the quality of life of an individual with hearing impairment.

Monaural versus Binaural Amplification

A great deal of research has been devoted to the following question: if hearing loss is bilateral, are two hearing aids better than one? The intuitive answer is, of course, that two hearing aids would best offset the effects of a hearing loss in both ears. The question persists, however, because many patients choose monaural fittings. This decision is usually based on cosmetic or financial reasons, although some patients claim to hear just as well with one hearing aid as with two, whereas others appear to hear better with a single instrument.

The benefits of binaural amplification have been well documented in the audiologic literature.[9,16] The most commonly cited benefits are as follows:

- improved word identification, particularly in adverse listening conditions
- improved localization of the sound source
- a sense of "balanced hearing"
- the need for less gain
- elimination of the head shadow effect
- increased perception of high-frequency consonants

In addition to these advantages, a number of recent studies have appeared that document the effects of auditory deprivation in unaided ears. The research indicates that word identification scores in the unaided ear decrease over time relative to the scores in the aided ear in a monaural fitting. In contrast, this decrement is not observed in either ear in a binaural fitting. Further, limited recovery in word identification is observed following the subsequent provision of amplification to the deprived ear.

With regard to this evidence, binaural amplification should be recommended unless specifically contraindicated. Valid contraindications to binaural amplification are as follows:

- unilateral hearing loss
- medical complication in one ear
- one ear cannot be aided owing to insufficient residual hearing
- binaural amplification results in diminished word identification

Hearing Aid Selection

The selection of an appropriate hearing aid requires a combination of factors. First, it is essential that the practitioner thoroughly assess the patient's motivation, concerns, communicative goals, and listening needs before selecting a hearing instrument.[24] Issues such as cosmetics, dexterity, and financial constraints may significantly limit the style and technological options. Additionally, one's communicative goals or typical listening environment may dictate the need for a directional microphone or other advanced features. To aid in this endeavor, the use of self-assessment inventories, which will be discussed in the following section, may be of benefit.

To aid in the selection of a hearing aid, multiple researchers have developed prescriptive formulae that are based on audiometric thresholds or suprathreshold information.[25] The goal of these formulae is to provide the listener with the optimum aided speech intelligibility possible. However, these formulae are not expected to be a panacea but are merely intended to provide the dispenser with a valid approximation, a logical starting point. From here, minimal modifications should be required. Currently, there are a number of prescriptive formulae available, but no standard has been adopted.

Assessment of Hearing Aid Outcomes

Any time treatment is rendered, some measure of efficacy of the service should be performed. In the age of managed care, it is exceedingly important to document the success (or failure) of rehabilitation services. Currently, there are a number of means to assess the efficacy of amplification, which include electroacoustic analysis, functional gain measures, real-ear measures, and self-assessment inventories. Each of these procedures will be discussed briefly in the following sections.

Electroacoustic Analysis On arrival of the hearing instruments from the manufacturer and before dispensing the devices to the patient, it is essential that the devices be evaluated to determine if they are functioning appropriately.[3] According to the American Speech-Language Hearing Association, these devices should be assessed via an electroacoustic analysis according to the ANSI S3.22 standard and a listening check.[26] In this first test, the hearing aid is attached to a 2 cm^3 coupler and evaluated in a test box that analyzes the sound pressure present within this cavity. This analysis provides information regarding the gain characteristics of the hearing aid at each frequency, frequency response, maximum output, battery drain, and distortion of the hearing aid. These results can be compared to the specifications of the hearing instrument provided by the manufacturer. If discrepancies occur, the hearing instruments should be returned to the manufacturer for repair or replacement.

Additionally, a listening check of the instruments should be performed. This check will provide information regarding the sound quality, circuit noise, and any intermittent function of the devices. Once again, if problems are noted, the devices should be returned to the manufacturer.

Functional Gain Measures Functional gain measures are designed to assess the realistic benefit patients derive from their hearing instrument by determining behaviorally the amount the hearing aid improves the patient's thresholds of audibility.[9,25] This procedure is conducted in the sound field, using a speech signal and/or warble tones, to determine the difference in the patient's aided and unaided hearing thresholds. Although this test procedure is designed to assess the patient's "real-world" benefit, it is plagued by a number of limitations. These limitations include poor test–retest reliability, limited frequency response information, lack of maximum output information, and a dependence on the patient's behavioral responses.

Real-Ear Measures Owing to the number of limitations presented by functional gain measures, real-ear measures are currently the most accepted means of verifying hearing aid performance.[3,25,27,28] Real-ear measurement systems use a specially designed probe microphone to measure the sound pressure in the ear canal at the eardrum. This probe microphone is inserted along the side of the patient's ear canal into a position located within 6 to 8 mm of the tympanic membrane. A broadband noise, speech composite noise, an International Collegium of Rehabilitative Audiology (ICRA) noise, or tonal series is presented via a loudspeaker that is located approximately 1 meter away from the patient. Measurements are made and graphically displayed with and without the hearing aid in place. Aided test results are usually compared to and adjusted to match the prescribed

target. At a minimum, this testing should be conducted at an input level of 60 to 70 dB SPL, which is typical of average conversational speech. Additionally, testing of the saturation level of the hearing aid should be conducted to ensure that the maximum output of the hearing aid is set below the patient's threshold of uncomfortable loudness. For nonlinear hearing aids, testing at a high input level (80 to 85 dB SPL) and a low input level (50 dB SPL) should also be included. At each of these input levels, patient feedback regarding the volume of the input signal should be solicited. For a comprehensive discussion regarding real-ear technology and a variety of clinical procedures based on this technique, the reader is referred to a text by Mueller and colleagues[25] and a text by Valente.[27]

Self-Assessment Inventories Perhaps the most compelling evidence of the success of a hearing aid fitting comes directly from the patient. After a brief period of hearing aid use, most patients are quite capable of determining whether the prescribed amplification has improved their ability to hear. This is understandable considering that the vast majority of individuals who purchase hearing aids cope with the hearing impairment before deciding to pursue amplification. After very little experience wearing the new hearing aids, most patients are able to precisely describe the listening situations in which their hearing aids are effective or ineffective. To quantify this valuable patient report, a number of self-assessment inventories have been developed for use by audiologists and hearing aid dispensers.[3] These inventories are widely used to evaluate the initial fitting and the effectiveness of subsequent modifications.

Patient Accommodation to Amplification An individual with hearing impairment requires a period of adjustment to become accustomed to wearing hearing aids. The average patient has developed a hearing loss over a period of years before deciding to pursue amplification. During this time, considerable adaptation has occurred in the central auditory system.[29] The sudden introduction to amplified sound initiates compensatory neurophysiologic changes that may take 5 to 6 weeks to develop. In this time period, adjustments to the hearing instruments are often needed to assist the patient in the process of accommodation. Many hearing aid programs encourage patients to use a wearing schedule that gradually extends the length of time that the hearing aids are worn each day. Some programs also restrict hearing aid use to quiet listening environments during the initial period of accommodation with systematic exposure to more demanding conditions as accommodation progresses.

A common problem arises when a patient is initially unhappy with the performance of the hearing aid. According to federal law, a customer may return a hearing aid to the dispenser within 30 days for a refund. Unfortunately, 30 days is often an insufficient period of time to allow for complete accommodation and/or modification of the instruments. Therefore, prospective hearing aid patients should receive proper counseling before, during, and after the initial fitting to avoid making a premature decision.

Pediatric Issues in Amplification Children present a unique set of issues in the selection and fitting of amplification. These issues primarily exist because of the small size of children's ears in comparison with adult ears, the noncompliance of some children, the increased demands in their communication needs, and a lack of complete audiometric data on which to base these decisions. To address these matters, audiologists have developed a set of standards for selecting and fitting hearing aids for young children. As mentioned previously, BTE instruments are the most popular hearing aid style for children. This style allows for maximum flexibility.[16] First, this design allows for an inexpensive means for accommodating changes in the size of the auricle and ear canal as the child grows. While these changes are being made, the child is still able to use the hearing instrument. Second, these devices are more resistant to feedback problems owing to the increased distance between the microphone and the receiver. Furthermore, using an earmold with superthick tubing may also decrease the likelihood of feedback. Third, this style offers options that aid in the retention of the hearing aid in the ear canal, such as wire retention attachments, kiddie tone hooks, and Huggie-Aids. Finally, a BTE hearing instrument offers many additional features that may not be available on smaller styles.

In addition to the style of hearing aid, a number of other hearing aid options should be considered.[16] Every hearing aid dispensed to a child should have direct audio input, telecoil, and microphone/telecoil options available. These features will provide maximum adaptability with assistive listening devices, such

as FM systems. Additionally, tamper-resistant volume controls and battery doors should be considered. For maximum flexibility, digital hearing aids with feedback suppression, programmable features, and multiple memories that offer both omnidirectional and directional microphone technology are added assets. Finally, owing to the likelihood of the hearing instruments becoming lost or damaged in young hands, a loss and damage warranty for the hearing instruments is highly encouraged.

Noting the difficulties in fitting and the communication needs of young children, a prescriptive formula called the Desired Sensation Level (DSL) was developed.[3,30] This formula calculates a prescriptive target for average conversational speech and maximum output of the hearing instrument from a variety of sources, whether it is pure-tone thresholds obtained under telephonic dynamic headphones (standard earphones for audiometric testing) or ABR thresholds. Additionally, if hearing aid verification and adjustments cannot be made using real-ear measures, these adjustments may be made using a 2 cm^3 coupler in the hearing aid test box by using real-ear to coupler differences (RECDs). Real-ear to coupler difference is the dB difference between the level of sound in the ear canal versus the level that is measured in the 2 cm^3 coupler. This measurement is obtained by taking an unaided measurement in the child's ear canal. If this measurement cannot be obtained, the formula provides average RECDs for children of various ages. Hence, with this prescriptive formula, appropriate amplification can be provided with little dependence on the cooperation of the child.

ADVISING THE PARENTS OF YOUNG CHILDREN WITH HEARING IMPAIRMENT

Although the otolaryngologist is principally concerned with the medical or surgical management of aural disease and disorders, it is important that one has an understanding and an appreciation for the various educational options and communication modalities available to the families of children with hearing impairments.[31,32] Since otolaryngologists are one of the professionals responsible for the identification and diagnosis of hearing impairments, they are often the first, if not only, resource that the parent has for information regarding these difficult choices. The debate over the best communication mode and training has raged since the fourteenth century. Unfortunately, there is no one communication modality that is right for every child.[9,15,31,32] This decision must be made on an individual basis, and consideration should be given to issues such as the characteristics of the child (ie, age of diagnosis and accompanying learning disabilities), available community resources, and the commitment of the family to the child and the chosen communication modality. A basic review of the various communication options will be covered in the following paragraphs. Major methods are summarized in Table 5–3. It is important to keep in mind that variations on these methodologies also exist. (See also Chapter 23, "Cochlear Implants.")

Oral Approaches Two major forms of oral English in practice today are the auditory-verbal approach and the auditory-oral approach.[31,32] These two approaches are based on the idea that all children with hearing impairment can realistically attain receptive and expressive language competence regardless of the degree of hearing loss. The auditory-verbal philosophy places an emphasis on the child's residual hearing through the use of amplification with the goal of developing listening skills through natural communication. In this approach, the child is placed in mainstreamed education beginning in preschool. The auditory-oral approach, on the other hand, emphasizes the development of amplified residual hearing and spoken language, using speechreading cues as a supplement to the auditory signal. In this approach, the child is usually enrolled in an oral program until he or she can be appropriately mainstreamed.

Both of these approaches have their advantages and disadvantages. The biggest benefit of these programs is that they provide the child with the greatest access to the hearing world. Geers and Moog reported that children who attended oral education programs attained better speech production, speech perception, and overall spoken language skills than those students who attended total communication programs.[33] Additionally, evidence suggests that these students attain literacy scores twice the national average for children with hearing impairments. Unfortunately, the success of these programs is dependent on early identification of the hearing loss and amplification use, as well as consistent, quality aural habilitation training. Another major drawback to these

TABLE 5–3. Summary of Strengths and Limitations of Various Approaches for Auditory Habilitation of Children with Hearing Impairment and Deafness

Educational Option	*Benefits*	*Limitations*
Auditory-verbal	Greatest access to hearing world Better speech production, speech perception, and spoken language Higher literacy rates Early mainstream education	Dependent on early identification and intervention Isolation from the Deaf community
Auditory-oral	Greatest access to hearing world Better speech production, speech perception, and spoken language Higher literacy rates	Dependent on early identification and intervention Isolation from the Deaf community
Cued speech	Reading and writing skills on par with hearing peers Access to hearing world	Few educational programs available Few transliterators Isolation from the Deaf community
Total communication	Allows the child access to all means of communication	Few programs put this into practice May overstimulate the child
Bilingual-bicultural	Designed to teach language and culture of the Deaf and hearing communities Promotes increased literacy and academic skills	Little information available regarding its success
Signing Exact English	Uses English syntax and grammatical features	Denial of Deaf culture
American Sign Language	Natural mode of communication for the Deaf child Allows membership in the Deaf community and may improve self-confidence	Syntax not conducive to the development of English language

approaches is isolation from the Deaf community owing to a lack of training in sign language.

Manual Approaches In stark contrast to the aforementioned oral approaches are the manual approaches. American Sign Language (ASL) is the common language of the Deaf community in the United States. It is a vast lexicon of hand shapes and motions, or signs, and has its own syntax and grammar.[9,31,32] Additionally, this mode of communication places a heavy emphasis on the facial expressions and body language of the signer. It is a unique language, having no simple translation to the oral English language. Proponents of ASL believe that it is an easier, more natural mode of communication for the child with hearing loss. Additionally, this system allows for membership into the Deaf community, which has the capability of improving the child's self-esteem and confidence. A major limitation of ASL is that its syntax is not conducive to development of the English language, which may hinder spoken language and literacy skill acquisition.

In an attempt to alleviate the difficulties in learning English language through manual communication, educators developed English-based sign systems.[9,34] The most popular form of the English-based sign systems is Signing Exact English (SEE2). Signing Exact English uses much of the same vocabulary as ASL but adds grammatical features and follows English syntax. This system is primarily geared toward preschool and lower elementary schoolchildren to provide them with access to English during the language-learning years. Opponents of this system

believe that it is a denial of Deaf culture by inflicting the standards of the hearing world on the Deaf child.

Bilingual-Bicultural Approach Recently, momentum has been growing for the development of bilingual-bicultural education for children who are Deaf.[35] This approach is designed to educate the child on the mores, customs, and practices, as well as the languages, of the hearing world and the Deaf culture. In these programs, children are taught ASL as their first language, which provides a foundation for which English is later taught. This early access to language is designed to promote increased literacy and academic skills. However, since these programs are relatively new, little information is available regarding its success.

Combination Approaches A combination of the oral and manual approaches is referred to as a combination approach. The most popular combination approach is total communication. Total communication involves the use of one or more modes of communication at any given time in the child's educational program, whether it be manual, oral, auditory only, or written.[9,31,32,36] The design of this approach is to use whatever communication modality is most appropriate for the child at that stage of development or for that given situation, allowing the child access to all means of communication. Despite its promise, this approach has a number of limitations. First, few programs actually put this philosophy in practice owing to the biases of the instructor and the difficulty of combining all of these methods at the same time. Additionally, the use of all of these modalities may overstimulate the child and be a detriment to the development of communication.

A less popular combination approach is cued speech. Cued speech is a visual communication system that employs eight hand shapes placed at four different locations near the mouth.[9,37] These hand shapes are designed to supplement spoken language and speechreading cues since many sounds may be indistinguishable on the lips. The purpose of this approach is to allow the child to see and hear the English language as it is spoken. Wandel found that children who used this system developed reading and writing skills on par with their hearing peers.[38] The limitations of this approach are that few programs provide this type of education and few transliterators (individuals who cue what an instructor says) are available. Additionally, these individuals are unable to communicate with the Deaf community unless ASL is also learned.

A variety of educational methods are available to the individual with hearing impairment. For the vast majority of families, this important educational decision is often made during a period of emotional turmoil. These parents view the identification of a hearing loss as a loss of their dreams of a normal child and may grieve accordingly. Unfortunately, a lack of understanding of the hearing loss, its implications, and remedial interventions can only exacerbate this emotional reaction. Although most parents would have liked to receive this information, they report that few professionals offer the supportive counseling and information they need.[39]

The concept of critical period of development during which the central nervous system exhibits maximal plasticity is central to any discussion about the education of the individual with hearing impairment. Should one maximize language development or the use of residual hearing within a given time period? With the recent implementation of universal newborn hearing screening, it is logical to ask if early identification and management causes a significant improvement in the acquisition of communication skills. Yoshinago-Itano et al provided evidence that children with hearing loss who are identified prior to 6 months of age and receive immediate intervention have greater language development, better receptive and expressive vocabularies, and higher social-emotional aspects of development than infants who are identified later, regardless of the degree of hearing loss or mode of communication.[10,40] Additionally, no significant differences in language development were noted in terms of time of identification for those children identified after 6 months. This finding suggests that language stimulation within the first 6 months of life is critical for neural development.

In summary, each of the currently available educational methods has its proponents and opponents. Whenever appropriate, and with the support of parents and other caregivers, otolaryngologists should encourage and facilitate maximal use of residual hearing and language development in an attempt to help these children reach their full communicative potential. It is essential that the otolaryngologist provide the parent with supportive, unbiased information regarding the benefits and limitations of the aforementioned modalities. No

matter what the family's educational decision, early intervention services should be highly encouraged.

GLOSSARY

ABLB: Alternate binaural loudness balance. A traditional diagnostic auditory procedure for detecting "loudness recruitment" used in differentiating cochlear versus retrocochlear auditory dysfunction in unilateral hearing loss. The task is to balance the sensation of loudness for the better versus the poorer hearing ear. Loudness recruitment is a cochlear auditory sign.

ABR (BAER): Auditory brainstem response. Electrical activity, evoked (stimulated) by brief-duration sounds, arises from the eighth cranial nerve and auditory portions of the brainstem. The ABR is usually recorded from the surface of the scalp and external ear with disk-type electrodes and processed with a fast signal-averaging computer. Auditory brainstem response wave components are labeled with Roman numerals (eg, I, III, V) and described by the latency after the stimulus (in ms) and the amplitude from one peak to the following trough (in microvolts).

AC: Air conduction. Audiometric signals presented via earphones to the ear canal.

air–bone gap: The difference in pure-tone thresholds for air- versus bone-conducted signals. With calibrated audiometers, the normal ear and the sensorineurally impaired ear show no air–bone gap, whereas conductive hearing losses are characterized by an air–bone gap.

ASL: American Sign Language. A manual mode of communication that is commonly used by the Deaf community.

audiologist: A hearing care professional who is educated and trained clinically to measure auditory system function and to manage nonmedically persons with auditory and communicative impairments. Minimal educational requirements for audiologists are a master's degree and certification and/or state licensure.

BC: Bone conduction. Audiometric signals presented via an oscillator to the skull (eg, mastoid bone or forehead).

BCL: Békésy Comfortable Level. A Békésy audiometry procedure conducted at a comfortable loudness level versus a threshold level.

Békésy audiometry: An audiometric procedure performed with a Békésy audiometer for differentiating cochlear versus retrocochlear auditory dysfunction. Békésy audiometry is based on the comparison of responses to pulsed versus continuous tones varied across a wide frequency range. Four patterns of Békésy responses were classified by Jerger.

BOA: Behavioral observation audiometry. A pediatric behavioral audiometry procedure in which motor responses to sounds (eg, eye opening, head turning) are detected by a trained observer.

BTE: Behind-the-ear hearing aid design.

CIC: Completely-in-the-canal hearing aid design.

configuration: Term used to describe the shape or pattern of an audiogram, that is, how hearing loss varies as a function of the audiometric test frequency. There are three main configurations: rising (low-frequency loss), sloping (high-frequency loss), and flat.

CROS: Contralateral routing of signals. A hearing aid configuration in which a microphone is located on the poorer ear and the sounds are transduced and delivered electrically to the normal or mildly impaired ear.

crossover: Sound stimulus presented to one ear (the test ear) travels around the head (by air conduction) or across the head (by bone conduction) to stimulate the other (nontest) ear. See interaural attenuation.

dB HL: A decibel scale referenced to accepted standards for normal hearing (0 dB is average normal hearing for each audiometric test frequency).

dB nHL: A decibel scale used in auditory brainstem response measurement referenced to average behavioral threshold for the click stimulus of a small group of normal-hearing subjects.

dB SL: Sound level or intensity described in reference to an individual patient's behavioral threshold for an audiometric frequency or some other measure of hearing threshold (eg, the speech reception threshold).

dB SPL: A decibel scale referenced to a physical standard for intensity (eg, 0.0002 dynes/cm^2).

dichotic: Simultaneous presentation of a different sound to each ear.

DPOAE: Distortion-product otoacoustic emission.

DPgram (DPOAEgram): A graph of distortion-product otoacoustic emission amplitude in the ear canal (in dB SPL) as a function of the frequencies of the stimulus tones (in Hz).

ECochG: Electrocochleography. Evoked responses originating from the cochlea (the summating potential or SP and the cochlear microphonic or CM) and the eighth cranial nerve (the action potential or AP).

ENG: Electronystagmography. A test of vestibular function in which nystagmus is recorded with electrodes placed near the eyes during stimulation of the vestibular system.
ENoG: Electroneurography. Myogenic activity recorded from the facial muscles, usually in the nasolabial fold, in response to electrical stimulation of the facial nerve as it exits the stylomastoid foramen.
FM system: A device in which the acoustic information received by a remote microphone is transmitted via FM radio waves to a receiver used by the listener.
ITC: In-the-canal hearing aid design.
ITE: In-the-ear hearing aid design.
interaural attenuation: Insulation to the crossover of sound (acoustic or mechanical energy) from one ear to the other provided by the head. Interaural attenuation varies depending on whether the signal is presented by air conduction (interaural attenuation > 40 dB) or bone conduction (interaural attenuation < 10 dB). Insert earphones offer maximum interaural attenuation.
malingering: Feigning or exaggerating a hearing impairment. Also referred to as functional or nonorganic hearing loss.
masking (masker): A controlled background noise presented usually to the nontest ear in an audiometric procedure to prevent a response from the nontest ear (owing to crossover when interaural attenuation is exceeded).
masking dilemma: A problem encountered in audiometric assessment of patients with severe conductive hearing loss. The level of masking noise necessary to overcome the conductive component and adequately mask the nontest ear exceeds interaural attenuation levels. The masking noise may then cross over to the test ear and mask the signal (eg, pure tone or speech). In the masking dilemma, enough masking is too much masking. The masking dilemma can be reduced by the use of insert earphones. The sensory acuity level (SAL) test is also helpful for measuring ear-specific bone-conduction hearing thresholds in patients presenting the masking dilemma.
MCL: Most comfortable level. The intensity level of a sound that is perceived as comfortable.
MLD: Masking level difference. An audiometric procedure that compares a threshold response with masking noise presented in versus out of phase with a pure tone or speech signal. Release from masking is a normal phenomenon reflecting auditory brainstem integrity.
OAE: Otoacoustic emission. Sound generated by energy produced by the outer hair cells in the cochlea and detected with a microphone placed within the external ear canal.
PB: Phonetically balanced. Word lists developed in the late 1940s that contain all of the phonetic elements of general American English speech that occur with the approximate frequency of occurrence in conversational speech.
PI: Performance intensity. A measure of speech recognition or understanding as a function of the intensity level of the speech signal. See rollover.
PTA: Pure-tone average. The arithmetic average of hearing threshold levels for 500, 1,000, and 2,000 Hz or the speech frequency region of the audiogram. The PTA should agree within ± 7 dB of the speech reception threshold (SRT).
RECD: Real-ear to coupler difference. This is the decibel difference between the level of sound in the ear canal versus the level that is measured in the 2 cm^3 coupler.
rollover: A decrease in speech recognition performance (in percent correct) at high signal intensity levels versus lower levels. Rollover is an audiometric sign of retrocochlear auditory dysfunction.
SAL: Sensory acuity level. An audiometric procedure developed by James Jerger (1970) for assessing bone-conduction hearing in patients with serious conductive hearing loss. Air-conduction thresholds are determined without masking and then with masking presented by bone conduction to the forehead. The size of the masked shift in hearing thresholds corresponds to the degree of conductive hearing loss component.
SAT (SDT): Speech awareness threshold (speech detection threshold). The lowest intensity level at which a person can detect the presence of a speech signal. The SAT approximates the best hearing level in the 250 to 8,000 Hz audiometric frequency region.
SEE2: Signing Exact English. A manual form of communication that uses most of the same vocabulary as American Sign Language but adds grammatical features and follows English syntax.
SISI: Short increment intensity index. A clinical procedure developed by Jerger for assessing the ability to detect a 1 dB increase in intensity. High SISI score is consistent with cochlear auditory dysfunction.
SNR: Signal to noise. The signal-to-noise ratio is the difference between the intensity level of a sound or

electrical event and background acoustic or electrophysiologic energy.

SRT: Speech reception level. The lowest intensity level at which a person can accurately identify a speech signal (eg, two-syllable spondee words). See PTA.

SSI: Synthetic sentence identification. A measure of central auditory function that involves identification of syntactically incomplete sentences (a closed set of 10 sentences) presented simultaneously with a competing message (an ongoing story about Davy Crockett).

SSPL: Saturation sound pressure level. A measure of the maximum power output (MPO) of the hearing aid.

SSW: Staggered spondaic word test. A measure of central auditory function developed by Katz that uses spondee words presented dichotically.

Telecoil: An option on a hearing aid that enables it to receive the electromagnetic signals directly from the telephone. This option may be used to reduce the presence of feedback while the wearer is on the telephone.

TEOAEs: Transient evoked otoacoustic emissions. Otoacoustics emissions elicited by brief (click or tone burst) stimuli.

tone decay test: A clinical measure of auditory adaption in which a tone is presented continuously to a hearing-impaired ear until it becomes inaudible. There are numerous versions of tone decay tests. Excessive tone decay is a sign of retrocochlear auditory dysfunction.

TROCA: Tangible reinforcement operant conditioning audiometry. A pediatric behavioral audiometry technique that reinforces a response to auditory signals with food. TROCA is used mainly with mentally retarded or developmentally delayed children.

UCL/LDL: Uncomfortable level or loudness discomfort level. The intensity level of a sound that is perceived as too loud.

VRA: Visual reinforcement audiometry. A pediatric behavioral audiometry procedure that reinforces localization responses to acoustic signals with a visual event (eg, an animal playing).

REFERENCES

1. Hall JW, Mueller HG. Audiologists' desk reference. Vol I. San Diego: Singular Publishing Group; 1997.
2. Carhart R, Jerger JF. Preferred method for clinical determination of pure-tone thresholds. J Speech Hear Disord 1959;24:330–45.
3. Mueller HG, Hall JW. Audiologists' desk reference. Vol II. San Diego: Singular Publishing Group; 1998.
4. Jerger J, Musiek F. Report of the Consensus Conference on the Diagnosis of Auditory Processing Disorders in School-Aged Children. J Am Acad Audiol 2000;11:467–74.
5. Jerger JF. Clinical experience with impedance audiometry. Arch Otolaryngol 1970;92:311–24.
6. Hall JW. Handbook of auditory evoked responses. Needham (MA): Allyn & Bacon; 1992.
7. Pou AM, Hirsch BE, Durrant JD, et al. The efficacy of tympanic electrocochleography in the diagnosis of endolymphatic hydrops. Am J Otol 1996;17:607–11.
8. Hall JW. Handbook of otoacoustic emissions. San Diego: Singular Publishing Group; 2000.
9. Northern JL, Downs M. Hearing disorders in children. 4th ed. Baltimore: Williams & Wilkins; 1991.
10. Yoshinaga-Itano C, Sedley Al, Coulter DK, Mehl AL. Language of early- and later-identified children with hearing loss. Pediatrics 1998;102:1161–71.
11. Hall JW. Screening for and assessment of infant hearing impairment. J Perinatol 2000;20:12–20.
12. Joint Committee on Infant Hearing. Year 2000 position statement: principles and guidelines for early hearing detection and intervention. Am J Audiol 2000;9:9–29.
13. Hall JW. Central auditory processing disorder (CAPD) in Y2K. Hear J 1999;52:35–42.
14. Staab WJ, Lybarger SF. Characteristics and use of hearing aids. In: Katz J, editor. Handbook of clinical audiology. 4th ed. Baltimore: Williams & Wilkins; 1994. p. 657–722.
15. Maxon AB, Brackett D. The hearing impaired child: infancy through high-school years. Boston: Andover Medical Publishers; 1992.
16. Beauchaine KL, Donaghy KF. Amplification selection considerations in the pediatric population. In: Bess FH, Gravel JS, Tharpe AM, editors. Amplification for children with auditory deficits. Nashville: Bill Wilkerson Speech and Hearing Center; 1996. p. 145–60.
17. Chasin M. The acoustic advantages of CIC hearing aids. Hear J 1994;47:13–7.
18. Knowles Electronics. Optimizing the emerging market for completely-in-the-canal instruments. Available at: http://www.knowles-electronics.com/Knelec/html/f6_3.htm (accessed December 17, 2000).
19. Sandlin R. Hearing aid amplification: technical and clinical considerations. San Diego: Singular Publishing Group; 2000.

20. Bray V. Instant fit hearing aids. Available at: http://AudiologyOnline.com/newroot/resources/chat.home/cfm?tid=2/ (accessed December 17, 2000).
21. Schweitzer C. Development of digital hearing aids. Trends Amplific 1997;2:41–78.
22. Valente M, Fabry DA, Potts LG. Recognition of speech in noise with hearing aids using dual microphones. J Am Acad Audiol 1995;6:440–9.
23. Crandell CC. Hearing aids: their effects on functional health status. Hear J 1998;51:2–6.
24. Dempsey JJ. Hearing aid fitting and evaluation. In: Katz J, editor. Handbook of clinical audiology. 4th ed. Baltimore: Williams & Wilkins; 1994. p. 723–35.
25. Mueller G, Hawkins D, Northern J. Probe microphone measurements: hearing aid selection and assessment. San Diego: Singular Publishing Group; 1992.
26. ASHA Ad Hoc Committee on Hearing Aid Selection and Fitting: guidelines for hearing aid fitting for adults. ASHA 1997;40 Suppl 18:123–30.
27. Valente M. Strategies for selecting and verifying hearing aid fittings. 2nd ed. New York: Thieme; 2002.
28. Valente M, Van Vliet MA. The Independent Hearing Aid Fitting Forum (IHAFF) protocol. Trends Amplific 1997;2:6–35.
29. Hurley RM. Onset of auditory deprivation. J Am Acad Audiol 1999;10:529–34.
30. Seewald R. The desired sensation level (DSL) method for hearing aid fitting in infants and children. Phonak Focus 1995;20:3–19.
31. Oticon, Inc, and the Academy of Dispensing Audiologists. Open doors: options in communication and education for children who are deaf or hard of hearing [booklet]. Somerset (NJ): Oticon Inc; 1998.
32. Reamy CE, Brackett D. Communication methodologies: options for families. Otolaryngol Clin North Am 1999;32:1103–16.
33. Geers A, Moog J. Speech perception and production skills of students with impaired hearing from oral and total communication education settings. J Speech Hear Res 1992;35:1384–93.
34. Gustason G. Educating children who are deaf or hard of hearing: English-based sign systems. Reston (VA): The Council for Exceptional Children; 1997. No. EDO-EC-97-3.
35. Baker S, Baker K. Educating children who are deaf or hard of hearing: bilingual-bicultural education. Reston (VA): The Council for Exceptional Children; 1997. No. EDO-EC-96-8.
36. Hawkins L, Brawner J. Educating children who are deaf or hard of hearing: total communication. Reston (VA): The Council for Exceptional Children; 1997. No. EDO-EC-97-6.
37. Caldwell B. Educating children who are deaf or hard of hearing: cued speech. Reston (VA): The Council for Exceptional Children; 1997. No. EDO-EC-97-2.
38. Wandel J. Use of internal speech in reading by hearing and hearing impaired students in oral, total communication, and cued speech programs [dissertation]. New York: Teacher's College, Columbia Univ.; 1989.
39. Kricos P. The counseling process: children and parents. In: Alpiner J, McCarthy P, editors. Rehabilitative audiology: children and adults. Baltimore: Williams & Wilkins; 1993. p. 211–33.
40. Yoshinago-Itano C. Benefits of early intervention for children with hearing loss. Otolaryngol Clin North Am 1999;32:1089–102.

CHAPTER 6

Evaluation of the Vestibular (Balance) System

Neil T. Shepard, PhD, David Solomon, MD, PhD, Michael Ruckenstein, MD, Jeffery Staab, MD

In considering the evaluation of the patient with complaints of vertigo, light-headedness, imbalance, or combinations of these descriptors, one must look beyond just the peripheral and central parts of the vestibular system with its oculomotor connections. The various pathways involved in postural control, only part of which have direct or indirect vestibular input, should be kept in mind during an evaluation. Additionally, significant variations in symptoms and test findings can be generated by migraine disorders[1,2] and/or anxiety disorders,[3–5] yet these are diagnosed primarily by history, requiring a specific line of questioning. These issues, as well as the traditional investigation avenues, will be integrated to present an overall view of the evaluation of the "dizzy" patient. Although the focus of this chapter is the laboratory investigations, it is important to be reminded that the principles and techniques behind the laboratory investigations can be applied in detailed direct office examinations of these patients.[6]

The evaluation of the dizzy patient should be guided by what information is needed to make initial and subsequent management decisions. When various tests are reviewed and correlated with high-level activities of daily living, virtually no significant relationships exist for the chronic dizzy patient. Tests considered extent- and site-of-lesion studies, electronystagmography (ENG),[7] rotational chair,[8] and specific protocols in postural control assessment[9] give results that are unable to be used to predict symptom type, magnitude, or the level of disability of an individual patient. Conversely, patient complaints cannot be used to predict the outcomes of these tests. In a limited manner, more functionally oriented evaluation tools, computerized dynamic posturography (CDP)[10] and dynamic visual acuity testing,[11] provide for some correlation among results, patient symptoms, and functional limitations. Add to the testing specific or general health inventories such as the Dizziness Handicap Inventory[12] and predictive assessment of disability is improved but remains significantly limited. It is hypothesized that the reason for this dichotomy in test results versus functional disability and symptom complaints is the inability of the tests to characterize adequately the status of the central vestibular compensation process.[13–15] It is the exception, not the rule, that vestibular and balance laboratory tests return results that would drive management of the dizzy patient. It would be extremely rare that these studies return a diagnosis. Therefore, the routine use of these tools to determine how to proceed with the management of a dizzy patient is a false line of reasoning and not productive in the majority of the patients. What, then, is the role of the laboratory studies of ENG, rotational chair, and CDP? It is the determination of the extent and site of lesion within the peripheral and central parts of the vestibular system and the functional limitations in static and dynamic postural control (this may relate directly to gait abnormalities). The use of this information is in the confirmation of the suspected site of lesion and diagnosis derived from the patient history and direct physical examination, including aspects of the direct office vestibular evaluation. This does not imply a prioritized order to the testing versus the office visit as with chronic dizzy patients it can be very useful to triage them to at least core laboratory evaluations (to be discussed below) prior to the office visit.

To summarize the above discussion, the following are required elements to management decisions for the chronic dizzy patient: detailed neurotologic history, office vestibular and physical examination, and formal audiometric testing given the inescapable anatomic relationship between the peripheral parts of the auditory and vestibular

systems. The following are considered important; however, they are less likely to drive directly the management in the typical case: laboratory vestibular and balance function studies, neuroradiographic evaluations, and serologic tests. It is important to realize that there will be patients for whom unexpected findings on any one of these latter studies will either alter the complete course of the management or add dimensions to the management not originally considered. But for the majority of patients, the vestibular and balance tests will be confirmatory in nature. Consider the following contrasting examples to illustrate this discussion:

- *Case 1* — Testing is confirmatory. A man, 35 years of age, was seen for complaints of sudden onset, 6 months previously, of vertigo, nausea, and vomiting in a crisis event with continuous symptoms lasting 3 days, steadily showing slow improvement and no accompanying auditory symptoms. The continuous vertigo resolved into head movement–provoked spells of light-headedness with imbalance and occasional vertigo lasting seconds to a minute after a movement. All planes of motion would provoke symptoms. Symptoms had continued to improve but still occurred on an infrequent daily basis. He presented with no neurologic focal complaints, and past medical history was noncontributory. Audiometric evaluation was completely normal bilaterally, as was his contrast magnetic resonance imaging (MRI) study of the head. His detailed vestibular examination both with and without visual fixation was remarkable for a positive head thrust test to the left, right-beating post–head shake nystagmus, and spontaneous right-beating nystagmus with visual fixation removed. The full neurologic and oculomotor components of the examination together with the Hallpike maneuver were normal. The history with the examination was strongly suggestive of uncompensated left peripheral vestibular hypofunction secondary to vestibular neuritis. Laboratory vestibular function testing revealed spontaneous right-beating nystagmus with visual fixation removed, a 76% left reduced vestibular response with normal oculomotor testing, and postural control assessment. In this case, the test was, as is typical in most cases, confirmatory of the clinical suspicions from the history and direct examination. Management decisions made at the time of the office visit, to initiate treatment with vestibular and balance rehabilitation therapy (VBRT) and to discontinue vestibular suppressive medication, were not in any manner altered with the obtaining of the laboratory findings. The vestibular function and balance tests were well justified given the length of symptoms and the fact that the testing has better sensitivity for some oculomotor findings than the direct examination, specifically saccade velocity testing and quantification of smooth pursuit. Although it was not needed in this case, sensitivity to mild peripheral vestibular function asymmetry is also better with the laboratory testing. In this case, the magnitude of the peripheral asymmetry made it detectable by both the direct examination and the caloric irrigation studies.
- *Case 2* — Testing drives management. A 31-year-old man presented with the onset of head motion–provoked vertigo with more or less constant imbalance with standing and walking. He denied any vestibular crisis event or auditory complaints. His symptoms were more concentrated in sagittal plane movement and when rolling left or right from a supine position. These symptoms had been ongoing for several years with intervals when the vertigo was resolved and the imbalance was reduced but not absent. He reported an MRI from several years prior to this evaluation that was normal, with a cervical MRI positive for mild disk abnormalities. Audiologic examination was normal. Other than the development of mild paresthesia of the right hand and arm over the last year, he had no other neurologic complaints, and his past medical history was noncontributory. His direct office examination was remarkable for anterior semicircular canal benign paroxysmal positional vertigo (BPPV). The remainder of the examination was normal. He was treated in the office with a canalith repositioning procedure and referred on for a formal VBRT program. Secondary to the length of time of the symptoms and the complaints of persistent imbalance (although this is a common report with BPPV), vestibular and balance function testing was requested. The laboratory studies showed anterior canal BPPV with no other indications of peripheral vestibular system involvement. Pursuit tracking tests were normal, but saccade testing (these will be discussed below) was positive for mild right internuclear ophthal-

moplegia. Postural control abnormalities were collectively consistent with those seen in demyelinating disorders. Secondary to these findings and his report on a return visit of paresthesia starting in the left foot, a new MRI was obtained that showed multiple hyperintense spots throughout the brainstem region. He was referred to neurology and is being followed with a diagnosis of probable multiple sclerosis with BPPV. Unlike case 1, this case was driven strongly by the results of the vestibular and balance function tests. The test results were able to reveal abnormalities too subtle to be detected in a direct examination. This is clearly the exception to the impact that the testing has on a more routine basis in the decisions regarding management of the dizzy patient.

HISTORY

Before consideration of the actual laboratory tests, some discussion of the neurotologic history is needed. Given the various tools for assessment, the history is the single most important factor in determining the course of management and therefore requires discussion.[14,16–22] The differentiation among the various peripheral vestibular disorders is particularly dependent on historical information and the conclusions that the physician draws from the interview. Most vestibular disorders cannot be distinguished from one another simply by vestibular testing or other diagnostic interventions. Failure to discriminate properly among these disorders on historical grounds may be the source of considerable ongoing distress for the patient and may lead to improper management by the physician. Since subsequent treatment decisions will be based on the clinical diagnosis, it is particularly appropriate to spend additional time during the history to clarify important features. In addition, balance function study results are best interpreted in light of a proper clinical history. The main reason that the patient is seeking medical attention for his or her balance disturbance should be identified. Although little specific diagnostic information may be gained, it is often helpful to hear an account of the patient's perception of his or her illness prior to pursuing more specific questions. One can often gain a sense of the degree of functional disability that the vestibular symptoms have produced. The psychosocial impact of the illness may also become clear in the patient's initial comments. Sometimes patients will volunteer that their symptoms are trivial or have resolved completely, but they simply want to make sure that they have not suffered a stroke or developed a brain tumor. If the patient is not permitted to share this information freely, important aspects of the individual's care may be overlooked.

Once the patient has shared the main concerns raised by his or her condition, the first specific questions should be phrased to gain information regarding the initial onset of symptoms. The characteristics and intensity of the balance disturbance at that time provide useful insight for the differential diagnosis. If the patient can recount specifics surrounding the symptom onset such as date and time of day, along with the activity interrupted by intense vertigo, the probability is high that the patient has suffered a significant peripheral labyrinthine insult. Surprisingly, if the professional does not specifically inquire, patients who are preoccupied with recent symptoms and disability may neglect to report a very profound vestibular crisis that occurred initially. If the onset was more insidious and the patient is unable to provide any account of an initial event, an acute peripheral disorder is less likely. Other important issues to discuss at this stage of the interview include the association of physical trauma, barotrauma, or an intercurrent illness prior to the onset of vertigo. Patients should also be asked about previous remote episodes of vertigo.

It is very important to question the patient regarding the association of a hearing loss or other auditory symptoms with the onset of vertigo. A complete audiometric assessment should be performed early in the evaluation. The presence of an associated sensorineural hearing loss, whether stable, progressive, or fluctuating, is the single strongest incriminating factor in identifying a pathologic labyrinth.[23] The presence of other otologic symptoms such as aural fullness and tinnitus may also be helpful in lateralization. Head trauma associated with dysequilibrium and fluctuating sensorineural hearing loss or a history of previous otologic surgery or familial hearing loss is also important.

Most patients with acute vertigo will improve with supportive, expectant management. When this is not the case, they may present for evaluation after several months or years of dizziness. The patient should be asked to describe the progression of symptoms over time, along with the nature and duration of typical spells. Specifically, one wishes to know if

the spells are continuous or occur in discrete episodes. If the symptoms are episodic, it is extremely important to distinguish whether they are spontaneous or motion provoked. If the symptoms are brief and predictably produced by head movements or body position changes, the patient most likely has a stable lesion but has not yet completed central nervous system compensation. Those who describe these symptoms sometimes also note a chronic underlying sense of dysequilibrium or light-headedness. The chronic symptoms may be quite troublesome, but any intense vertigo should be primarily motion provoked. These patients are suitable candidates for vestibular rehabilitation. It is important to point out that historical information is essential in deciding who might benefit from rehabilitation therapy.

If the episodic spells described by the patient are longer periods of intense vertigo that occur spontaneously and without warning, this is probably progressive or unstable peripheral dysfunction. One must also suspect a progressive lesion if the vertigo is accompanied by fluctuating or progressive sensorineural hearing loss. Such patients are managed with medical therapy, and if this fails, they constitute the best candidates for surgical intervention. Such patients are not candidates for vestibular rehabilitation, except as an adjunctive modality.

As discussed above in generalities, two of the critical elements in developing a differential diagnosis involve the determination of spontaneity versus visual or head-movement provocation together with the temporal characteristics of the symptoms. These two features play such a major role in the development of the working diagnosis that a more detailed discussion is appropriate. Table 6–1, which is not intended to be exhaustive, suggests the use of these two aspects of the patient history and the possible sources to consider in the diagnosis. To differentiate among the diagnostic options further, the other historical information and auditory function can be immediately useful.

Additional history regarding current or prior use of medications should be elicited. Many patients are under the mistaken impression that vestibular suppressants will prevent spells of vertigo and take them habitually. Because oversedation from these centrally acting drugs may retard central nervous system compensation for vestibular lesions, one should consider tapering or discontinuing these medications whenever possible. Medications that must be continued should be directed toward particular symptoms that specifically interfere with the patient's recovery process.

Other psychosocial aspects can be important in understanding the patient's situation. Complicating features of anxiety, depression, or excessive dependence on psychotropic medications should be identified. It is desirable to understand the degree of functional disability produced by the patients' vestibular complaints, especially with respect to their professional and favorite social activities. The stability and commitment of their psychological support system should also be evaluated.

LABORATORY TESTS OF VESTIBULAR AND BALANCE FUNCTION

The principal tests that are found in most dedicated balance centers are ENG, rotational chair, and CDP. Other evaluation tools are available for specific tasks. These include head on body studies (autorotation tests),[24] special protocols that use eye movement recordings that can quantify linear horizontal and vertical as well as torsional activity,[25] dynamic visual acuity,[11] and protocols to assess the otolith organs specifically.[26] Although these additional studies can be useful, the primary discussion will be confined to the principal tools of ENG, rotary chair, and CDP.

Assuming that all of the studies are available, it is not necessary for each patient to have all aspects of the major evaluative tools as a selection of the tests may be adequate. One suggested strategy for determining what tests are needed is based on a core set of studies that would be appropriate for all patients, and the use of those studies together with the neurotologic history determines when the other tests are needed. This means that the selection of the tests occurs as the testing is being performed and is not predetermined at the time of referral. When consideration is given to the discussion above regarding the utility for the testing in the management decision process, the better approach is to let the testing determine the studies needed given the strengths of the individual tools that will be discussed below. A suggested core would be the ENG with full oculomotor studies (by computerized system), full history, and screening evaluation of postural control. The basic site-of-lesion investigation and first approximation of the functional assessment of balance can be obtained in this manner.[14] The rationales for proceeding with total-body rotational chair testing are as follows:

TABLE 6–1. Suggested Differential Diagnostic Entries Based on Temporal Course and Symptom Characteristics

Temporal Course of Symptoms	*"Dizziness" Characteristics*	*Auditory Characteristics*	*Differential to Consider*
Episodic (s)	HM-HP provoked	Normal	Benign paroxysmal positional vertigo, uncompensated stable peripheral lesion
Episodic (s–min < 60)	Spontaneous	Normal or mild fluctuations	Migraine, transient ischemic attacks, anxiety disorders
			Meniere's disease (spells over 20 min)
Episodic (min–h < 24)	HM-HP provoked	Normal	Uncompensated stable peripheral lesion, migraine
Episodic (min–h < 24)	Spontaneous	Fluctuant + progressive	Labyrinthine disorders, eg, Meniere's disease; autoimmune disorders
Episodic (min–h < 24)	Spontaneous	Normal or mild fluctuations	Migraine, anxiety disorders, cardiovascular disorders
Days	Spontaneous resolving in 1 to 3 d to HM-HP provoked	Normal	Vestibular neuritis, vascular events
Days	Spontaneous resolving in 1 to 3 d to HM-HP provoked	Sudden loss at time of vertigo onset	Labyrinthitis, posteroinferior cerebellar artery or anteroinferior cerebellar artery distribution strokes
Days typically < 7	Spontaneous or HM-HP provoked relatively constant overall	Normal or mild fluctuation	Migraine
Continuous	May or may not be exacerbated by HM-HP	Normal	Central vestibular disorders, anxiety disorders, nonvestibular disorders such as sensory or motor neuropathy of lower limbs

HM = head movement; HP = head position.

- When the ENG is normal, and oculomotor results are either normal or observed abnormalities would not invalidate rotational chair results. Chair testing is used here to expand the investigation of peripheral system involvement and compensation status.
- When the ENG suggests a well-compensated status (no spontaneous or positional nystagmus), despite the presence of a clinically significant unilateral caloric weakness and ongoing symptom complaints. Chair testing is used here to expand

the investigation of compensation in a patient with a known lesion site and complaints suggesting poor compensation.
- When the caloric irrigations produce nystagmus below 10 degrees/second bilaterally, when caloric irrigations cannot be performed, or when results in the two ears may not be compared reliably owing to anatomic variability. Chair testing is used in these cases to verify and define the extent of a bilateral weakness or to investigate further the relative responsiveness of the peripheral vestibular apparatus bilaterally when caloric studies are unreliable or unavailable.
- When a baseline is needed to follow the natural history of the patient's disorder (such as possible early Meniere's disease) or for assessing the effectiveness of a particular treatment, such as that of chemical ablation of one or both vestibular end-organs.

In a like manner, decisions for use of full postural control studies can be made based on the history and the screening tools for postural control assessment. Therefore, based on the results of the Clinical Test of Sensory Integration and Balance (CTSIB)[27,28] and the patient's full neurotologic history, CDP would be called for under the following conditions:

- If the CTSIB was abnormal with a fall reaction, significantly increased sway, or sudden change of strategy for maintaining stance during a trial. Based on the comparison between CTSIB and CDP, we know that the patient who can perform CTSIB in the normal range is very unlikely to fail the sensory organization portion of CDP. On the other hand, CDP is needed to delineate further postural control difficulties when abnormal performance is noted on the CTSIB.
- If the patient has a major complaint involving unsteadiness when standing or walking, independent of the CTSIB results. Dynamic posturography, especially motor control testing (MCT), is used in this case to investigate further postural control. This patient would also be a candidate for postural evoked response testing, a third special protocol that can be used with CDP in conjunction with lower limb electromyography (EMG) recording, described below.
- If the patient had a history of known pathology involving the postural control pathways that may influence the patient's overall performance, even though unsteadiness is not a major complaint.

Electronystamography

Traditional ENG, using electro-oculography for eye movement recordings, is a process that estimates the position of the eye as a function of time indirectly. The estimates are reliable whether recorded with eyes open or closed and in a darkened or well-lit environment. Changes in eye position are indicated by the polarity of the corneal-retinal potential (dipole) relative to each electrode placed near the eyes (Figure 6–1, A). These electrodes are typically placed at each lateral canthus and above and below at least one eye with a common electrode on the forehead (Figure 6–1, B). Since the primary purpose of the vestibular apparatus is to control eye movements, the movements of the eyes may be used to examine the activity of the vestibular end-organs and their central vestibulo-ocular pathways. More recently, the technique of infrared video-oculography has begun to replace the standard electro-oculography. This technique has the advantages of direct estimate of the eye position as a function of time and reduction in electrical artifacts. A typical horizontal ENG trace of right-beating nystagmus (could be from either electrodes or video techniques) is shown in Figure 6–2. In this figure, a rapid upward deflection of the trace (fast or saccade component of the nystagmus) is seen followed in each case by a slower downward drift of the trace (slow component of the nystagmus). The convention in the recordings is that the upward trace deflections represent rightward (horizontal trace) or upward (vertical trace) eye movements. Conversely, downward deflections of the trace represent leftward and downward eye movements. Calculation of the principal parameter used in the analysis of nystagmus, slow-component eye velocity, is illustrated in Figure 6–3. The ENG evaluation is a series of subtests performed to assess peripheral and central portions of the vestibular system. It is important to understand that assessment of the peripheral part of the vestibular system with ENG is significantly limited, typically reflecting function of the horizontal semicircular canal with restricted information from vertical canals and otolith organs. With the use of computerized ENG systems, which afford significant visual stimulus control and quantitative analysis, evaluation of the central vestibulo-ocular pathways can be quite thorough.

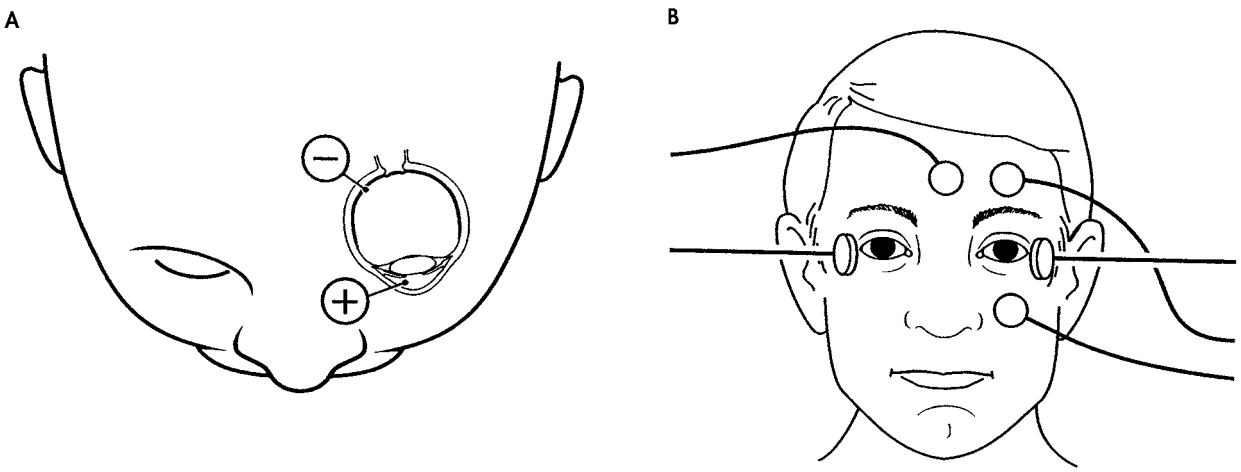


FIGURE 6–1. A, Shown schematically is the corneal-retinal potential that forms the basis for electro-oculography as an indirect recording technique for eye movement position as a function of time. *B*, Typical electrode montage (arrangement) for monitoring eye position as a function of time in both the lateral and vertical dimensions.

The ENG consists of the following groups of subtests: oculomotor evaluation, typically with smooth pursuit tracking, saccade analysis, gaze fixation, and optokinetic stimulation; spontaneous nystagmus; rapid positioning (Hallpike-Dix maneuver); positional nystagmus; and caloric irrigations. The slow-component eye velocity of the nystagmus (see Figure 6–3) is the measurement of interest as it reflects the portion of the nystagmus that is generated by the vestibulo-oculomotor reflex (VOR) (with the fast component generated from the pons area of the brainstem in response to the position of the eye in the orbit). Although traditionally part of the ENG, the Hallpike maneuver should be analyzed by direct examination, not requiring quantification with recording, and should not be analyzed by the typical two-dimensional (horizontal and vertical) recorded eye movements from an ENG. The oculomotor tests are quantified according to the eye movements generated during the task and analyzed by comparison to normative data when a computerized system is used.

Oculomotor Tests Just as the eyes serve as the window for investigating the function of the peripheral part of the vestibular system, they provide a

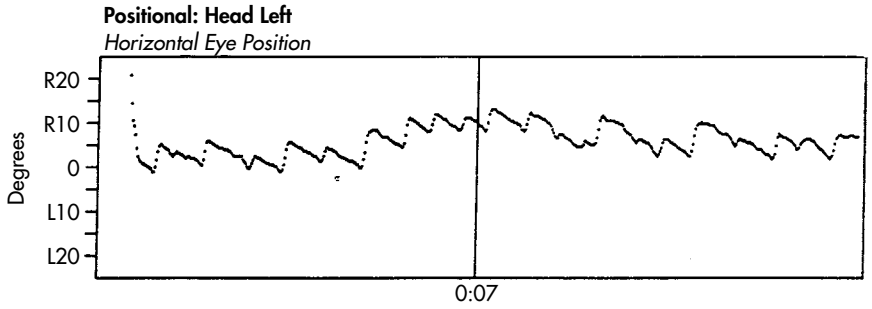


FIGURE 6–2. Right-beating horizontal nystagmus is demonstrated from a patient's eye movement recording with the patient's eyes closed in the head turned left from a supine position. The vertical axis is in degrees of movement, with the horizontal axis in time going from left to right, showing a total of 14 seconds.

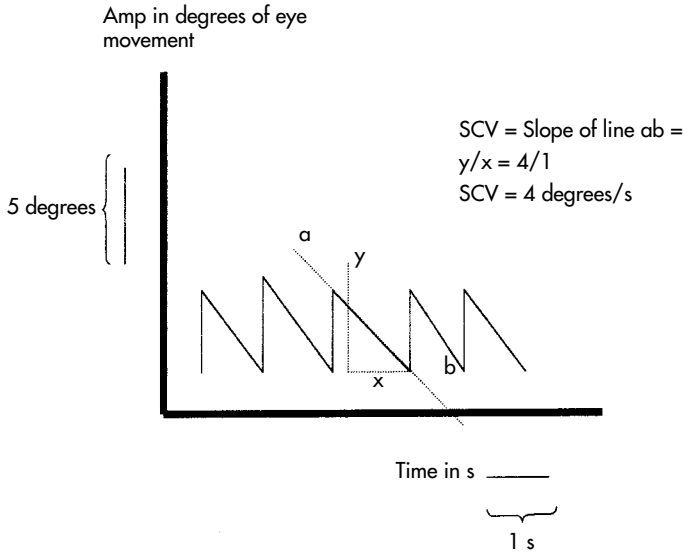


FIGURE 6–3. Right-beating nystagmus is illustrated along with the calculation of slow-component eye velocity, the slope of the line ab. As discussed in the text, upward trace deflections represent rightward eye movement, with downward indicating leftward eye movement. Amp = amplitude; SCV = slow-component eye velocity. Reproduced with permission from Shepard NT. Evaluation and management of balance system disorders. In: Katz J, Medwetsky L, Burkhard B, editors. Handbook of audiology. 5th ed. Baltimore: Williams & Wilkins; 2000.

means to investigate the oculomotor pathways in the brainstem and cerebellum that are required for the function of the VOR.

A variety of testing paradigms can assist in identifying abnormalities in the central oculomotor control systems that may produce the patient's complaints. The smooth pursuit test evaluates the ability to track a moving object with smooth eye movements and head still. Smooth pursuit is the most sensitive of the oculomotor tests but provides poor lesion site localization within the multiple pathways involved in pursuit generation. Abnormalities with pursuit are typically taken as an indication of possible vestibulo-cerebellar region involvement, the final common portion of the multiple pathways for pursuit production.[18] Saccade testing evaluates rapid movement of the eye used to place an object of interest on the most sensitive portion of the retina, the fovea. Saccade testing is not as sensitive as pursuit but when tested with different paradigms can provide for differentiating information concerning brainstem versus posterior cerebellar vermis involvement. Suggestions for possible frontal or parietal lobe involvement can also be obtained from saccade testing. Gaze fixation evaluates the ability to hold the eyes in a fixed direction of gaze without drifting off the target, typically straight ahead and to the right, left, up, and down. Gaze fixation provides general suggestions of brainstem/cerebellar involvement in most instances of abnormal results. There are specific abnormalities of fixation of gaze that can be indicative of specific cerebellar degenerative disorders (saccade intrusions). Optokinetic stimulation measures jerk nystagmus eye movements created by repeated objects moving across the subject's visual field and filling at least 80% of the visual field. Optokinetic stimulation is the least sensitive, probably owing to the combination of both smooth pursuit and saccade systems, allowing the optokinetic nystagmus to be generated by a combination of foveal and peripheral retinal stimulation. At present, it serves best as a cross-check with significant abnormalities seen during pursuit or saccade testing. Illustrated in Figures 6–4 and 6–5 are typical recordings and analysis of smooth pursuit and reactionary saccade testing.

Compared with gaze fixation and saccade testing, smooth pursuit is significantly more sensitive to the presence of pathology. Sensitivity performance of these tests was evaluated in a group of 134 patients diagnosed with multisystem atrophy, olivopontocerebellar atrophy, Friedreich's ataxia, or multiple sclerosis.[14] All had clearly defined brainstem/cerebellar lesions on MRI. Eighty-seven percent had evidence for central vestibulo-ocular pathway dysfunction on oculomotor testing, as indicated by abnormal performance on any single or combined test of gaze fixation, saccades, and smooth pursuit measures. Smooth pursuit measures alone identified central system involvement in 73% of these cases, independent of results on saccade and fixation testing, compared with only 40% identified by saccade

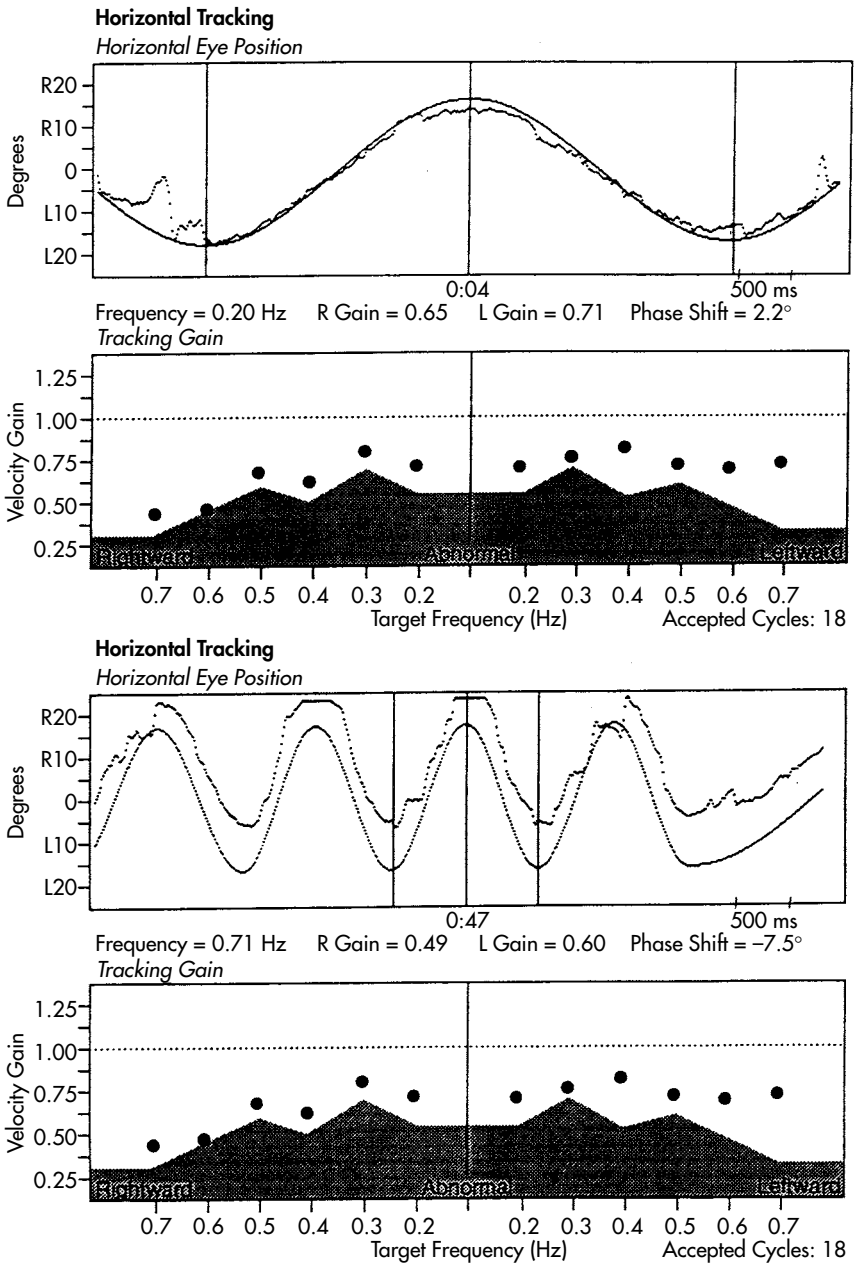


FIGURE 6–4. An example of smooth pursuit eye movement recording with the target at two of the multiple frequencies tested. Shown in both top and bottom panel sets is a plot of horizontal eye position as a function of time (500 ms time mark shown) for sinusoidal tracking (*dotted trace*). In the same plot (*smooth line*), the target position in degrees as a function of time is given. The panel below the eye and target position plots gives the value of velocity gain as a function of frequency of target movement. The *shaded region* represents abnormal performance. The top panel set is for a target frequency of 0.2 Hz, and the bottom panel set is for 0.71 Hz. Reproduced with permission from Shepard NT and Telian SA.[14]

testing alone and 47% for fixation testing. These tools also can suggest central nervous system involvement in patients without obvious structural abnormalities identified by MRI. As it is possible to have abnormal central nervous system function without structural changes on neuroradiographic studies, it is difficult to provide overall true sensitivity figures. Statistically, as normal values are typically based on two standard deviations from the mean performance, a small percentage of those with suspected central involvement will be false positives. An estimate of 86% specificity (14% false positives) for the oculomotor studies is derived from a group of patients suspected of having pure peripheral lesions.[14] Given the advanced nature of the disease states of the groups of patients used for the sensitivity and specificity figures, it is highly likely that both figures are overestimates of performance with actual sensitivity and specificity figures in the 75 to 80% range.

In a review conducted by Shepard and Telian[14] of 2,266 patients seen over a 32-month period, 14% had abnormalities suggesting brainstem and/or cerebellar involvement without any indications of involvement of the peripheral part of the system.

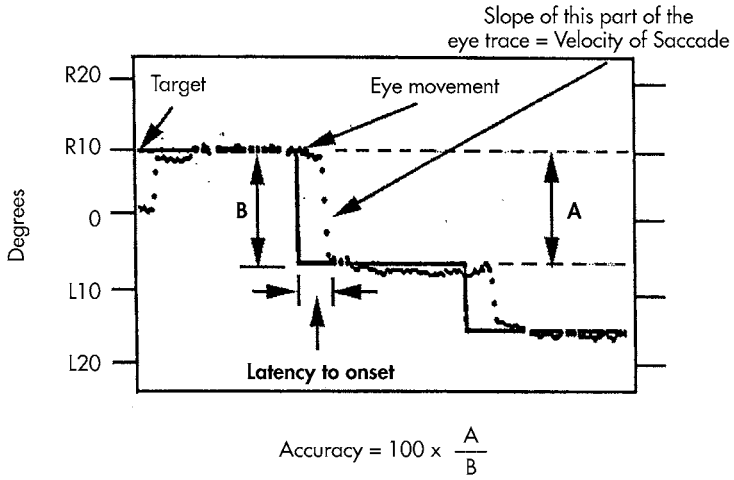


FIGURE 6–5. Schematic of eye and target position in degrees as a function of time demonstrating the calculations of velocity, latency, and accuracy of the saccade eye movement for analysis. The three parameters of velocity, latency, and accuracy are then used to characterize the patient's performance. Reproduced with permission from Shepard NT and Telian SA.[14]

The percentage of those suspected of involvement of the central part of the system increased to 21% when combinations of peripheral and central abnormalities were included. Therefore, since exact sensitivity and specificity performance figures cannot be determined, it seems best to employ test protocols that may reduce the false-positive rate (increase specificity) without decreasing test sensitivity. The three crucial principles to assist in these goals are (1) the use of computerized, quantitative analysis; (2) application of age-appropriate normative data comparisons; and (3) repetition of initially abnormal test measures to obtain optimal performance.

Spontaneous Nystagmus This test is performed with the patient sitting, head straight and eyes closed. The purpose is to record eye movements when visual fixation is removed, without any provocative head movements or positions. Jerk nystagmus is the principal abnormality of interest in most situations. Other forms of abnormal eye movements, such as pendular nystagmus (sinusoidal horizontal repeating eye movements with no distinguishable slow and fast components), may be seen. Clinically significant, direction-fixed nystagmus is interpreted to indicate possible pathology within the peripheral part of the vestibular system if the oculomotor evaluation is normal (Figure 6–6, A).

Head-Shake Nystagmus The head-shake protocol is commonly associated with the direct office examination yet can be usefully incorporated as part of the ENG.[6] As with the office examination, reciprocal horizontal head movements at about 2 to 3 Hz for 10 seconds are presented with fixation removed. When the head is brought to a stop, recording of eye movements is continued long enough to capture nystagmus that had been provoked or no nystagmus if the test is negative. The presence of three or more horizontal nystagmus beats is suggestive of a peripheral asymmetry. In most cases, the direction of the fast component of the nystagmus is away from a paretic lesion; however, this is not always the case, and the nystagmus may beat toward the side with a lesion. Vertical nystagmus resulting from the horizontal head shake is a positive sign for brainstem and/or cerebellar involvement and is referred to as cross-coupling nystagmus. Although the sensitivity for identification of peripheral involvement is reported to be only fair, the ease of the test and its short time for execution make it reasonable to add to the ENG battery.[29]

Hyperventilation Nystagmus This is another of the office examination tools that can be easily and effectively used during an ENG protocol. The use of hyperventilation testing has a dual purpose. The first is identification of nystagmus suggestive of eighth cranial nerve involvement on the side to which the fast component of the nystagmus is directed. The nystagmus results from myelin being absent from a section of the nerve secondary to a demyelination process from disease or a mass lesion.[30] The second use is to identify an individual more likely to be experiencing an anxiety disorder. This is recognized by the production of symptoms of light-headed-

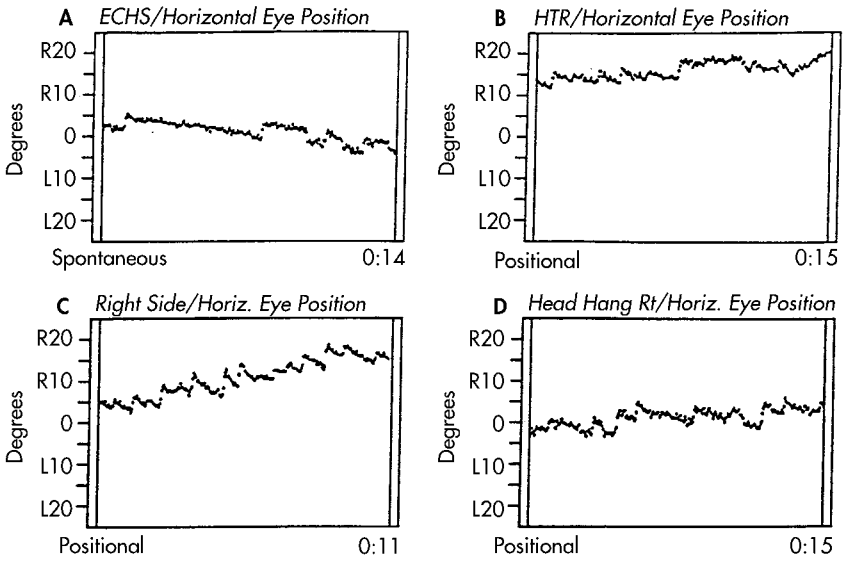


FIGURE 6–6. *A* through *D* plot horizontal eye position in degrees as a function of time, each showing right-beating nystagmus. The time in the lower right corner of each panel is the total elapsed time in seconds for that recording, not the total time shown. Each panel represents 7 seconds of tracing. *A*, ECHS = eyes closed head straight, spontaneous. *B*, HTR = eyes closed head turned right, sitting position. *C*, Eyes closed, lying on the right side (right decubitus position). *D*, Eyes closed, extension of the neck (by approximately 30 degrees) with the head turned to the right. These are 4 of a total of 11 positions tested on this patient. Reproduced with permission from Shepard NT and Telian SA.[14]

ness/imbalance within 30 seconds of starting the hyperventilation, with no nystagmus visualized. Hyperventilation is more likely to provoke dizziness, light-headedness, autonomic symptoms, and acute anxiety attacks in patients with certain anxiety disorders than in the general population.[31] It also may provoke light-headedness and autonomic arousal without significant anxiety in patients with autonomic dysfunction (eg, hyperventilation syndrome).[32] Anxiety disorders or dysautonomia should be suspected in patients who experience a marked reproduction of their symptoms, without nystagmus, during hyperventilation.

Hallpike Maneuver The Hallpike maneuver is a well-known office procedure used primarily to elicit evidence for BPPV. Since this is typically a problem with a single semicircular canal at a time, the description of the nystagmus is critical to identification of the canal involved. Each canal has an individual and unique eye movement signature that can be used to recognize which canal is causing this disorder.[33] In the variants of BPPV, only the horizontal canal variant does not have torsional eye movement; all others involve torsional movements and therefore cannot be recorded with standard electrooculography techniques or typical standard video systems as both are two-dimensional in the printed output.

The Hallpike maneuver is also typically part of the standard ENG protocol (Figure 6–7). Classically, positive Hallpike responses produce a torsional nystagmus (pure rotational movement of the eye about an axis through the center of the globe going from anterior to posterior when that axis is oriented parallel to the axis perpendicular to the plane of the canal involved). The fast component is directed toward the involved ear (left torsional for left ear and right torsional for right ear, from the patient's perspective, relative to the superior pole of the eye). To detect this action adequately, the movements must be viewed by the examiner (directly or with video equipment), not reviewed after the fact by two-dimensional recordings from standard surface electrode or standard video recording techniques.

Positional Nystagmus Positional nystagmus is the most common abnormality noted with ENG evaluation. In this study, the patient is moved slowly into stationary positions. The eye movements are moni-

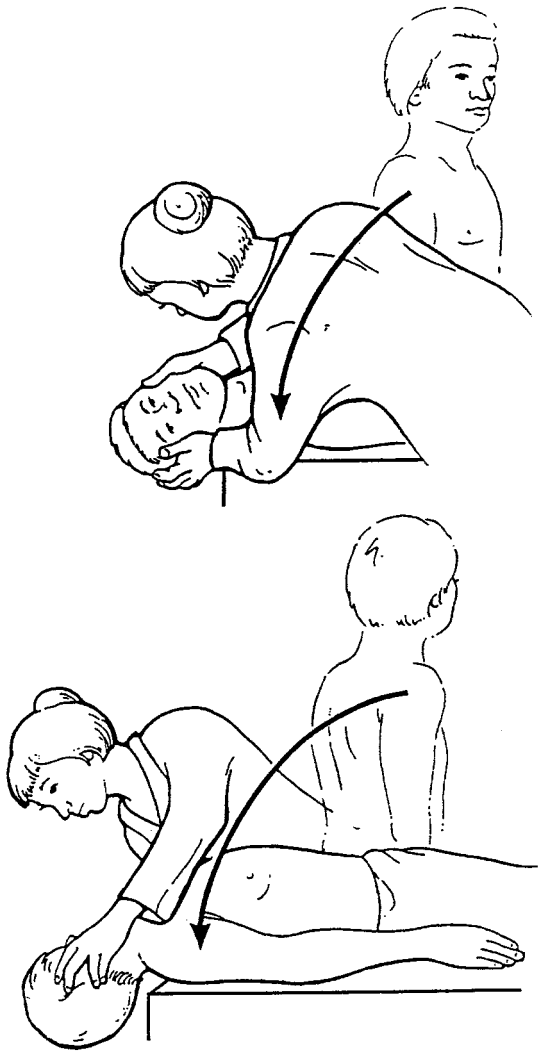


FIGURE 6–7. Illustrations of the technique for the Hallpike maneuver, for right side (*top*) and left side (*bottom*). Note that the patient's eyes are open and fixating on the examiner. Reproduced with permission from Shepard NT and Telian SA.[14]

tored as in the spontaneous nystagmus test and should be done without visual fixation (see Figure 6–6, B to D). The more common positions include sitting head turned right; sitting, head turned left; supine; supine, head turned left; supine, head turned right; right decubitus (right side); left decubitus (left side); and preirrigation position (head and shoulders elevated by 20 to 30 degrees up from the horizontal plane). In cases for which no cervical region injuries or active pathologies are reported, use of head hanging (neck extended by up to 20 degrees) straight, right, and left adds three additional positions for testing prior to preirrigation position. The purpose of this test is to investigate the effect of different head positions within the gravitational field. Positional nystagmus is typically classified by the direction of the fast component of the nystagmus but measured by the velocity of the slow component. It may be either direction fixed (always beating in the same direction) or direction changing (both right- and left-beating nystagmus observed during the examination). Direction-changing nystagmus may be subclassified, when appropriate, into geotropic (toward the pull of gravity, toward the underneath ear) or ageotropic (away from the pull of gravity, away from the underneath ear).

The clinical interpretation of positional nystagmus is taken as indicative of involvement of the peripheral part of the system (with two notable exceptions) as long as the oculomotor studies, done in a thorough manner, requiring a computerized system, are normal. The exceptions to this situation are (1) when the direction-changing nystagmus is observed while the patient remains in one head position, that is, without a change in gravitational orientation, and (2) if persistent vertical nystagmus is noted with no horizontal component. These are typically interpreted as indicative of central pathway involvement (usually low posterior fossa), independent of the oculomotor results, unless an alternative explanation is apparent. If spontaneous, direction-fixed nystagmus is present and no significant change in that nystagmus (average slow-component velocity) is noted, then the nystagmus in the positions is taken as a reflection of the spontaneous findings and not considered as positional. This situation would suggest little or no influence of the otolith organs in the modulation or generation of the nystagmus but would not otherwise alter the general interpretation of involvement of the peripheral part of the system, assuming normal oculomotor studies.

Since positional nystagmus can be produced with central or peripheral lesions, the presence of abnormal oculomotor studies (the most sensitive test indication of central nervous system involvement) precludes the use of positional nystagmus as an indicator of a lesion of the peripheral part of the system. Clinical significance of the positional nystagmus is typically determined by a combination of the number of positions in which it is present and the velocity range of the slow component of the nystagmus. One suggested set of significance criteria based on

studies of the presence of positional nystagmus in a normal population is as follows: slow-component eye velocity > 5 degrees/second, slow-component eye velocity < 6 degrees/second with persistent nystagmus in 4 or more of the 8 to 11 positions, and slow-component eye velocity < 6 degrees/second with sporadic nystagmus in all positions tested and direction changing within a given head position.[34]

Two other conditions that can produce positional nystagmus without oculomotor abnormalities must be considered before involvement of the peripheral part of the system is suggested. Migraine headache conditions can produce these findings, in addition to mild caloric asymmetries.[35,36] No particular patterns of positional nystagmus are indicative of migraine. Since migraines and dizziness related to migraine are diagnoses of exclusion, it is important to keep this in mind if the total presentation of the patient is not convincing for involvement of the peripheral part of the system. Somewhat less well documented in producing positional eye movement abnormalities are anxiety disorders. The most commonly recognized eye movements associated with anxiety, especially related to the testing process, causing symptoms are macrosquare wave jerks (square waves of a 5- to 15-degree subtended arc with normal intersaccade intervals).[18] This type of intrusion saccade is noted only with fixation removed. Classic positional jerk nystagmus of a direction-fixed or direction-changing nature can also be provoked by anxiety disorders.[31]

Caloric Irrigations The caloric test is the study that is most likely to indicate the side of a peripheral lesion with objective, repeatable eye movement data. The stimulus employed is nonphysiologic compared to the normal function of the system during head motion, when one side is stimulated and the other is simultaneously inhibited. Nevertheless, the caloric test is the only portion of the test battery that provides a measure of unilateral labyrinthine function. There are three primary delivery methods for caloric irrigations: closed-loop water (circulates in a thin latex balloon that expands in the external auditory canal), open-loop water (runs into the external auditory canal and drains out), and air irrigation (Figure 6–8). During any of the methods, the fluid or air is set at temperatures above or below that of the body; typically, 44°C for the warm and 30°C for the cool are used for the open-loop water systems. All are reasonably reliable when the tympanic membrane is intact. When a tympanic membrane perforation or short ventilation tube is present, the closed-loop water irrigation method is preferable.

In both terrestrial and weightless environments, there appears to be at least two mechanisms operating to produce the caloric VOR response to the temperature changes. The one that seems to predominate in routine testing involves gravity and the density changes that occur in the endolymph of the horizontal canal when it is heated or cooled (density decreased or increased, respectively). The head is positioned so that the horizontal canal is oriented

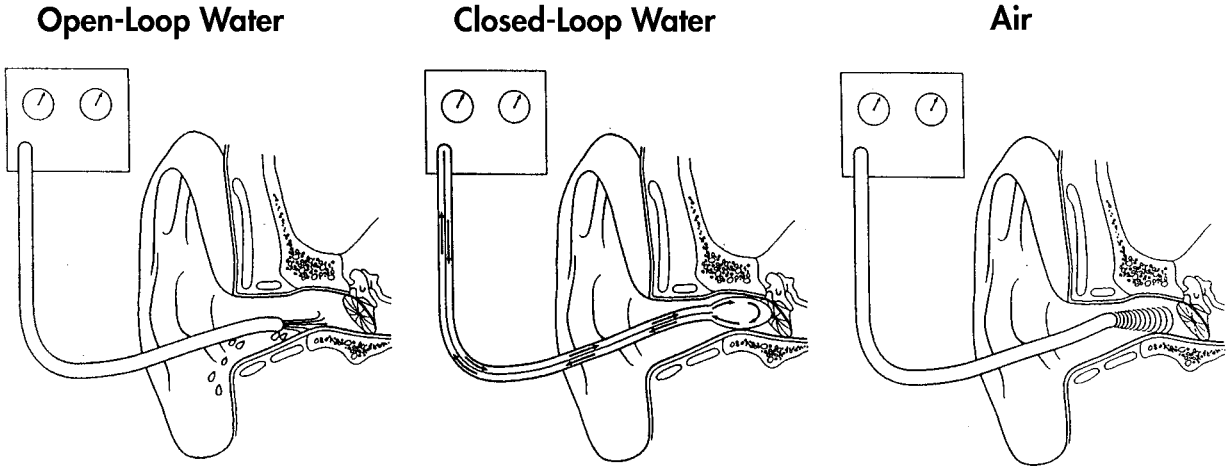


FIGURE 6–8. Three techniques for caloric irrigations are shown. Reproduced with permission from Shepard NT and Telian SA.[14]

parallel to the gravitational vector, with the nose of the patient upward and the head tilted 20 to 30 degrees upward from the horizontal plane. During a warm irrigation, the less dense fluid attempts to rise upward. This produces a deviation of the cupula toward the utricle owing to the pressure differential across the cupula, causing stimulation of the eighth nerve. The reverse action occurs for the more dense area of cooled fluid, causing inhibition. This results in the well-known mnemonic "COWS," which refers to the direction of the fast component of the nystagmus: *C*old *O*pposite, *W*arm *S*ame (relative to the side of irrigation).

The traditional interpretation of caloric stimulation uses a relative comparison of maximum, average slow-component eye velocity on the right versus the left. These values are used to provide a percentage comparison of response magnitude (reduced vestibular response) and direction bias of eye movement (directional preponderance) (Figure 6–9). Traditionally, the four maximum values are combined to produce the two outcomes through the use of Jongkee's formulas.[37]

Although four irrigations are typical, there are situations in which use of "ice water" (4°C) is needed and other situations in which only two irrigations of either warm or cool are sufficient.[14,38] Directional preponderance is interpreted as a bias in the central part of the system, making it easier to produce nystagmus in one direction than another. This bias is most often a result of asymmetric peripheral function, for which the central compensation process is incomplete, and is less likely to be a result of a lesion of the central part of the system.

It is critical in the interpretation of the caloric test to recognize its limitations. The most dominant of these is that the test is one of the horizontal semicircular canal function. Participation by either of the

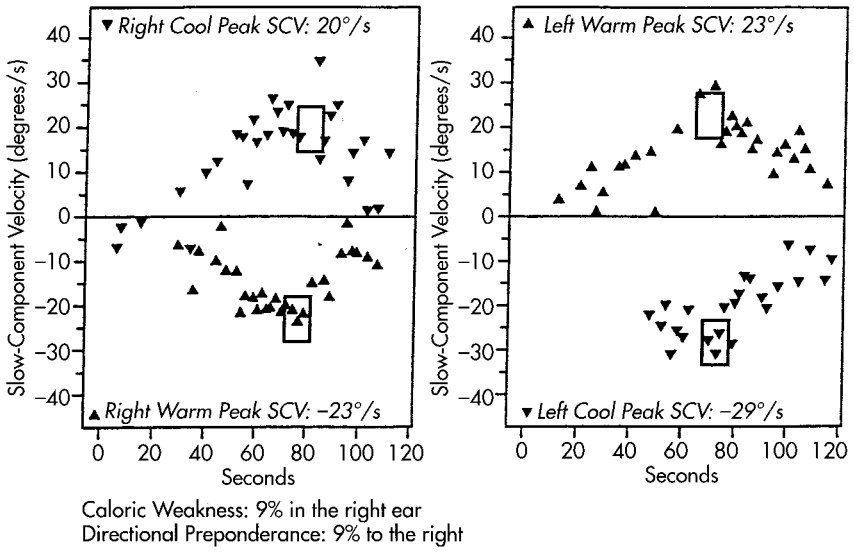


FIGURE 6–9. Plots of slow-component eye velocity (SCV) from nystagmus provoked by open-loop water irrigations as a function of time. Each triangle represents one SCV movement of the eye from the nystagmus trace. Responses for the right ear are shown on the left; those for the left ear are on the right. The orientation of the triangles represents either cool (30°C) or warm (44°C) irrigations. The plots are arranged so that right-beating nystagmus SCVs are on the bottom (right warm, left cool) and left-beating nystagmus SCVs are on the top (right cool, left warm). The velocity values given in the top or bottom of each plot represent the average maximum SCV calculated for the nystagmus beats within the rectangle shown on each plot. These maximum, average SCV values were used to calculate the caloric weakness and directional preponderance values shown at the bottom of the figure. Nine percent in the right ear means a 9% weaker response on the right compared with the left. Nine percent to the right means a 9% greater response for right-beating nystagmus compared with left-beating nystagmus. For purposes of calculations, SCVs from right-beating nystagmus are assigned a negative number and those for left-beating nystagmus are assigned a positive number. Reproduced with permission from Shepard NT and Telian SA.[14]

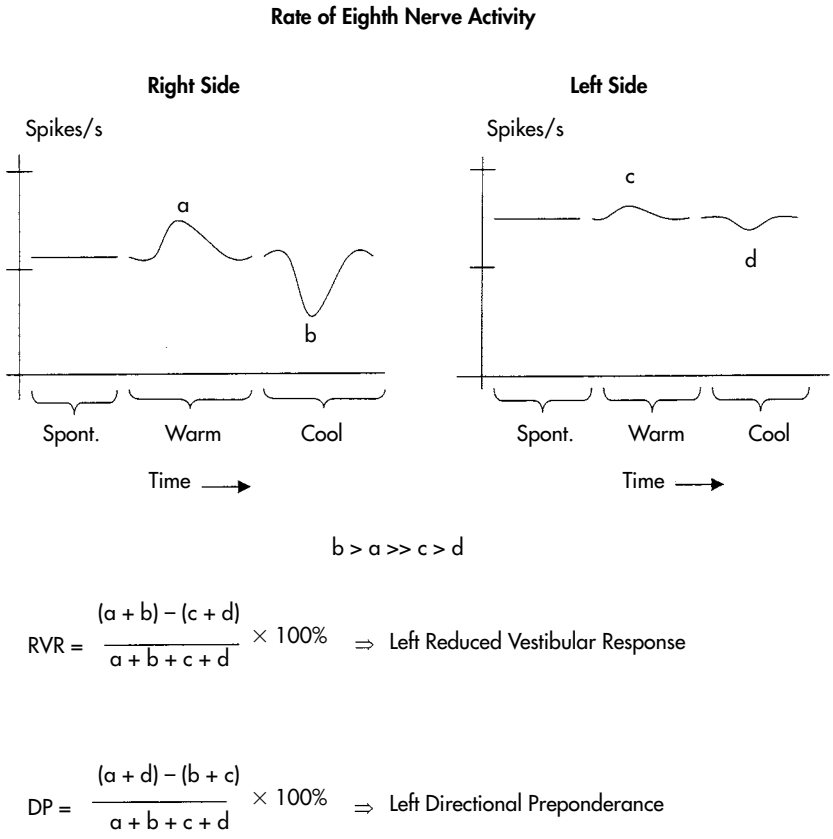


FIGURE 6–10. A schematic representation of a caloric irrigation test giving a left-side reduced vestibular response (RVR) with left directional preponderance (DP). The two plots at the top show in graphic form the estimated firing rate activity in spikes per second on the right and left eighth cranial nerves during the caloric test. The three intervals shown during the test are preirrigation (spontaneous) then the result from warm or cool open-loop water irrigations performed in an alternating bithermal protocol (right warm, followed by left warm, followed by left cool, and finally right cool). The amplitudes of the responses to the irrigations are given with the variables a, b, c, d. The spontaneous preirrigation neural activity on the eighth nerve is greater on the left than the right; however, the responses to irrigations are significantly less on the left than the right. For the caloric response traces, an upward deflection indicates an excitatory response on the nerve with the downward deflection indicating an inhibitory (reduction in firing rate) response. Given the relative values of the responses shown in the figure, the corresponding calculations of RVR and DP would produce the indicated result.

vertical canals is minimal at best. The test is also limited in its frequency response for the peripheral system (discussed in more detail below). Finally, the test is one of relative responsiveness of the peripheral system (horizontal canal) to an attempt at driving the system with an exogenous stimulus. It is not a test that is intented to reveal in a direct manner the ongoing strength of the neural activity of the right versus the left peripheral part of the system. As a result of this, it is possible to have a reduced vestibular response on the same side as that to which spontaneous and positional nystagmus beats and to which the directional preponderance is oriented. Yet nystagmus always beats toward (fast component of the nystagmus toward) the side of greater neural activity. This presents as a dichotomy in that the side with the reduced vestibular response must have the greater neural activity ongoing to have the spontaneous and positional nystagmus, as well as the directional preponderance oriented toward that same side. This is illustrated in the simple model presented in Figure 6–10. Figure 6–11 shows the actual results in a patient with classic Meniere's syndrome. The hearing loss and all auditory complaints for the

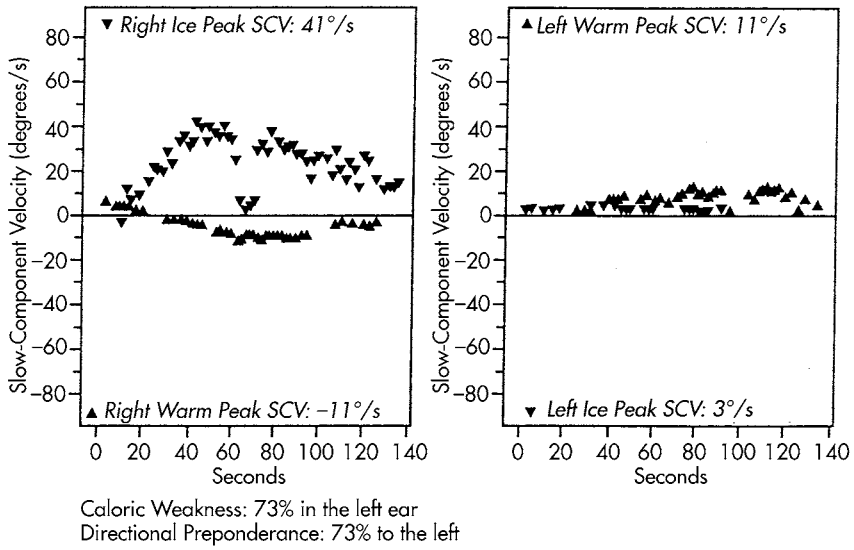


FIGURE 6–11. Plots of slow-component eye velocity (SCV) from nystagmus provoked by open-loop water irrigations as a function of time. Each triangle represents one SCV movement of the eye from the nystagmus trace. Responses for the right ear are shown on the left and those for the left ear on the right. The orientation of the triangles shown on the figure represents either ice water (4°C) or warm water (44°C) irrigations. The plots are arranged so that right-beating nystagmus SCVs are on the bottom (right warm, left ice water) and left-beating nystagmus SCVs are on the top (right water ice, left warm). The velocity values given in the top or bottom of each plot represent the average maximum SCV calculated for the nystagmus beats within the rectangle shown on each plot. These average maximum SCV values were used to calculate the caloric weakness and directional preponderance values shown at the bottom of the figure. For purposes of calculations, SCVs from right-beating nystagmus are assigned a negative number and those from left-beating nystagmus are assigned a positive number (note that left ice water gave a left-beating nystagmus, not a right, as would be expected—hence a positive averaged value). Reproduced with permission from Shepard NT and Telian SA.[14]

patient in this example were on the left. During the ENG, spontaneous left-beating nystagmus was noted and persisted in all positions with fixation removed. This indicates that the neural activity on the left must be greater than that on the right. Yet the caloric results shown in Figure 6–11 demonstrate a significant left reduced vestibular response. Not surprisingly, given the spontaneous nystagmus, a left directional preponderance is shown. This result is referred to as a left paretic lesion with an irritative status. This suggests that although the pathologic process is destructive, it is causing an irritative focus. The two most common disorders to result in an irritative lesion, with or without a paretic component, are early mass lesions of the cerebellopontine angle and Meniere's disease.

An additional feature of the example given in Figure 6–11 is the consequence of using a combination of unequal temperatures, warm and ice water irrigations. As long as the same protocol is performed on both sides, the use of warm and ice water is not a problem. The purpose is to increase the absolute value of the caloric response (as shown for the right irrigations in Figure 6–11) to avoid percentage calculations with numbers at or less than 10 degrees/second. However, as in this example, if both sides are not equivalent in their response to the irrigation, the increased response to ice on one side alone can lead to an artificially increased or falsely produced directional preponderance. In Figure 6–11, the actual directional preponderance, when estimated with the added effect of ice taken into account, is about 36%, not 73%.[14]

The use of the formulas provided in Table 6–2 requires the following assumptions to be met. First, anatomic symmetry between the external auditory canals, middle ears, and lateral aspects of the temporal bones must be present. Once the first assumption is fulfilled, then the second is that, technically, the same temperature of fluid or air must be delivered to

TABLE 6–2. Formulas for the Calculation of Reduced Vestibular Response and Directional Preponderance

Reduced vestibular response (RVR) (values –100% to +100%)	=	$\frac{[(RW + RC) - (LW + LC)] \times 100\%}{RW + RC + LW + LC}$
Directional preponderance (values unlimited)	=	$\frac{[(RW + LC) - (LW + RC)] \times 100\%}{RW + RC + LW + LC}$

RW = right warm; RC = right cool; LW = left warm; LC = left cool.

the plane of the tympanic membrane on each side. These both deal with the establishment of equivalent temperature gradients for stimulating the horizontal semicircular canals. If the first assumption is violated secondary to congenital or acquired processes, the test can still be performed, but the responses, right versus left, cannot be compared. The results can be used to indicate the presence or absence of a response. There are other protocols for calorics that will not be covered in this chapter; however, one in particular, the Torok Monothermal Caloric test, was developed to use data from only one ear irrigation at a time and not via a comparison.[39]

Since the peripheral part of the vestibular system functions across a frequency range, it is reasonable to question what portion of that range a caloric response occupies. The equivalent angular acceleration response falls in the lower-frequency range of response of the system (0.002 to 0.004 Hz). Therefore, absence of a caloric response to warm, cool, or ice water irrigations cannot be taken as an indication of complete lack of function. Testing by rotational chair is the tool needed to help define the true extent of bilateral lesions of the peripheral part of the vestibular system. In addition to the bilateral issue, the frequency response effect for the caloric test allows for a peripheral lesion to be present, yet not affect the caloric, but result in abnormalities on other tests such as the rotational chair (see below) or head on body rotation testing.[24]

Rotational Chair Testing

Rotational chair testing has been used to expand the evaluation of the peripheral part of the vestibular system. As with the ENG findings, the rotational chair evaluation can assist in site-of-lesion determination, counseling the patient, and confirmation of clinical suspicion of diagnosis and lesion site but is not likely to alter or impact significantly the course of patient management, excepting the patient with bilateral peripheral weakness.

The review of 2,266 patients was used to investigate the clinical utility of rotary chair testing in the evaluation of the peripheral part of the vestibular system. Among this group of patients, 16% had completely normal ENG studies. Among those with normal ENG results, rotational chair testing indicated abnormalities suggesting possible pathology in the peripheral part of the system in 80% of the cases. In all cases of bilateral caloric weakness, the chair findings help confirm the bilateral reduction in function of the peripheral part of the system. The test further defines the extent of the bilateral lesions. This additional information plays an important role in designing a vestibular rehabilitation program for these patients. It is the patient with bilateral lesions for whom rotational chair testing can make a direct impact on the management course in a rehabilitation program. There are also patients who for reasons other than vestibular dysfunction have mildly reduced caloric responses. Of these patients, the majority have normal rotational chair responses, suggesting that indication of bilateral paresis was a false-positive finding or that the extent of the bilateral lesions is mild and restricted to only the low-frequency range of function. Rotational chair testing is the only tool currently available for defining the extent of suspected bilateral peripheral lesions.

Patients diagnosed by clinical presentation and hearing test results, independent of balance function test results, with Meniere's disease, labyrinthitis, or vestibular neuritis numbered 311 during the period of this study.[14] Of this group, ENG was abnormal in 90%, suggesting a test sensitivity of this value. Rotary chair testing had a sensitivity suggesting involvement of the peripheral part of the system of only 66%. However, it is important to note that the

66% of patients identified by chair testing did not completely overlap with those identified by ENG as the combination of ENG and chair had a sensitivity of 100% in this group. Since there is no objective gold standard for identification of balance system lesions, there is currently no way to develop specificity figures, that is, the percentage of patients who do not have peripheral or central abnormalities who will have normal test results. Nor is there a good means for developing accurate sensitivity figures. Therefore, as with the discussion of oculomotor performance, the above group of 311 patients were well advanced in their respective disorders, and, clearly, the use of ENG and rotary chair testing together does not have a 100% sensitivity. A better estimate would be in the 80% range based on clinical diagnosis breakdown of the 2,266 patients.

From the above discussion, there appears to be good support for obtaining, in a subgroup of patients, the adjunctive information available from total-body, low-frequency rotary chair testing in the investigation of function of the peripheral part of the vestibular system. It is in this vein that the criteria for use of the rotational chair, discussed above, were developed.

As with the ENG test, the protocols for rotational chair testing are not new, and some have been used for well over a century. When considering the use of the chair protocols, it is important to remember that the sides of the peripheral part of the vestibular system have a "push-pull" arrangement, such that if one side is stimulated with angular acceleration, the opposite side is inhibited in its neural activity. Therefore, the chair is not a tool that can be easily used to isolate one side of the peripheral part of the system from the other for evaluation as each stimulus affects both sides simultaneously.

As with the ENG, electro-oculography or video-oculography can be used to monitor and record the outcome measure of interest, jerk nystagmus, which is generated in response to the angular acceleration stimulus. The VOR is the slow component of the jerk nystagmus and, as with ENG, is the portion of the eye movement for which velocity is calculated for analysis.

It must be remembered that with total-body rotational testing, the stimulus is being delivered to the head via movement of the whole body. Thus, the head must be secured to the chair with a restraint system (Figure 6–12). To make analysis as simple as possible, it is assumed that whenever the chair moves, the head is also making the same movement. Because of the potential for movement of the skin relative to the skull, this assumption becomes increasingly faulty at frequencies of 1 Hz or greater. Because of this, most commercial and clinical research systems have restricted test frequencies to 1 Hz or less. Shown in Figure 6–12 is a generic chair system, consisting of a chair on a computer-controlled electric torque motor. The head is held firmly to the chair, and the system is in an enclosure to allow testing in darkness with the eyes open. Means for providing visual stimuli, such as an optokinetic stimulus, are usual.

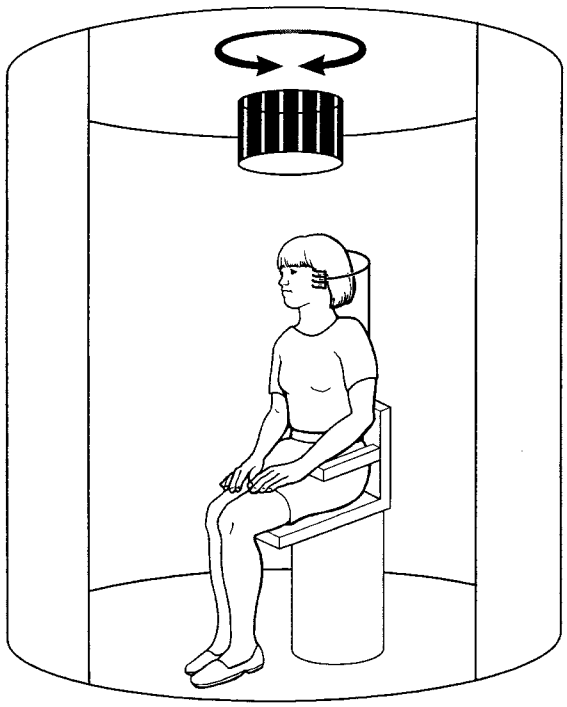


FIGURE 6–12. Generic rotational chair setup. The chair is on a computer-controlled motor within an enclosure and can be rotated in either direction. A device for holding the head to the chair is shown. A means for producing optokinetic stimulation is shown as the drum in the ceiling. Reproduced with permission from Shepard NT and Telian SA.[14]

Sinusoidal Rotation Typically, starting with the lower frequencies, the chair is moved with sinusoidal waveforms at a specific frequency from 0.01 to 1.28 Hz. Signal averaging is used to improve the signal-to-noise ratio and thereby improve the reliabil-

ity and validity of the analysis. Multiple cycles of a given frequency are delivered, producing repetitive to-and-fro movement of the chair in a sinusoidal harmonic acceleration paradigm. The slow-component eye velocity response from each cycle of stimulation is added to subsequent responses and divided by the number of cycles used, providing an average response for the test frequency. The frequency is then changed, and the process is repeated. Ideally, the more cycles that can be averaged, the more reliable the signal. Pragmatically, the number of cycles needs to be considered in light of the period (length of time for a single cycle) of the stimulus. As the period at 0.01 Hz is 100 seconds, to do more than three cycles becomes prohibitive. Unfortunately, the very low frequencies (less than 0.08 Hz) produce the weakest response from the VOR system and therefore have the poorest signal-to-noise ratio. In general, the very low frequencies are also most likely to produce unpleasant neurovegetative symptoms such as nausea and vomiting. The frequencies from 0.16 Hz and above can be completed quickly, allowing responses from as many as 10 cycles to be averaged. For all frequencies tested, the peak chair velocity is typically fixed at 50 to 60 degrees/second. Therefore, as the frequency is increased, the subject experiences increasing acceleration with decreasing excursion of the chair.

Three parameters are measured during rotational chair testing to characterize the function of the VOR and thereby evaluate the function of the peripheral part of the vestibular system. These parameters are phase, gain, and asymmetry. Figure 6–13 shows a schematized version of an averaged slow-component velocity response from multiple cycles of stimulation at a single frequency. The chair velocity is also shown, which correlates to head velocity, assuming that the head is properly stabilized. The parameters that characterize VOR function are developed by comparing the slow-component eye velocity profile to the head velocity profile.

Phase. This parameter of the VOR is the least intuitive of the three but has the greatest clinical significance owing to its ability to document dysfunction of the peripheral part of the system. Phase measurements objectify the timing relationship between head movement and reflex eye movement. Figure 6–13 illustrates findings expected from a normally functioning VOR system at test frequencies below

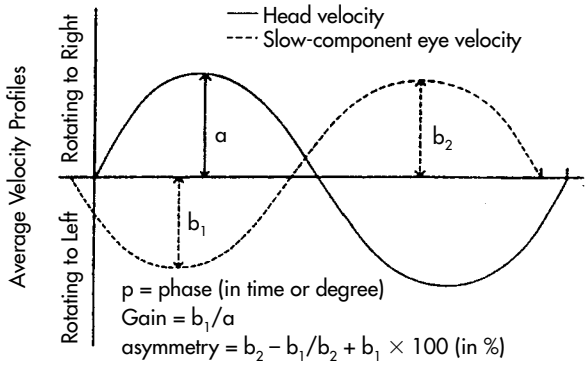
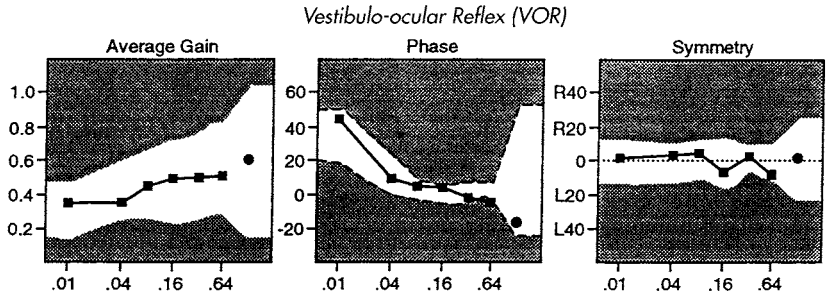


FIGURE 6–13. Averaged head and eye velocity plotted as a function of time or degrees is shown with indications for the development of the three parameters used to characterize performance with chair stimulation by sinusoids. b_1 = the maximum excursion of the slow-component eye velocity trace for rotating to the right; a = the maximum excursion of the head velocity trace for rotating to the right; b_2 = the maximum excursion of the slow-component eye velocity trace for rotating to the left. Reproduced with permission from Shepard NT and Telian SA.[14]

0.08 Hz. Under these circumstances, the compensatory eye movements can lead the head movement, as shown in Figure 6–13. The amount of this phase lead is called the phase angle and is typically measured in degrees. The center panel of Figure 6–14 shows a plot of phase angle versus frequency of rotation in a patient with normal rotary chair findings. The phase results can be used to calculate the system time constant from any frequency tested; however, assumptions can be applied that significantly simplify the calculations if the frequencies are restricted to 0.04 Hz and below.[17] In general, in this restricted frequency range, the relationship between phase angle and time constant is an inverse proportion. As phase angle increases, time constant decreases. In general, phase angle and its complement time constant should be thought of as a measure of the timing relationship of head and eye movement and are parameters that have specific values indicating normal VOR functioning.[17,40]

As seen in the center panel of Figure 6–14, the normal range for phase (based on two standard deviations above and below the mean) is indicated by the clear area. An increase in the phase lead outside this range implies an abnormally low time constant. From experimental studies of the velocity



Freq. Hz	Vel. d/s	Gain	Phase deg.	Symmet.%
■0.010	60	0.35	45	R 2.4
■0.040	60	0.37	11	R 3.9
■0.080	60	0.46	6	R 4.9
■0.160	60	0.50	5	L 6.0
■0.320	60	0.50	0	R 3.4
■0.640	60	0.52	-4	L 7.1
●1.280	50	0.61	-15	R 2.4

FIGURE 6–14. Normal rotary chair results from a patient. The plot on the left shows gain (eye velocity divided by head velocity) as a function of frequency of chair sinusoidal stimulation. The center plot shows phase angle in degrees as a function of frequency, and the plot to the right gives symmetry data in percent as a function of frequency. The *shaded areas* represent abnormality, and the individual values for gain, phase, and symmetry are given in the table at the bottom of the figure. Normal ranges are two standard deviations above the mean. Reproduced with permission from Shepard NT and Telian SA.[14]

storage integrator (vestibular nucleus region) that regulates the VOR system time constant, we know that damage to the labyrinth or the vestibular portion of the eighth cranial nerve causes a decrease in the time constant. Hence, increased phase lead, implying an abnormally low time constant, strongly suggests pathology in the peripheral part of the system. Caution in this interpretation must be used as other sources for the increased phase lead can result from damage in the vestibular nuclei within the brainstem and possible influence of migraine headache conditions. Therefore, other clinical information is needed to help localize the lesion to the labyrinth or eighth nerve. The significance of an abnormally low phase lead (abnormally high time constant) can suggest a lesion in the nodulus region of the cerebellum, an area that influences the velocity storage integrator in the brainstem.[41]

Gain The second parameter of the VOR measured in rotary chair testing is gain (eye velocity divided by head velocity) (see Figure 6–13). Gain measures give an indication of the overall responsiveness of the system. Unilateral peripheral weaknesses can cause a mild reduction in gain, especially at the lowest frequencies. However, the principal clinical use of gain measures is to define the extent of a bilateral reduction in responsiveness of the peripheral part of the system when using the sinusoidal protocol. The gain value helps verify that severely reduced or absent responses to caloric irrigations accurately reflect a bilateral weakness and did not result from an artifact of alertness or some other test pitfall. Figure 6–14 (left panel) shows normal results for gain as a function of the sinusoidal frequency of the stimulation.

Asymmetry In Figure 6–13, the schematic representation of asymmetry involves a comparison between the slow-component eye velocity to the right (positive values) compared with the left (negative values). It is important to recognize that these values are calculated and named by the direction of the eye movement that is produced by the VOR, that is, the slow component of the nystagmus. The situation is reversed when discussing directional preponderance from caloric irrigations. Directional preponderance values are calculated by slow-component velocity but, by convention, are named by the direction of the fast component of the nystagmus. Therefore, a patient who exhibits a right-beating directional preponderance (left slow-component velocity greater than right slow-component velocity) may show a left

greater than right asymmetry on rotational chair testing, indicating that during chair testing, left slow-component velocity was greater than right slow-component velocity, consistent with the directional preponderance. Directional preponderance and asymmetry from chair testing will not always both be abnormal. Both directional preponderance (caloric testing) and asymmetry (rotary chair testing) give an indication of bias within the system, favoring larger slow-component velocities in one direction versus the other. A bias usually results from a peripheral lesion with incomplete dynamic compensation in the central nervous system. Less commonly, it may indicate the presence of an uncompensated lesion in the central pathways. Whenever a VOR asymmetry is noted on rotary chair testing, the finding may be attributed to abnormalities in either peripheral part of the system: either a peripheral weakness on the side of the stronger slow-component velocity response or an irritative lesion on the opposite side. For example, a patient with an uncompensated right peripheral weakness will generally demonstrate a right greater than left slow-component velocity asymmetry. This is owing to an ability to produce a greater rightward compensatory eye movement when rotated toward the intact left side and a less intense leftward compensatory eye movement response produced after rotation toward the weaker right side. Figure 6–14 (right panel) shows a normal result for the asymmetry measurement.

Step Test This protocol is performed, as in the sinusoidal case, with the test booth in total darkness. A fixed chair velocity between 60 and 40 degrees/second is achieved by applying an acceleration impulse with a magnitude near 100 degrees/second^2. Once the desired velocity is reached, the acceleration is returned to 0 degrees/second^2, and the patient continues at the desired velocity. The VOR response to the initial acceleration stimulus is known as per-rotary nystagmus. The slow-component velocity intensity decays over time if the chair velocity is constant, and the subject falsely perceives that the chair is slowing down. The decay in slow-component eye velocity over time can be used to estimate the system's time constant, as discussed above. Use of a constant-velocity step protocol of 60 to 100 degrees/second works best to calculate the time constant. After 45 to 60 seconds of fixed-velocity rotation, a second impulse is applied to the chair. This is a deceleration step, usually of equal magnitude to the initial acceleration, bringing the chair to a rapid stop. Although the chair is now stationary, the subject will perceive motion in the opposite direction. The VOR response will produce nystagmus beating in the direction opposite to that produced by the initial acceleration, known as postrotary nystagmus. The decay of the slow-component eye velocity over time should be similar to that seen after the acceleration impulse, and, ideally, both should give similar estimates of the system's time constant. The entire procedure is then repeated with the initial rotation in the opposite direction. If a constant velocity of 180 to 240 degrees/second is used, the time constant calculation is not as accurate; however, differences in gain for acceleration to the right versus the left are made apparent in this manner, helping to identify a periphery with a weaker response.

The results are heavily influenced by the noise in both the recording and physiologic systems and the arousal of the patient prior to the acceleration as averaging is not used in this paradigm. As a result, there will be patients for whom the estimates of time constant from this protocol may not agree with estimates from the sinusoidal protocol. Given that both the step test and the sinusoidal acceleration tests are influenced by the noise issue, the two can be employed in parallel to increase the accuracy of estimates of the system time constant and identification of which periphery, if any, is functioning in an abnormal manner.[8,17,40,42,43]

Postural Control Assessment

Just as all patients who are being evaluated in the laboratory need tests for peripheral and central vestibulo-ocular pathway involvement, they also require some assessment of postural control ability. However, just as in the use of ENG and rotational chair testing, not all patients would need high-technology, formal postural control assessment. There are several different general approaches to formal postural control testing, each with specific technical equipment requirements and goals for the testing.[10,14,44] To reduce the scope of this discussion, comments will be restricted to the most common formal assessment tool used in the United States, CDP as formulated in the EquiTest® equipment (NeuroCom International, Inc., Clackamus, OR).[45–48] Briefly, the equipment detects vertical and horizontal forces

from the feet by two independent force plates on which the subject stands. The force plates can be made to translate forward or backward and rotate toes up or down to provoke movement of the subject's center of mass. The rotation movements can also be stimulated by the subject's own movements to create information from the ankle that is inaccurate. The visual surround can also be made to move stimulated by the subject's movements or independent of subject activity. Two principal testing protocols (discussed below) are used in patient evaluation: the Sensory Organization Test (SOT), using patient-stimulated floor and visual surround movements, and MCT using the translation and rotations of the support surface to cause subject movements. As with rotational chair testing, not all patients need to have full formal CDP. Suggested criteria as to when the use of full CDP would be most clinically revealing were presented in the introductory discussion above.

The survey of 2,266 patients with consecutive balance disorder, introduced in the ENG and rotational chair discussions, can again be used to assess the percentage of patients with abnormalities on CDP (EquiTest) and to study the types of patients most likely to have abnormal posturography results. In general, when all patients with indications of involvement of the peripheral part of the system are considered, only 30 to 35% show abnormalities on formal CDP. However, there was a group of 4 to 5% of this total population of 2,266 patients for whom CDP was the only abnormal finding across all of the ENG and rotational chair tests. This group was primarily over 65 years of age, with the chief complaint of unsteadiness with standing and walking. These patients had no complaints when sitting or lying and no perceptions of vertigo or other abnormal movements. These statistics played a principal role in the development of the above discussed criteria for when to proceed for full CDP. The protocols for EquiTest assess various aspects of quiet and dynamic postural control abilities. Although certain aspects of postural control are prerequisites for gait activities and certain abnormalities of postural control may cause problems in ambulation, posturography cannot be used in isolation to evaluate gait deficits. Gait evaluation has its own complete set of parameters that must be tested under a completely different set of conditions.

Sensory Organization Test Briefly, the SOT measures the ability to perform volitional, quiet stance during a series of six specific conditions (Figure 6–15). The first three provide for uninterrupted, accurate foot support surface information on a surface with adequate friction that is larger than the foot size. In condition 1, the eyes are open, whereas in condition 2, the eyes are closed. Under condition 3, the visual surround moves in a pattern that is stimulated by the anterior/posterior sway movements of the patient. Conditions 1 and 2 are a modified Romberg's test (a qualitative clinical test of patient stability when standing with feet together, arms folded, and eyes opened or closed) as the feet are at their normal separation rather than close together. Condition 3 presents a situation of visual conflict, for which visually accurate information is provided but of no significant help in maintaining quiet stance. Condition 3 presents misleading optokinetic and foveal visual cues about the position of the body in space. Conditions 4, 5, and 6 use the same sequence of the three visual conditions but with the foot support surface giving misleading information. As with the movement of the visual surround in condition 3, when testing under conditions 4, 5, and 6, sway movements of the patient in the sagittal (anterior/ posterior) plane drive the movement of the support surface in a rotational manner about an axis parallel to the ankle joint. In this way, somatosensory and proprioceptive information is not removed in conditions 4, 5, and 6 but is of limited use in maintaining upright stance. These conditions provide a disrupted relationship between body position and the ankle angle (that angle made between the upper surface of the foot and the anterior portion of the lower leg). Typically, after the simpler conditions 1 and 2, three trials are given for each of the more challenging conditions. The average performance is taken as representative of the patient's postural control ability under that sensory condition.

The equilibrium score is a percentage representing the magnitude of sway in the sagittal (pitch) plane for each trial of each condition. Details of how this score is obtained will not be repeated here.[48] However, it is important to realize that this score is based on a normal value of 12.5 degrees of anterior/posterior sway about the ankle joint, typically 8 degrees forward and 4.5 degrees backward. It is assumed that this range of sway is available to patients during the test. Some patients may not have this normal range because of physical restrictions at the ankle, or, because of limits of sway, patients have adopted secondary strategies owing to their sense of

SENSORY ORGANIZATION PROTOCOL

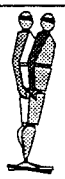
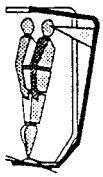
Condition	Vision	Support	Patient Instructions
1 	Normal	Fixed	Stand quietly with your eyes OPEN
2 	Absent	Fixed	Stand quietly with your eyes CLOSED
3 	Sway Ref	Fixed	Stand quietly with your eyes OPEN
4 	Normal	Sway Ref	Stand quietly with your eyes OPEN
5 	Absent	Sway Ref	Stand quietly with your eyes CLOSED
6 	Sway Ref	Sway Ref	Stand quietly with your eyes OPEN

FIGURE 6–15. The six sensory organization test conditions, showing which sensory input cues are available or accurate for each condition. Reproduced with permission from NeuroCom International, Inc., Clackamus, OR.

imbalance and fear of a potential fall. It is important to recognize the patient who has a reduction in limits of sway. If the limits of sway are reduced more than 50%, the interpretation of the patient's results may be inaccurate.[14] Figure 6–16 (top panel) is an example of the graphic representation of these results in a patient with normal CDP findings.

As with the overall balance system evaluation, this study is also interpreted using pattern recognition. The combinations of the six conditions that are abnormal are used to define a pattern of abnormality that can then be functionally interpreted. Table 6–3 presents the most common patterns and the related nomenclature.

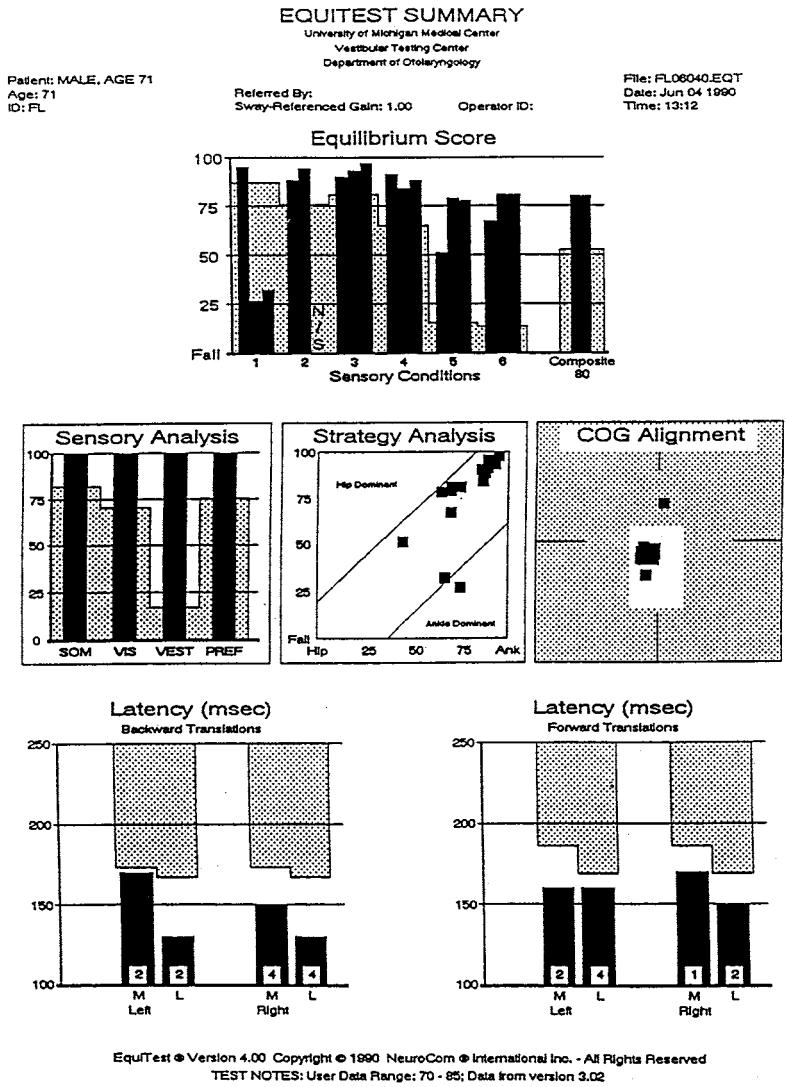


FIGURE 6–16. Results of dynamic posturography testing in a patient with all results interpreted as normal. The bar graph at the top plots a percentage equilibrium score for each of the six sensory organization test conditions (see Figure 6–15). A score of 100 indicates no sway in the sagittal plane, with "Fall" indicating that sway reached a magnitude equal to the theoretical limits of sway for the patient in the sagittal plane. The composite graph is the numeric average of the scores from the six conditions. The left panel of the center three shows a ratio analysis of the six conditions relative to four of the most frequent abnormal patterns. SOM = somatosensory dysfunction; VIS = visual dysfunction (typically combined with vestibular dysfunction); VEST = vestibular dysfunction; PREF = visual preference pattern. See the text for a discussion of these and other patterns. The middle panel in the center row plots a measure of horizontal shear force (an indication of body segment movement strategy) on the abscissa against the equilibrium scores on the ordinate for each of the six test conditions. The right panel of the center row shows a measure of alignment of the center of gravity (COG) in the sagittal and lateral planes for each of the six conditions tested. The bar graphs in the bottom row (from movement coordination portion of the testing) plot latency to onset of active recovery to induced forward sway (left graph) and induced backward sway (right graph). The latencies are given in milliseconds for left and right leg for two sizes of platform translations. See the text for interpretation of these results. Reproduced with permission from NeuroCom International, Inc. Reproduced with permission from Shepard N, Telian SA. Evaluation and management of balance system disorders. In: Katz J, editor. Handbook of audiology. 4th ed. Baltimore: Lippincott Williams & Wilkins; 1994.

TABLE 6–3. Common Patterns of Abnormal Performance on the Sensory Organization Test

- Vestibular dysfunction pattern: abnormal on conditions 5 and 6 (alternatively condition 5 alone).
  Vestibular dysfunction pattern indicates the patient's difficulty in using vestibular information alone for maintenance of stance. When provided with accurate visual and/or foot somatosensory information, stance is within a normal range.
- Visual and vestibular dysfunction pattern: abnormal on conditions 4, 5, and 6.
  Visual and vestibular dysfunction pattern indicates the patient's difficulty in using accurate visual information with vestibular information or vestibular information alone for maintenance of stance. When provided with accurate foot support surface cues, stance is within a normal range.
- Visual preference pattern: abnormal on conditions 3 and 6 (alternatively condition 6 alone).
  Visual preference pattern indicates the patient's abnormal reliance on visual information, even when inaccurate. When provided with accurate foot support surface information together with accurate or absent visual cues, or absent vision and vestibular information alone, stance is within a normal range.
- Visual preference and vestibular dysfunction pattern: abnormal on conditions 3, 5, and 6.
  Visual preference and vestibular dysfunction pattern indicates the patient's difficulty in using vestibular information alone and the patient's abnormal reliance on visual information, even when inaccurate. When provided with accurate foot support surface information together with accurate or absent visual cues, stance is within a normal range.
- Somatosensory and vestibular dysfunction pattern: abnormal on conditions 2, 3, 5, and 6.
  Somatosensory and vestibular dysfunction pattern indicates the patient's difficulty in using foot support surface information with vestibular information or vestibular information alone for maintenance of stance. When provided with accurate visual information, stance is within a normal range.
- Severe dysfunction pattern: abnormal on four or more conditions not covered in the above descriptions, for example, 3, 4, 5, and 6; 2, 3, 4, 5, and 6; or 1, 2, 3, 4, 5, and 6.
  Severe dysfunction pattern indicates the patient's difficulty with stance independent of the sensory information (vestibular, visual, and/or somatosensory) provided. Note that these situations often involve a dominant feature such as significantly abnormal conditions 5 and 6, or they may involve equally distributed difficulties on all conditions affected.
- Inconsistent pattern: abnormal on conditions 1, 2, 3, 4, or any combination and normal on conditions 5 and 6.
  Inconsistent pattern indicates that performance of the patient is difficult to explain with normal or typical pathophysiologic conditions and could imply volitional or nonvolitional exaggerated results.

By far the most common pattern is the vestibular dysfunction pattern, comprising approximately 45% of all abnormalities on this test in our facility. The most important aspect of interpretation for the SOT is that it provides information as to which input system cues the patient is unable to use for performing the task of maintaining postural control. In other words, it provides a relative measure of the patient's ability to use the sensory input cues of vision, vestibular, and proprioceptive/somatosensory to maintain quiet upright stance. The test does *not* provide relative information as to which of the sensory systems has lesions, causing postural control abnormalities. Therefore, the SOT of CDP provides no site-of-lesion information; it is strictly a test of functional ability. The test in no way implies that there is a central or peripheral vestibular system lesion, nor does it imply central or peripheral pathway lesions in the visual or somatosensory/proprioceptive systems. The information should be interpreted only to reflect which input information the patient is able (or, conversely, unable) to use for the task at hand.

Dynamic posturography is useful for identification of patients who may be, for whatever reason, exaggerating their condition. Recent work by several

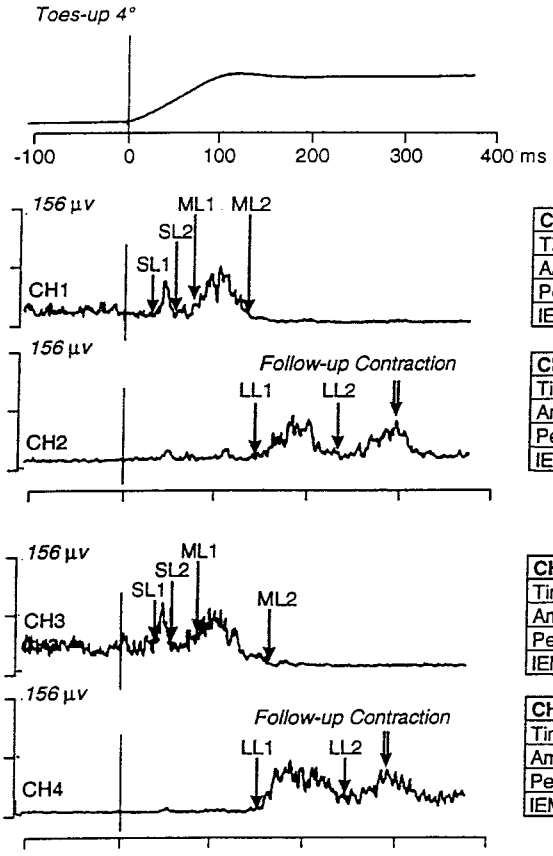
investigators (not all with EquiTest) has attempted to quantify the use of this tool to identify these patients and list qualitative factors that would raise questions in this dimension.[49–51]

Motor Control Test Information about the ability to react to unexpected perturbations in the subject's center of mass position is obtained with the MCT. The center of mass perturbations is created by abrupt anterior or posterior horizontal translations of the support surface. Typically, three increasingly large translations in both directions are administered. The increasing size of the translation creates a stimulus intensity series. The profile of the surface movement is varied for each patient based on height so that all translations are normalized to a 6-foot-tall person.[48] This allows for direct comparison of results across patients. After the three posterior and the three anterior translations, unexpected rotations about the ankle are used. Contrary to the horizontal translations, the typical muscle response that is mapped to the stimulus provoked by rotary stimuli is destabilizing. The patient must then be able to adapt to the new stimulus on repeated trials. Five randomly timed toes-up or toes-down rotations provide relative information about the patient's ability to adapt to this familiar but destabilizing stimulus. For this protocol, as with the SOT, floor reaction force detected by the force plates in the support surface is measured. The principal output parameter is the latency to onset of active recovery from the unexpected translations (see Figure 6–16, bottom panels). Other information obtained from the protocol includes weight distribution onto right or left leg and a relative measure of strength as a function of the size of the perturbation.[48]

This study is used less as a functional evaluation than the SOT and more to evaluate the long-loop neural pathway. This pathway begins with inputs from the ankle, knee, and hip regions (tendon and muscle stretch receptors) and projects to the motor cortex and back to the various muscles of postural control, including upper and lower body. When an abnormal latency to onset of active recovery from induced sway is noted, then problems in the long-loop pathway should be considered. The explanation may be as simple as ongoing joint or back pain, a congenital condition of the back or lower limbs, or an acquired lesion involving the neural pathways of the afferent or efferent tracts. Therefore, abnormalities of the movement coordination test, related to latency, are nonspecific indicators of potential problems in the long tracts or the musculoskeletal system needed to coordinate recovery from unexpectedly induced sway in the sagittal plane. Other abnormalities from this portion of the testing include inappropriate weight bearing or an inability to scale properly the strength of the response to the increasing size of the perturbations. Such findings may provide information that helps explain the patient's complaints of dysequilibrium. These abnormalities are unlikely to directly implicate involvement of the neural pathways if the latency findings are normal. In many cases, the weight shift or scaling problems may be maladaptive behaviors developed in response to the initial symptoms of the vestibular disorder.[9]

Postural Evoked Responses Postural evoked responses are used to define patterns of muscle response that have been associated with specific lesion sites and dieases. Muscle activity from the distal lower extremities is stimulated by sudden toe-up rotations of the support surface (the force plate platform of EquiTest). The muscle activity stimulated by this dorsiflexion movement at the ankle is recorded with surface EMG electrodes.[47,52] The responses from the medial gastrocnemius and the anterior tibialis are recorded. To improve the signal-to-noise ratio of the evoked EMG activity, the rotation is repeated, with random, interstimulus intervals, and the EMG responses are rectified and averaged over 15 to 20 responses. This allows for clearer identification of onset and offset times of muscle contraction following the stimulus. There are three specific responses obtained, as illustrated in Figure 6–17. The short and medium latency responses from the contraction of the gastrocnemius are shown in traces from channels 1 and 3 of Figure 6–17. The third response is the long latency obtained from the contraction of the anterior tibialis, shown in channels 2 and 4 of Figure 6–17.

The EMG patterns for contraction from the gastrocnemius and the anterior tibialis muscles are compared with those that have been associated with specific pathologies, such as multiple sclerosis, Parkinson's disease, or other neurologic lesions. Patterns have been described for lesions in the anterior cerebellum and the basal ganglia, as well as for spinal cord compression. When the contraction pattern is



CH1	SL1	SL2	ML1	ML2
Time (ms)	38	56	79	143
Amp (mv)	20	19	25	10
Peak (mv)	61		79	
IEMG (mvS)	0.6		2.8	

CH2	LL1	LL2
Time (ms)	160	241
Amp (mv)	10	7
Peak (mv)	35	
IEMG (mvS)	1.4	

CH3	SL1	SL2	ML1	ML2
Time (ms)	39	57	87	165
Amp (mv)	36	32	42	10
Peak (mv)	94		84	
IEMG (mvS)	1.1		3.3	

CH4	LL1	LL2
Time (ms)	163	253
Amp (mv)	15	20
Peak (mv)	72	
IEMG (mvS)	3.9	

FIGURE 6–17. Postural evoked response results for a normal subject. The plot at the top shows the position profile of the rotation of the surface the subject is standing on as a function of time. Deflection upward indicates toe-up rotation. The four plots on the left give the averaged, rectified surface electromyographic responses in amplitude of contraction (in microvolts) as a function of time. Channels 1 and 3 (CH1 and 3) are results from the left and right gastrocnemius muscles, respectively. Channels 2 and 4 (CH2 and 4) are the results from the left and right anterior tibialis muscles, respectively. The short (SL), medium (ML), and long latency (LL) responses are indicated at onset with a number 1 and offset with a 2. The tables at the right give the numeric values for each response, indicating onset and offset times and amplitudes, peak amplitude contraction, and integrated amplitude (IEMG). Reproduced with permission from Shepard NT and Telian SA.[14]

unrecognized, the interpretation is based on knowledge of the underlying neural pathways considered responsible for the specific muscle activity. In general, these involve mediation of the short latency response via the spinal cord (H-reflex). The medium latency response is primarily controlled via the spinal cord, with amplitude size determined by the brainstem and basal ganglia. The functional stretch reflex, the long latency response, involves brainstem and cortical activity. Defined patterns of muscle response have been associated with specific lesion sites and disease classifications.[53] Normative results for the paradigm have been developed across age and have been shown to have sensitivity and specificity of 68% and 87%, respectively, for identifying the specific disease entities reflected by the defined patterns of abnormal responses.[9,52] As with the MCT, the EMG evaluation does not distinguish afferent from efferent disruptions that may underlie the abnormal muscle responses. With additional clinical investigations of sensitivity in the lower limbs and/or the use of lower limb somatosensory evoked

responses, pathology affecting sensory input may be distinguished from motor output abnormalities.

The CDP tests described above can provide data ranging from purely functional information to that of specific site-of-lesion, postural control pathways by changing the protocol and the output parameters. Therefore, for the patient with a classic peripheral vestibular lesion, the utility of CDP lies in the use of SOT for functional information related to functional dimensions of compensation not available from the extent- and site-of-lesion studies of ENG and the rotary chair. For the patient with a less well-defined lesion site or evidence of central nervous system involvement, the functional information is then supplemented with data that approach a more specific evaluation of the central and peripheral postural control pathways. Additionally, CDP can serve to help in the design and monitoring of vestibular and balance rehabilitation programs, primarily the SOT protocol with the emphasis on functional performance.

OFFICE "BEDSIDE" EVALUATIONS OF THE BALANCE SYSTEM

A variety of test procedures may be used in the office setting to assess the patient with balance disorder. These, like the laboratory studies, assist in the identification of the extent and site of the lesion. These straightforward clinical tests are essentially variations of the related laboratory studies but have less ability to quantify the outcomes. The theoretical basis behind many of these tests is well founded on physiologic considerations (please see Chapter 4). Owing to the subjective nature of these tools, the validity and reliability of these tests are reduced compared to the formal laboratory studies. Unfortunately, little clinical research exists to define the sensitivity/specificity and test reliability for these procedures.[6,14,17,18,26,54,55]

The bedside vestibular and balance evaluation is presented in Chapter 20. In the majority of patients, this evaluation produces the same impressions as the extensive laboratory studies discussed above. These tools are, however, not as sensitive to mild anomalies as the laboratory studies. It is also important that the bedside chair study and the modified CTSIB study of balance[14] be used with the ENG to expand the laboratory investigation when formal equipment such as the rotational chair and posturography are not readily available.

OTHER INVESTIGATIVE MODALITIES

A discussion of the investigation of patients with dizziness and balance disorders would not be complete without a brief description of use of neuroradiographic techniques and serologic studies and when the pursuit of psychological issues may be useful.

Psychological Disorders

Patients with dizziness often have psychological symptoms such as anxiety or depression. From a diagnostic standpoint, psychiatric disorders may be the sole cause of dizziness, but just as often they co-occur with neurotologic illnesses. Therefore, no specific characteristics of dizziness itself are pathognomonic of a psychiatric cause. Rather, the following associated features suggest that a psychiatric illness may be present, with or without vestibular pathology: (1) illusions of movement that are more prominent than true vertigo (typically a rocking or swaying), (2) functional impairment out of proportion to objective medical findings, (3) autonomic symptoms during bouts of dizziness (palpitations, chest pain, dyspnea, tremulousness, paresthesias), (4) changes in sleep or appetite, (5) excessive worry about dizziness, and (6) avoidance of situations associated with dizziness.[3] A recent investigation found that a brief psychological screening questionnaire, the BSI-53, was a useful adjunct to the neurotologic examination for detecting psychiatric disorders in patients with dizziness.[56] Recognition of any of the above features with a clinical impression of anxiety or depression would be sufficient to suggest further investigation by a psychiatrist or psychologist. This should be a professional with whom prior standing arrangements have been established since an understanding of the general area of balance and dizziness disorders is needed to evaluate appropriately the potential impact of a psychological component to the patient's complaints.

Neuroradiographics

The use of MRI, magnetic resonance angiography (MRA), and computed tomography (CT) scans is widespread. The following discussion focuses on when each of the studies would be of utility. Below are the bedside (office examination) signs that would warrant a referral for MRI as each of these is an indicator of possible central nervous system involvement. Some of these same findings are also appreciated on an ENG study. The most prominent

central indicator is that of pure vertical nystagmus (especially when noted with fixation present) or a pure torsional nystagmus with fixation removed.

Additional signs are as follows:

- direction-changing nystagmus in a given head position
- sustained or nonfatiguing positional nystagmus
- dissociated (disconjugate) eye movements most easily noted with saccade testing
- saccadic disruptions to smooth pursuit
- direction-changing gaze-evoked nystagmus
- abnormal posture when seated, inability to stand
- focal motor deficit
- dysarthria, dysphagia, diplopia, dysmetria (limb ataxia)
- Horner's syndrome (sympathetic paresis of the pupil, with unilateral failure to dilate in darkness and ptosis)
- loss of pinprick or temperature sensation on one side of the face and/or on the other side of the body
- intractable hiccups
- hyperventilation-induced nystagmus
- hemifacial or hemibody sensory or motor deficit
- visual inversion (90- to 180-degree reversals of the visual scene, "floor on the ceiling"), drop attacks, visual loss, and/or confusion

There are signs with acute-onset vertigo that should trigger an emergent imaging study. A diffusion-weighted MRI is the preferred technique over CT if it can be obtained within hours of the onset of symptoms. Cerebellar infarcts may mimic peripheral vestibular paresis. These may become hemorrhagic, and swelling during the first few days after the stroke can cause brainstem herniation and death. The signs that necessitate a scan are the following:

- unilateral or asymmetric hearing loss
- brainstem or cerebellar symptoms other than vertigo, most commonly vertical nystagmus, nystagmus not suppressed with fixation, and/or inability to stand unassisted (cerebellar infarctions are usually cardioembolic)
- stroke risk factors (diabetes, hypertension, tobacco use, history of myocardial infarct)
- acute onset associated with neck pain (suggestive of a vertebral artery dissection)
- direction-changing spontaneous nystagmus
- new-onset severe headache

Patients with a history of brainstem or cerebellar symptoms mentioned above lasting from typically 5 to 60 minutes may be having posterior circulation (vertebrobasilar) transient ischemic attacks. This is an indication for the use of MRA to rule out stenosis of the vertebral or basilar arteries. Magnetic resonance angiography can evaluate both intracranial and cervical vessels. Transcranial Doppler studies can be performed on the intracranial circulation and provide evidence of stenosis by showing increased flow velocity. The direction of blood flow can also be determined: with proximal vertebral artery occlusion there can be retrograde blood flow from the basilar into a vertebral artery.

Computed tomography without contrast is the examination of choice for delineating the bony anatomy of the temporal bone. There are several indications for this study in patients with the primary complaint of dizziness. This type of scan may detect a perilabyrinthine fistula secondary to cholesteatoma. In primary lesions of the petrous apex that invade the labyrinth or internal auditory canal, this scan can be used both to aid in the diagnosis and to determine the extent of involvement of the temporal bone. Cholesterol granulomas and cholesteatomas of the petrous apex will appear as homogeneous lesions that gradually expand the margins of the bone.[57–60] Fibrous dysplasia will manifest as an osseous lesion with a heterogeneous, "ground glass" appearance on scan. Sarcomas, such as chondrosarcomas, will manifest as bulky, irregular, and destructive lesions that display a "popcorn" pattern of calcification. Metastases to the temporal bone will appear as irregular, destructive lesions.[60] The degree of bony invasion and destruction of glomus jugulare tumors can also be assessed by this scan. A suggestion of the presence of an acoustic neuroma may be made by the observation of asymmetric widening of the internal auditory canal. In younger patients, congenital osseous dysplasias of the inner ear can be detected by CT scan.[61] A further use of CT, but with very fine (0.1 to 0.5 mm) cuts, is in the detection of superior semicircular canal dehiscence that has been shown to result in a sound- and/or pressure-provoked vertigo. As with other uses of CT and MRI, there are clues in the history and presenting signs that would lead to a request for this specialized study.[62]

Computed tomography of the temporal bone with contrast is rarely indicated for the evaluation of dizziness. The addition of contrast does not improve

the visualization of cholesteatomas, cholesterol granulomas, or other cystic lesions of the temporal bone. In contrast, chondrosarcomas and metastatic lesions will enhance. This examination remains the evaluation of choice to rule out the presence of an acoustic neuroma in patients who cannot undergo MRI.

Of all of the radiologic evaluations available, MRI remains the most useful in evaluating the patient complaining of dizziness. It is the undisputed evaluation of choice to rule out the presence of acoustic neuromas or other neoplasms of the cerebellopontine angle.[63,64] When evaluating the petrous apex, it provides invaluable assistance in determining the specific type of cystic lesion present.[57–60] Cholesterol granulomas are bright on both T_1-weighted and T_2-weighted images and do not enhance. In contrast, epidermoids (cholesteatomas) are bright only on T_2-weighted scans. They are also the imaging study of choice for evaluation of tumors of the jugular foramen region.

The fast spin echo, T_2-weighted MRI scan has been suggested as a useful and economical examination to screen for the presence of acoustic neuromas.[65,66] Although its accuracy has been validated, it has not found universal acceptance. This form of scan is also a useful complement to the CT scan in patients with suspected inner ear dysplasias.

Serologic Studies

Serologic studies can be used for investigation of central disorders and peripheral neuropathy. A listing of the more common of these is given below:

- Vasculitic disorders: erythrocyte sedimentation rate (ESR), complement levels, rheumatoid factor, antineutrophil cytoplasmic antibodies (ANCA) (Wegener's granulomatosis, polyarteritis nodosa), cryoglobulinemia, hepatitis B antigenemia, eosinophilia, systemic lupus erythematosus (antinuclear antibody [ANA] and anti–disease-specific DNA), Sjögren's (autoantibodies SSA and SSB)
- Peripheral neuropathy owing to infectious, metabolic, or neoplastic conditions: rapid plasma reagin (RPR), human immunodeficiency virus, glucose, hemoglobin A_{1c} (diabetes), SPEP, vitamin B_{12} deficiency

When evaluating patients with peripheral neuropathy, one should consider in the history factors known to be toxic to peripheral nerves, such as B_6 overdose, metronidazole, or cisplatinum. Distal sensory loss may also be on the basis of lumbar or cervical spine disease.

In patients with hearing loss associated with true vertigo, serologic testing is usually indicated once a retrocochlear lesion has been ruled out. In these cases, it is mandatory to rule out late secondary or tertiary syphilis using the FTA-ABS (fluorescent treponemal antibody-absorbed) or equivalent test. The VDRL (Venereal Disease Research Laboratory) or RPR test lacks sufficient sensitivity to be used as a screen in this population. Autoimmune testing is reserved for patients with rapidly progressive hearing loss, evolving over a period of weeks to months. Approximately 30% of these patients will manifest vestibular complaints.[67] There is no standardized testing protocol for the workup of these patients. A complete blood count with differential, ESR, electrolytes, blood urea nitrogen, creatinine, and urinalysis should be part of the general screen. The ANA assay has been the most revealing. The role of antiphospholipid antibodies in neurotologic disease has yet to be clarified, but they may prove to be significant mediators of labyrinthine pathology.[68] A Western blot test to detect antibodies directed against the inducible form of heat shock protein 70 is commercially available and may be predictive of the presence of corticosteroid-responsive hearing loss.[69,70] Other tests that may compose part of an autoimmune screen include assays for rheumatoid factor, complement levels, anti-Sjögren's antibodies, and c-ANCA, although the cost effectiveness of these tests has yet to be established. Authorities often advocate the performance of Lyme titers in patients with neurovestibular complaints, although there is little literature to support the association between Lyme disease and peripheral auditory and especially vestibular disorders.[71,72] Similarly, metabolic diseases (diabetes mellitus, hypothyroidism, hypercholesterolemia) have not been established as mediators of neurovestibular disease.

VESTIBULAR AND BALANCE REHABILITATION THERAPY

Although this chapter is dedicated to the evaluative process for the patient with complaints of imbalance, vertigo, or general dysequilibrium, it is during this process that the first suggestions for use of VBRT should be developed. Therefore, closing the chapter with a brief discussion on this topic and the

relationship between the evaluation tools and the use of VBRT is appropriate.

Vestibular and balance rehabilitation therapy is a symptom-driven program for determination of who is an appropriate candidate. Therefore, it is during the history that the first possible indications for the use of this management technique would be developed. For the most part, other than the SOT of posturography, the laboratory tests described above do not provide indications for the use of this treatment option. There are situations in which the symptom complaints are questionable for use of the program, yet findings on postural control studies suggest functional deficits that can be addressed with the VBRT techniques, and its use would be appropriate. In the case of bilateral vestibular paresis, the rotary chair is the only test that can provide information about the extent of the bilateral weakness. With those chair results, VBRT programs can be altered for more realistic goals. Lastly, indications from ENG or rotary chair of a unilateral reduced vestibular response to caloric irrigation or abnormal timing relationship between eye and head movement from rotary chair support the use of adaptation exercises to improve VOR gain. It is unlikely that repeat use of ENG or rotary chair following VBRT would be called for since the utility of these tools in determining compensation status is very limited, and the primary focus in a VBRT program is to enhance the compensation process in many dimensions. However, repeated use of various postural control and dynamic visual acuity assessments[11] would be useful as a monitor and/or final outcome measure of the effectiveness of a VBRT program since these studies are primarily functional in nature. The patients who are the most appropriate for VBRT on first review are suggested from any one of the following:

- Patients with symptoms provoked by head or visual motion
- Symptoms that are continuous with motion exacerbation
- Evaluations revealing balance or gait dysfunction with or without either of the traits listed above

The patients who are most probably not appropriate for use of VBRT as the initial management technique have the following characteristics to their symptoms:

- Symptoms of only spontaneous events that are more frequent than one time every 6 to 8 weeks and last longer than 15 minutes at a time.
- No provocative activity or balance dysfunction can be realized during the therapy evaluation; therefore, nothing is found on which to base exercise activities.
- Progressive central lesions involving gait and balance have not been shown to respond to therapy for balance and gait. These patients may benefit from exercises to reduce eye and head movement sensitivity and improve safety with ambulation-oriented goals.

The general principles of designing vestibular rehabilitation programs involve exposing the patient to the stimuli that provoke vertigo and cause slippage of the visual signal on the retina and challenging areas of deficiency in postural control and ambulation. First, the therapist must identify those activities or environmental situations that provoke symptoms. Second, the patient's functional deficits regarding balance and gait must be identified. These may be caused by the vestibular symptoms, lesions in nonvestibular regions, or maladaptive behavior that has developed in response to the symptoms. Lastly, it is desirable to challenge the sedentary lifestyle that the patient with vestibular disorder often adopts. An active lifestyle including regular exercise that accounts for age and other health constraints will serve as a maintenance program once active therapy is completed. Extensive literature is available to discuss the protocols, techniques, and efficacy of this form of management.[33,73–76]

REFERENCES

1. Tusa RJ. Diagnosis and management of neuro-otological disorders due to migraine. In: Herdman SJ, editor. Vestibular rehabilitation. Philadelphia: FA Davis; 2000. p. 298–315.
2. Neuhauser H, Leopold M, von Brevern M, et al. The interrelations of migraine, vertigo and migrainous vertigo. Neurology 2001;56:436–41.
3. Staab JP. Diagnosis and treatment of psychologic symptoms and psychiatric disorders in patients with dizziness and imbalance. Otolaryngol Clin North Am 2000;33:617–36.
4. Tusa RJ. Psychological problems and the dizzy patient. In: Herdman SJ, editor. Vestibular rehabilitation. Philadelphia: FA Davis; 2000. p. 316–30.
5. Yardley L. Overview of psychologic effects of chronic dizziness and balance disorders. Otolaryngol Clin North Am 2000;33:603–16.

6. Walker MF, Zee DS. Bedside vestibular examination. Otolaryngol Clin North Am 2000;33:495–506.
7. Shepard NT. Electronystagmography (ENG) testing. In: Goebel JA, editor. Practical management of the dizzy patient. Philadelphia: Lippincott Williams & Wilkins; 2000.
8. Shepard NT. Rotational chair testing. In: Goebel JA, editor. Practical management of the dizzy patient. Philadelphia: Lippincott Williams & Wilkins; 2000.
9. Shepard NT. Clinical utility of the motor control test (MCT) and postural evoked responses (PER). Clackamus (OR): A NeuroCom Publication; 2000. p. 3–19.
10. Monsell EM, Furman JM, Herdman SJ, et al. Technology assessment: computerized dynamic platform posturography. Otolaryngol Head Neck Surg 1997;117: 394–8.
11. Herdman SJ, Tusa RJ, Blatt P, et al. Computerized dynamic visual acuity test in the assessment of vestibular deficits. Am J Otol 1998;19:790–6.
12. Jacobson GP, Newman CW. The development of the Dizziness Handicap Inventory. Arch Otolaryngol Head Neck Surg 1990;116:424–7.
13. Jacobson GP, Newman CW, Hunter L, Balzer G. Balance function test correlates of the Dizziness Handicap Inventory. J Am Acad Audiol 1991;2:253–60.
14. Shepard NT, Telian SA. Practical management of the balance disorder patient. San Diego: Singular Publishing Group; 1996.
15. Zee DS. Vestibular adaptation. In: Herdman SJ, editor. Vestibular rehabilitation. Philadelphia: FA Davis; 2000. p. 77–90.
16. Jacobson GP, Newman CW, Kartush JM, editors. Handbook of balance function testing. St. Louis: Mosby-Year Book; 1993.
17. Baloh R, Honrubia V. Clinical neurophysiology of the vestibular system. 2nd ed. Philadelphia: FA Davis; 1990.
18. Leigh RJ, Zee DS. The neurology of eye movements. 3rd ed. Philadelphia: FA Davis; 1999.
19. Baloh RW, Halmagyi GM, editors. Disorders of the vestibular system. New York: Oxford University Press; 1996.
20. Furman JM, Cass SP. Balance disorders: a case-study approach. Philadelphia: FA Davis; 1996.
21. Bronstein AM, Brandt T, Woollacott M, editors. Clinical disorders of balance, posture and gait. London: Arnold; 1996.
22. Rosenberg ML, Gizzi M. Neuro-otologic history. Otolaryngol Clin North Am 2000;33:471–82.
23. Shonel G, Kemink JL, Telian SA. Prognostic significance of hearing loss as a lateralising indicator in the surgical treatment of vertigo. J Laryngol Otol 1991; 105:18–20.
24. O'Leary DP, Davis LL. High-frequency autorotational testing of the vestibulo-ocular reflex. Neurol Clin 1990;8:297–312.
25. Henn V, Straumann D. Three-dimensional eye movement recording for clinical application. J Vestib Res 1999;9:157–62.
26. Halmagyi GM, Curthoys IS. A clinical sign of canal paresis. Arch Neurol 1988;45:737–9.
27. Whitney SL, Herdman SJ. Physical therapy assessment of vestibular hypofunction. In: Herdman SJ, editor. Vestibular rehabilitation. Philadelphia: FA Davis; 2000. p. 333–72.
28. El-Kashlan HK, Shepard NT, Asher A, et al. Evaluation of clinical measures of equilibrium. Laryngoscope 1998;108:311–9.
29. Jacobson GP, Newman CW, Safadi I. Sensitivity and specificity of the head-shaking test for detecting vestibular system abnormalities. Ann Otol Rhinol Laryngol 1990;99:539–42.
30. Minor LB, Haslwanter T, Straumann D, Zee DS. Hyperventilation-induced nystagmus in patients with vestibular schwannoma. Neurology 1999;53: 2158–67.
31. Papp LA, Klein DF, Gorman JM. Carbon dioxide hypersensitivity, hyperventilation, and panic disorder. Am J Psychiatry 1993;150:1149–57.
32. Han JN, Stegen K, Schepers R, et al. Subjective symptoms and breathing pattern at rest and following hyperventilation in anxiety and somatoform disorders. J Psychosom Res 1998;45:519–32.
33. Herdman SJ, Tusa RJ. Assessment and treatment of patients with benign paroxysmal positional vertigo. In: Herdman SJ, editor. Vestibular rehabilitation. Philadelphia: FA Davis; 2000. p. 451–75.
34. Barber HO, Stockwell CW. Manual of electronystagmography. 2nd ed. St. Louis: CV Mosby; 1980.
35. Cutrer FM, Baloh RW. Migraine-associated dizziness. Headache 1992;32:300–4.
36. Johnson GD. Medical management of migraine-related dizziness and vertigo. Laryngoscope 1998;108 Suppl 85:1–30.
37. Fitzgerald G, Hallpike CS. Studies in human vestibular function: I. Observations of the directional preponderance of caloric nystagmus resulting from cerebral lesions. Brain 1942;65:115.
38. Jacobson GP, Calder JA, Shepherd VA, et al. Reappraisal of the monothermal warm caloric screening test. Ann Otol Rhinol Laryngol 1995;104:942–5.

39. Torok N. Differential diagnosis of the caloric nystagmus. Equilibrium Res 1973;3:70–9.
40. Wall C. The sinusoidal harmonic acceleration rotary chair test. Theoretical and clinical basis. Neurol Clin 1990;8:269–85.
41. Waespe W, Cohen B, Raphan T. Dynamic modification of the vestibulo-ocular reflex by the nodulus and uvula. Science 1985;228:199–202.
42. Stockwell CW, Bojrab DI. Background and technique of rotational testing. In: Jacobson GP, Newman CW, Kartush JM, editors. Handbook of balance function testing. St. Louis: Mosby Year Book; 1993. p. 237–48.
43. Stockwell CW, Bojrab DI. Interpretation and usefulness of rotational testing. In: Jacobson GP, Newman CW, Kartush JM, editors. Handbook of balance function testing. St. Louis: Mosby Year Book; 1993. p. 249–60.
44. Allum JH, Shepard NT. An overview of the clinical use of dynamic posturography in the differential diagnosis of balance disorders. J Vestib Res 1999;9:223–52.
45. Nashner LM. Practical biomechanics and physiology of balance. In: Jacobson GP, Newman CW, Kartush JM, editors. Handbook of balance function testing. St. Louis: Mosby-Year Book; 1993. p. 261–79.
46. Nashner LM. Computerized dynamic posturography. In: Jacobson GP, Newman CW, Kartush JM, editors. Handbook of balance function testing. St. Louis: Mosby-Year Book; 1993. p. 280–307.
47. Nashner LM. Computerized dynamic posturography: clinical applications. In: Jacobson GP, Newman CW, Kartush JM, editors. Handbook of balance function testing. St. Louis: Mosby-Year Book; 1993. p. 308–34.
48. Shepard NT, Schultz A, Alexander NB, et al. Postural control in young and elderly adults when stance is challenged: clinical versus laboratory measurements. Ann Otol Rhinol Laryngol 1993;102:508–17.
49. Allum JHJ, Huwiler M, Honegger F. Objective measures of non-organic vertigo using dynamic posturography. In: Taguchi K, Igarashi M, Mori S, editors. Vestibular and neural front: proceedings of the 12th International Symposium on Posture and Gait. Amsterdam: Elsevier; 1994. p. 51–5.
50. Cevette MJ, Puetz B, Marion MS, et al. Aphysiologic performance on dynamic posturography. Otolaryngol Head Neck Surg 1995;112:676–88.
51. Goebel JA, Sataloff RT, Hanson JM, et al. Posturographic evidence of nonorganic sway patterns in normal subjects, patients and suspected malingerers. Otolaryngol Head Neck Surg 1997;117:293–302.
52. Lawson GD, Shepard NT, Oviatt DL, Wang Y. Electromyographic responses of lower leg muscles to upward toe tilts as a function of age. J Vestib Res 1994;4:203–14.
53. Dichgans J, Diener HC. The use of short- and long-latency reflex testing in leg muscles of neurological patients. In: Struppler A, Weindl A, editors. Clinical aspects of sensory-motor integration: implications for neurological patients. New York: Springer-Verlag; 1987. p. 165–75.
54. Shumway-Cook A, Horak FB. Assessing the influence of sensory interaction on balance. Suggestion from the field. Phys Ther 1986;66:1548–50.
55. Beynon GJ, Jani P, Baguley DM. A clinical evaluation of head impulse testing. Clin Otolaryngol 1998;23: 117–22.
56. Ruckenstein MJ, Staab JP, Solomon D, Shepard NT. The BSI-53: its use in identifying patients with psychogenic dizziness. Annual Meeting of the American Academy of Otolaryngology-Head and Neck Surgery Foundation, Washington, DC, September 27, 2000.
57. Arriaga MA, Brackmann DE. Differential diagnosis of primary petrous apex lesions [published erratum appears in Am J Otol 1992;13:297]. Am J Otol 1991;12:470–4.
58. Jackler RK, Parker DA. Radiographic differential diagnosis of petrous apex lesions. Am J Otol 1992; 13:561–74.
59. Curtin HD, Som PM. The petrous apex. Otolaryngol Clin North Am 1995;28:473–96.
60. Muckle RP, De la Cruz A, Lo WM. Petrous apex lesions. Am J Otol 1998;19:219–25.
61. Jackler RK, Dillon WP. Computed tomography and magnetic resonance imaging of the inner ear. Otolaryngol Head Neck Surg 1988;99:494–504.
62. Minor LB, Solomon D, Zinreich JS, Zee DS. Sound- and/or pressure-induced vertigo due to bone dehiscence of the superior semicircular canal. Arch Otolaryngol Head Neck Surg 1998;124:249–58.
63. Doyle KJ. Is there still a role for auditory brainstem response audiometry in the diagnosis of acoustic neuroma? Arch Otolaryngol Head Neck Surg 1999; 125:232–4.
64. Ruckenstein MJ, Cueva RA, Morrison DH, Press G. A prospective study of ABR and MRI in the screening for vestibular schwannomas. Am J Otol 1996;17:317–20.
65. Shelton C, Harnsberger HR, Allen R, King B. Fast spin echo magnetic resonance imaging: clinical

application in screening for acoustic neuroma. Otolaryngol Head Neck Surg 1996;114:71–6.
66. Linker SP, Ruckenstein MJ, Acker J, Gardner G. An accurate, cost-effective approach for diagnosing retrocochlear lesions utilizing the T2-weighted, fast-spin echo magnetic resonance imaging scan. Laryngoscope 1997;107:1525–9.
67. McCabe BF. Autoimmune sensorineural hearing loss. Ann Otol Rhinol Laryngol 1979;88:585–9.
68. Lake GM, Sismanis A, Ariga T, et al. Antibodies to glycosphingolipid antigens in patients with immune-mediated cochleovestibular disorders. Am J Otol 1997;18:175–8.
69. Gottschlich S, Billings PB, Keithley EM, et al. Assessment of serum antibodies in patients with rapidly progressive sensorineural hearing loss and Meniere's disease. Laryngoscope 1995;105: 1347–52.
70. Bloch DB, San Martin JE, Rauch SD, et al. Serum antibodies to heat shock protein 70 in sensorineural hearing loss. Arch Otolaryngol Head Neck Surg 1995;121:1167–71.
71. Peltomaa M, Pyykkö I, Seppälä I, Viljanen M. Lyme borreliosis — an unusual cause of vertigo. Auris Nasus Larynx 1998;25:233–42.
72. Peltomaa M, Pyykkö I, Seppälä I, Viljanen M. Lyme borreliosis, an etiologic factor of sensorineural hearing loss? Eur Arch Otorhinolaryngol. [In press]
73. Whitney SL, Maureen MR. Efficacy of vestibular rehabilitation. Otolaryngol Clin North Am 2000;33: 659–73.
74. Clendaniel RA. Outcome measures for assessment of treatment of the dizzy and balance disorder patient. Otolaryngol Clin North Am 2000;33:519–34.
75. Shepard NT, Telian SA. Programmatic vestibular rehabilitation. Otolaryngol Head Neck Surg 1995; 112:173–82.
76. Smith-Wheelock M, Shepard NT, Telian SA. Physical therapy program for vestibular rehabilitation. Am J Otol 1991;12:218–25.

CHAPTER 7

Imaging of the Temporal Bone

Galdino E. Valvassori, MD

The temporal bone is unique in the human body because it contains, in the small volume of a cubic inch, a concentration of vital osseous and membranous structures surrounded by a more or less extensive system of pneumatic cells. Because of the different densities of its bony components and of the air- and fluid-filled spaces around and within them, the temporal bone lends itself to accurate visualization and assessment by various imaging procedures.

Conventional radiography, computed tomography (CT), magnetic resonance imaging (MRI), and arteriography are the techniques currently used to study the temporal bone and auditory-vestibular pathways. The otolaryngologist should know how to select the proper procedure for the anatomic area and the clinical problem under investigation.

CONVENTIONAL RADIOGRAPHY

Today, the use of conventional radiography is limited to evaluation of the mastoid pneumatization and assessment of the position and integrity of cochlear implant electrodes. The latter cannot be established by tomographic techniques because the wires often are visualized in several contiguous sections; therefore, their continuity cannot be demonstrated. Only three projections are of practical interest: the lateral or Schüller's, the frontal or transorbital, and the oblique or Stenvers'. The other special projections have historical significance but no useful clinical application.

Schüller's or Rungstrom's Projection

The Schüller's projection is a lateral view of the mastoid obtained with a cephalocaudad angulation of the x-ray beam of 25 to 30 degrees. The patient's head is turned so that the sagittal plane of the skull becomes parallel to the tabletop and the side under examination is closer to the film. Proper centering is obtained by placing the external auditory meatus of the side to be examined 1 cm above the center of the film or of the tabletop.

The extent of the pneumatization of the mastoid, distribution of the air cells, degree of aeration, and status of the trabecular pattern are the main features of this projection (Figure 7–1). The anterior plate of the vertical portion of the sigmoid sinus groove (corresponding to the most lateral part of the posterior aspect of the petrous pyramid) casts an almost vertical line, slightly concave posteriorly in its upper portion, superimposed on the air cells. At its upper extremity, this line joins another line that slopes gently forward and downward to form the sinodural angle of Citelli. The latter line is produced by the superior aspect of the lateral portion of the petrous pyramid. The more medial portion of the superior petrous ridge, from the arcuate eminence to the apex, has been displaced downward by the angulation of the x-ray beam and casts a line that extends forward and downward, crossing the epitympanic area and, more anteriorly, the neck of the mandibular condyle. Above this line, the upper portion of the attic with the head of the malleus is usually visible. Finally, the temporomandibular joint is outlined.

Transorbital Projection

This view can be obtained with the patient's face either to or away from the film. The patient's head is flexed on the chin until the orbitomeatal line is perpendicular to the tabletop. For better details, each side should be obtained separately, and the central x-ray beam should be directed at the center of the orbit of the side under examination and perpendicular to the film.

The petrous apex is outlined clearly but foreshortened because of its obliquity to the plane of the films. The internal auditory canal is visualized in its full length as a horizontal band of radiolucency

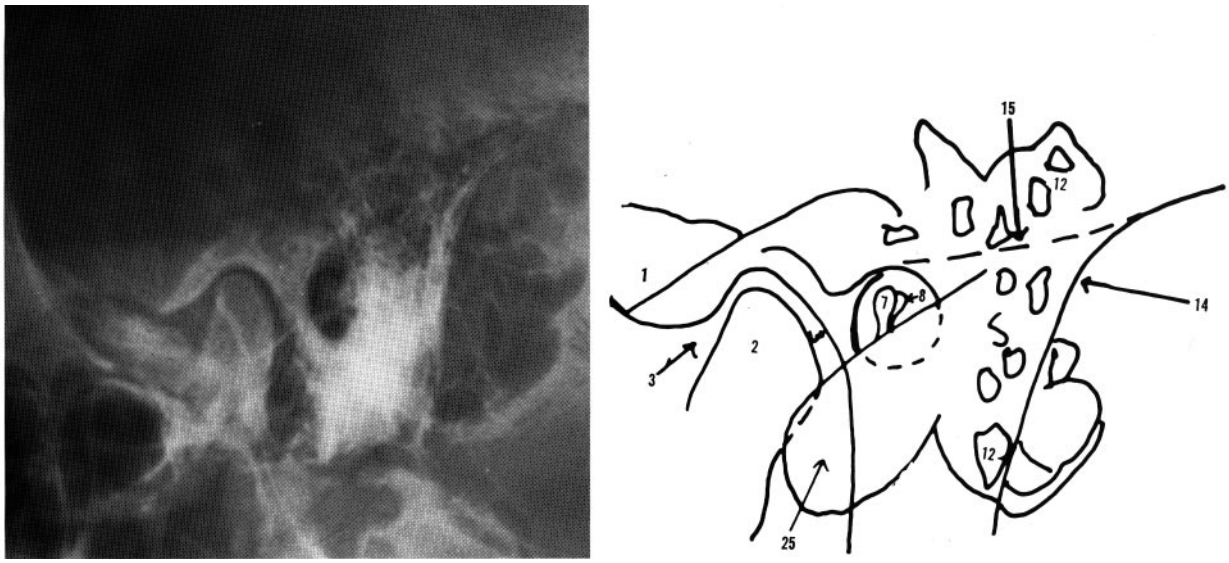


FIGURE 7–1. Schüller's projection: (1) root of the zygoma, (2) condyle of the mandible, (3) temporomandibular joint, (7) malleus, (8) incus, (12) air cells, (14) anterior plate of the sigmoid sinus, (15) dural plate, (25) petrous apex.

extending through the petrous pyramid (Figure 7–2). At the medial end of the canal, the free margin of the posterior wall casts a well-defined and smooth margin, concave medially. Often the radiolucent band of the internal auditory canal seems to extend medially to the lip of the posterior wall into the petrous apex. This band is not caused by the internal auditory canal but is produced by the medial extension of the upper and lower lips of the porus (opening) of the canal and by the interposed groove. Lateral to the internal auditory canal, the radiolucencies of the vestibule and of the superior and horizontal semicircular canals are usually detectable. The apical and middle coils of the cochlea are superimposed on the lateral portion of the internal auditory canal, whereas the basal turn is visible underneath the canal and the vestibule.

Stenvers' Projection

The patient is positioned facing the film, with the head slightly flexed and rotated 45 degrees toward

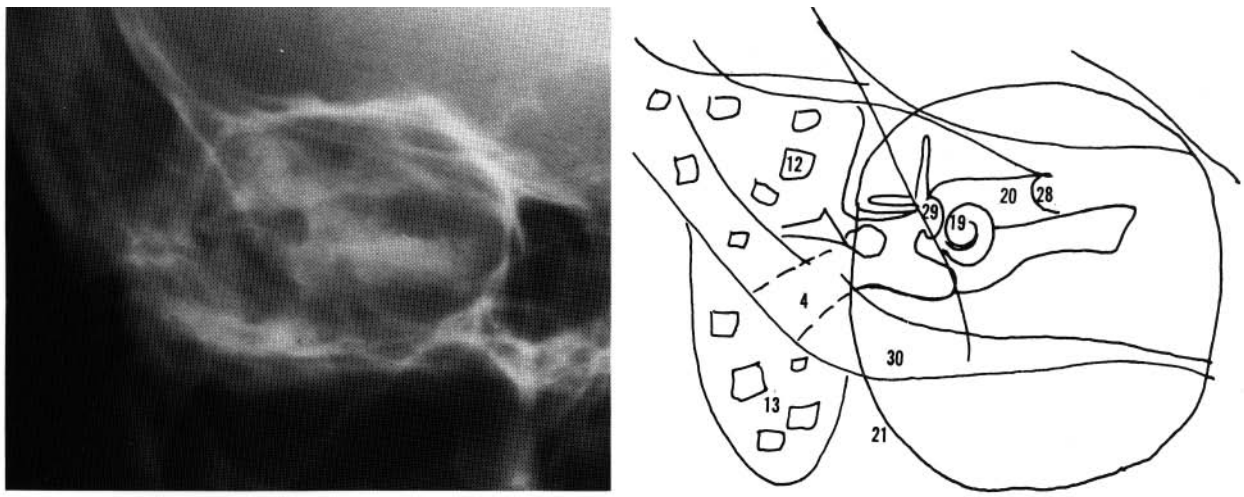


FIGURE 7–2. Transorbital projection: (4) external auditory canal, (12) air cells, (13) mastoid process, (19) cochlea, (20) internal auditory canal, (21) orbital rim, (28) medial lip of the posterior wall of the internal auditory canal, (29) vestibule, (30) base of the skull.

the side opposite the one under examination. The lateral rim of the orbit of the side under investigation should lie in close contact with the tabletop. The x-ray beam is angulated 14 degrees caudad.

The entire petrous apex is visualized in its full length lateral to the orbital rim (Figure 7–3). The porus of the internal auditory canal seen on face appears as an oval-shaped radiolucency open medially and limited laterally by the free margin of the posterior canal wall. Lateral to the porus, the internal auditory canal appears foreshortened. The vestibule and semicircular canals, especially the posterior, which lies in this projection in a plane parallel to the film, are usually recognizable. On the outside, the entire mastoid is outlined, with the mastoid process free from superimpositions.

COMPUTED TOMOGRAPHY

Computed tomography is a radiographic technique that allows measurement of small absorption coefficients and differentials not recognizable by previously available recording or displaying systems.

The scan is initiated at a chosen level, and the x-ray tube, collimated to a thin or pencil beam, rotates around the patient. The transmitted x-rays are picked up by detectors arrayed along the circumference of the tube trajectory, converted into electronic current, amplified, and transmitted to the computer for storing and processing. The computer analyzes these data and develops an image on a 360 by 360 picture element (pixel) matrix, where the brightness of each point is proportional to the attenuation coefficient.

The study is often performed after intravenous injection of iodinated contrast agents that produce an increase in density value or enhancement of several anatomic structures and pathologic tissues.

Recently, conventional CT has been replaced by helical or spiral CT, which allows rapid acquisition of volumetric data of the part of the body under examination. Images are acquired at an angled plane of section and then reconstructed by interpolating the volumetric data set as two-dimensional or high-quality three-dimensional reformations. The continuous acquisition of images is made possible by replacing the former electric cabling of the gantry with slip-ring design that allows continuous rotation of the source detector assembly and by the use of high heat capacity x-ray tubes. The advantage of spiral CT includes the elimination of respiratory misregistration, a decrease of motion artifact, and an obvious improvement in patient comfort owing to the shortening of the examination time. This is particularly important in children, whether they have been administered anesthesia or sedation, and in older patients who have difficulty in extending the head. Reconstruction of axial data in other planes produces satisfactory images provided that a collimation as small as 1 or 0.5 mm is used. A scanning time of less than 1 minute is sufficient to cover the mastoid and petrous portions of the temporal bone.

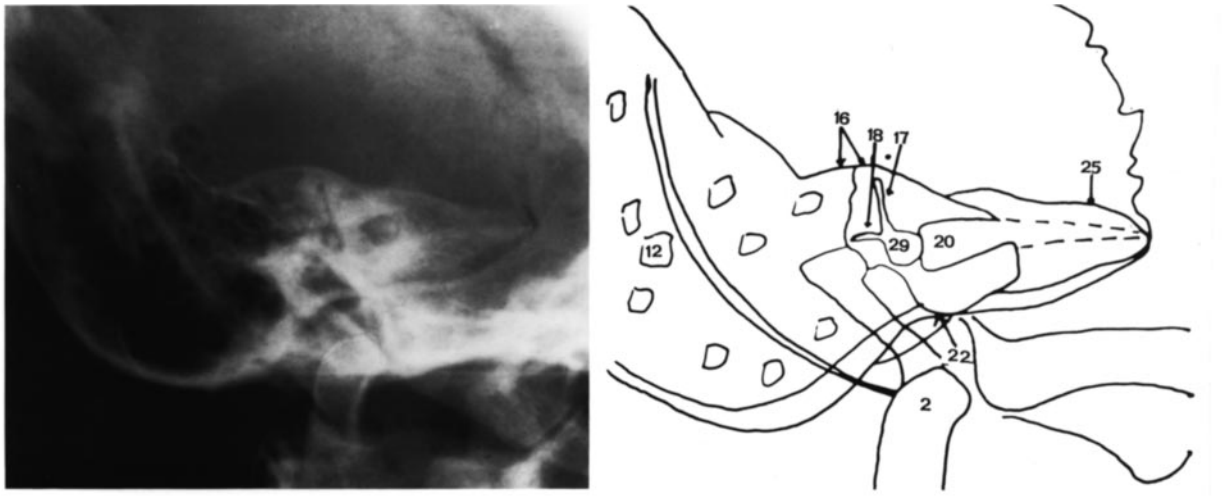


FIGURE 7–3. Stenvers' projection: (2) condyle of the mandible, (12) air cells, (16) arcuate eminence, (17) superior semicircular canal, (18) horizontal semicircular canal, (20) internal auditory canal, (22) basilar turn of the cochlea, (25) petrous apex, (29) vestibule.

New multidetector scanners will further decrease the scanning time. The CT images provide exquisite bony details and excellent demonstration of soft tissue density within the air space of the mastoids, external auditory canal, and middle ear but very limited identification of the type of substance producing the abnormal density. For instance, the density of cholesteatoma is identical to that of a tumor or granulation tissue and even of fluid.

The densitometric readings obtained with a cursor of variable size are often unreliable because of partial volume averaging within the small cavities of the ear. Time-density curves obtained by rapid sequential images of a preselected section after injection of a bolus of contrast are useful for the identification of vascular masses. At present, MRI provides a far more precise identification of the type of tissue within the air spaces and abnormal cavities in the temporal bone. Three-dimensional reconstruction of the temporal bone has not added diagnostic information but is of great value in surgical planning in large lesions of the base of the skull.

Computed Tomographic Projections of the Temporal Bone

Horizontal or Axial Projection This is the basic projection of the CT study.[1] It is comfortable for the patient, who lies supine on the table, and is easy to obtain and reproduce. It allows a good demonstration of the external, middle, and inner ears except for the structures parallel to the plane of section, such as the tegmen (Figure 7–4).

Coronal or Frontal Projection The patient lies on the table, either prone or supine, with the head extended. The gantry of the scanner is often tilted to compensate for an incomplete extension of the head. This projection is indispensable to complement the axial sections but is often difficult to obtain, particularly in young children and older people (Figure 7–5).

Twenty-Degree Coronal Oblique Projection This projection is a modification of the coronal projecton for the study of the medial wall of the tympanic cavity. The medial or labyrinthine wall of the middle ear forms an angle, open posteriorly, of 15 to 25 degrees with the midsagittal plane of the skull. The patient is first positioned as for the coronal projection, and the head is then rotated 20 degrees toward the side under examination so that the medial wall becomes perpendicular to the plane of section. For this projection, we position the patient prone with the head extended. This projection is particularly useful for the study of the oval window, promontory, and tympanic segment of the facial canal. It should be used in all patients with otosclerosis (Figure 7–6).

Sagittal or Lateral Projection It is impossible or extremely difficult to obtain direct sagittal sections. However, computer reconstructed sagittal images can be obtained from the raw data collected for the horizontal sections. This is particularly true if fast or spiral CT is used, and thin sections at 0.5 or 1 mm increments are obtained. These images are satisfactory for the demonstration of the mastoid segment of the facial canal and of the vestibular aqueduct (Figure 7–7).

MAGNETIC RESONANCE IMAGING

Magnetic resonance imaging is an imaging modality capable of producing cross-sections of the human body in any plane without exposing the patient to ionizing radiation. Magnetic resonance images are obtained by the interaction of hydrogen nuclei or protons of the human body, high magnetic fields, and radiofrequency pulses. The strength of the MRI signal to be converted into imaging data depends on the concentration of the free hydrogen nuclei or protons and on two magnetic relaxation times, T_1 and T_2, which are tissue specific. One of the characteristics of MRI is the possibility of changing appearance and therefore information of the images by changing the contribution of the T_1 and T_2 relaxation times. This is accomplished by varying the time between successive pulses (TR or by repetition time) and the time that the emitted signal or echo is measured after the pulse (TE or echo time).

Magnetic resonance imaging has undergone multiple changes and refinements to increase definition of the images and to decrease the acquisition time of the image.[2] First, the use of surface coils placed adjacent to the area of interest has increased the signal-to-noise ratio and consequently improved the image. Shorter acquisitions have been obtained by shortening the TR, using pulses with flip angles smaller than 90 degrees (gradient echo imaging), reducing the number of phase-encoded steps, and

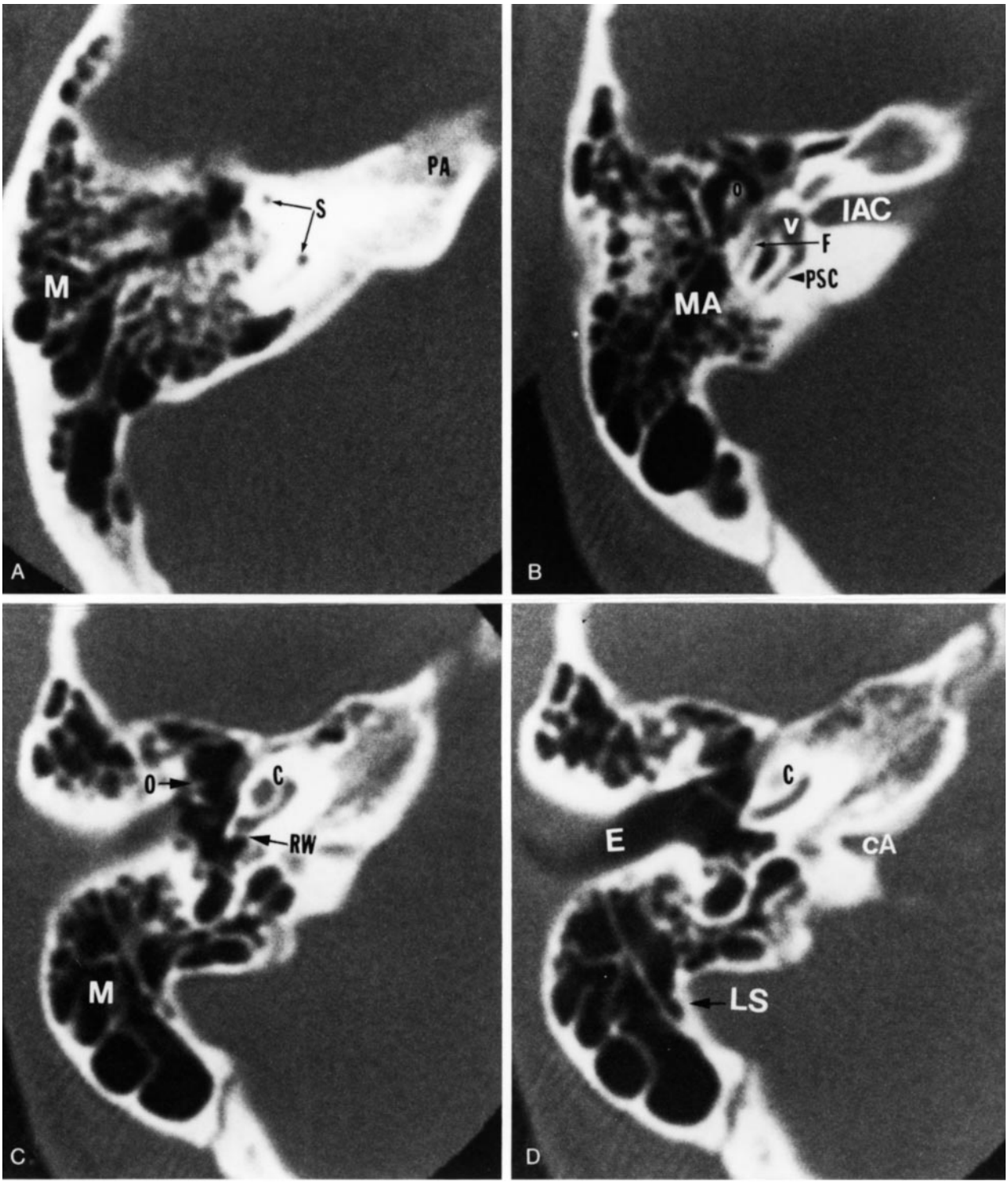


FIGURE 7–4. Horizontal computed tomographic sections of a normal right temporal bone, in sequence from top to bottom: *A*, M = mastoid; S = superior semicircular canal; PA = petrous apex. *B*, MA = mastoid antrum; O = ossicles; V = vestibule; F = facial nerve canal; IAC = internal auditory canal; PSC = posterior semicircular canal. *C*, C = cochlea; RW = round window. *D*, E = external auditory canal; CA = cochlear aqueduct; LS = lateral sinus plate.

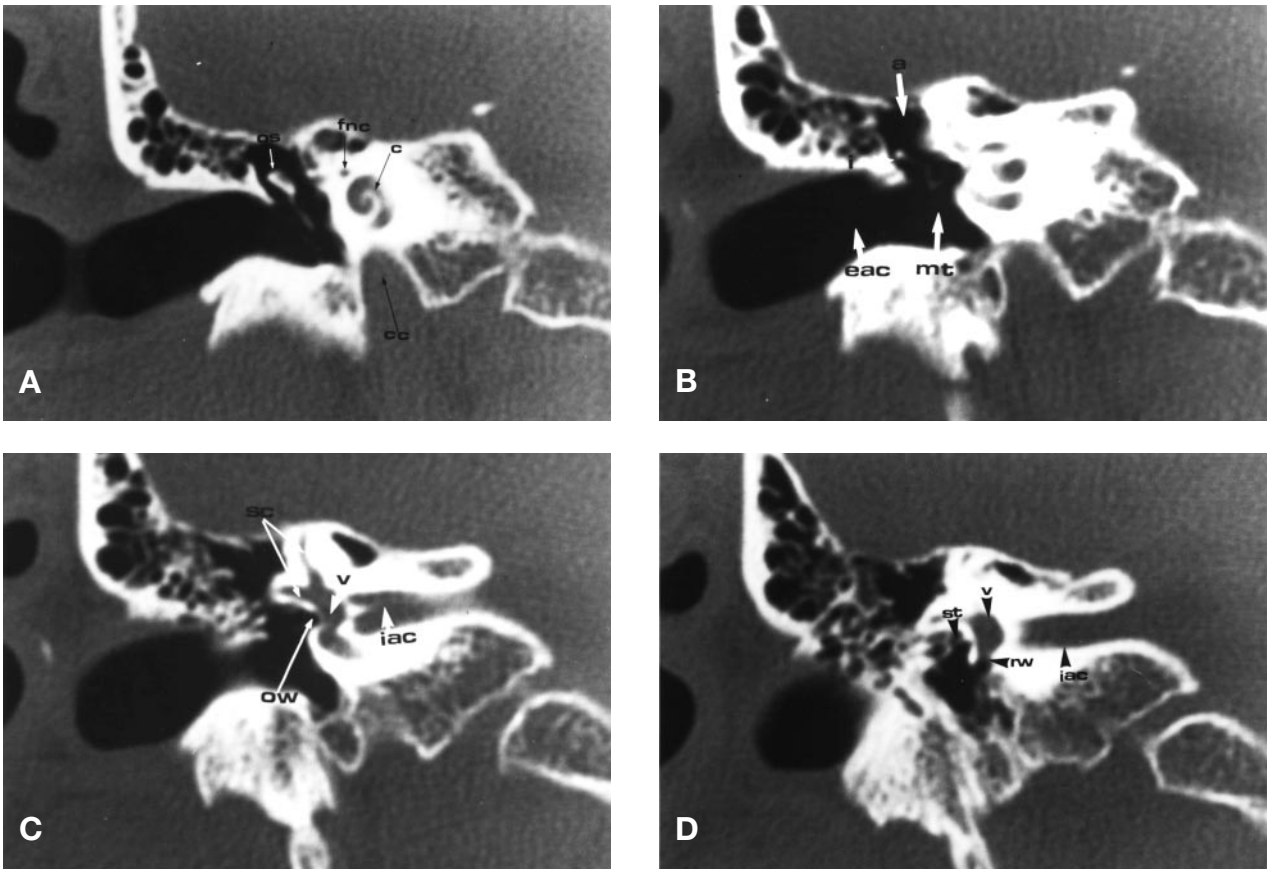


FIGURE 7–5. Coronal computed tomographic sections of a normal right temporal bone in sequence from front to back. *A*, OS = ossicles; C = cochlea; fnc = facial nerve canal; CC = carotid canal. *B*, eac = external auditory canal; a = attic; mt = mesotympanum; i = incus. *C*, OW = oval window; V = vestibule; SC = semicircular canals; iac = internal auditory canal. *D*, rw = round window; st = sinus tympani.

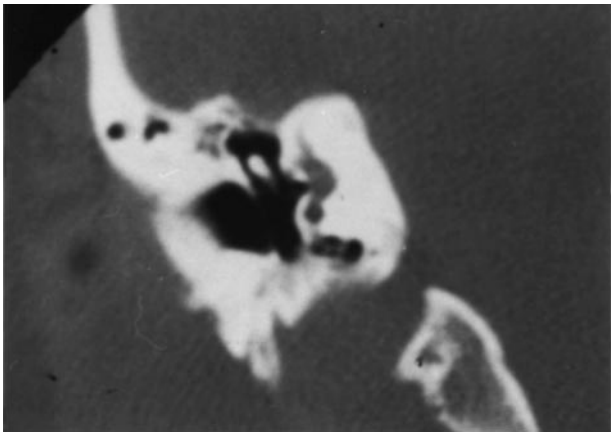


FIGURE 7–6. Twenty-degree coronal oblique section of a normal right temporal bone. The oval window forms a well-defined opening in the lateral wall of the vestibule underneath the horizontal semicircular canal and above the promontory.

increasing the data collected per excitation. The latter technique, known as fast spin echo (FSE), uses multiple echoes (4 to 16 for an excitation), therefore reducing the number of excitations necessary for forming an image.

Examination is performed with the patient supine and the plane extending from the tragus to the inferior orbital rim perpendicular to the tabletop. Different projections are obtained by changing the orientation of the magnetic field gradients without moving the patient's head. Axial, coronal, and sagittal projections are obtained.

T_1 images obtained by a short TR and TE offer the best anatomic delineation. T_2 images obtained with long TR and TE better differentiate normal from pathologic tissues.

Air, cortical bone, and calcifications contain few free protons and therefore appear in the images

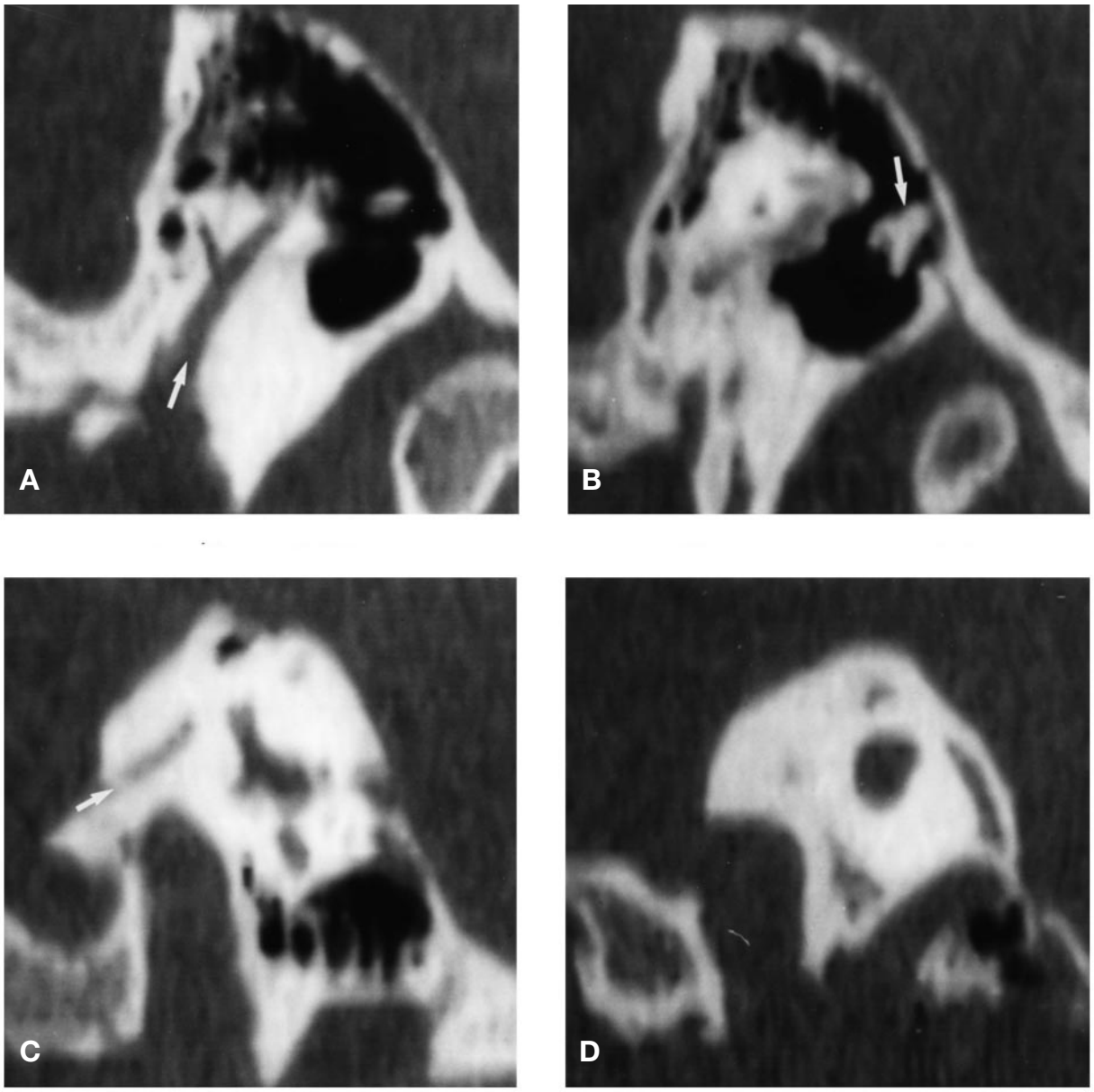


FIGURE 7–7. Reformatted sagittal computed tomographic sections of the right ear showing *A*, the external auditory canal, attic, and vertical segment of the facial canal (*arrow*); *B*, the ossicles (*arrow*) and the horizontal semicircular canal; *C*, the common crus, vestibular aqueduct (*arrow*), and the jugular fossa; *D*, the internal auditory canal.

as dark areas of no signal. Fat and body fluid are rich in free protons and produce signals of high intensity, fat in the T_1- and fluid in the T_2-weighted images. Blood vessels usually appear as areas of signal void because the stimulated protons of the circulating blood have moved out of the section before their emitted signals can be detected. Because cortical or nondiploic bone and air emit no signal, the normal mastoid, external auditory canal, and middle ear appear in the MRI as dark areas without pattern or structures within them. The petrous pyramid is equally dark except for a gray or white cast of the inner ear structures and internal auditory canal produced by the fluid within their lumens (Figure 7–8).

Pathologic processes are demonstrated by MRI whenever the hydrogen density and relaxation times of the pathologic tissues are different from the normal. The intravenous injection of ferromagnetic

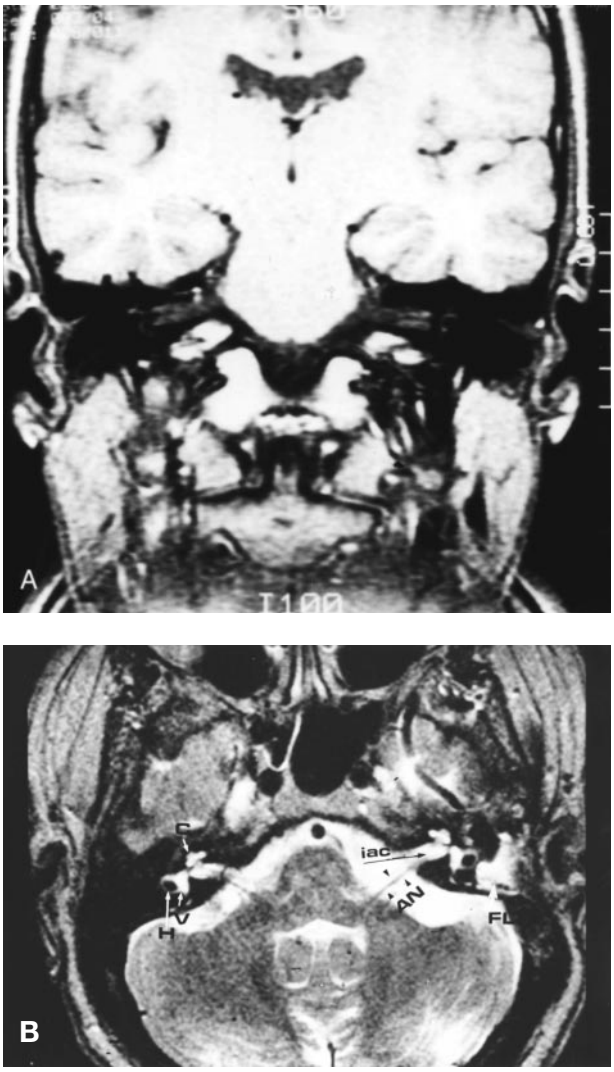


FIGURE 7–8. Magnetic resonance imaging sections of the normal temporal bone. *A*, Coronal T_1. *B*, Axial T_2 fast spin echo. The normal mastoid, external auditory canal, and middle ear cavity appear as dark areas of signal void. The internal auditory canals and inner ear structures are recognizable because of the signal produced by the fluid and the neural structures within their lumens. The cerebrospinal fluid has a low signal in T_1 but a high signal in T_2. AN = acoustic nerve; C = cochlea; FL = fluid in the mastoid and middle ear; H = horizontal semicircular canal; iac = internal auditory canal; V = vestibule.

contrast agents (gadolinium DTPA [diethylenetriamine pentaacetic acid] and gadolinium chelate) has improved the recognition and differentiation of pathologic processes. Because the contrast material does not penetrate the intact blood barrier, normal brain does not enhance except for structures such as the pituitary gland and several cranial nerves that lack a complete blood-brain barrier. Enhancement of brain lesions occurs whenever the blood-brain barrier is disrupted provided that there is sufficient blood flow to the lesions. Extra-axial lesions such a meningiomas and neuromas lack a blood-brain barrier and therefore undergo a strong enhancement.

Fluid, blood, and soft tissue masses within the temporal bone are identified readily by MRI as areas of abnormally high signal intensity. The exact location, extent, and involvement of bony structures such as the ossicles, scutum, and labyrinthine capsule cannot be detected, however. For this reason, CT remains the study of choice for the assessment of intratemporal pathology with the exception of the petrous apex.

The images shown in this chapter were obtained with a super-conducting magnet and a magnetic field of 15,000 gauss or 1.5 tesla.

ANGIOGRAPHY

Gradient echo techniques and flow-encoding gradients have enabled the development of magnetic resonance angiography (MRA). Time of flight angiography is a gradient echo technique in which the stationary tissues within the imaging plane are saturated with the magnetic field so that they will not produce a signal. Blood flowing within the same plane is unsaturated and will be the only tissue to produce a signal.

Phase-contrast angiography is acquired differently. Instead of saturating the stationary tissues with radiofrequency pulses, a bipolar gradient of magnetization is applied to the entire slice, first with a positive value and then with a negative value. In the stationary tissues, the two opposite gradients cancel each other out. In the flowing blood, however, the two opposite gradients cannot cancel each other out because the blood will have moved to a different plane in the region before the inverse gradient is applied.

The obtained slices are reconstructed into projection images, which can be rotated in different planes to separate vessels and eliminate superimposition. Magnetic resonance angiography of the intracranial vasculature has been particularly useful in the demonstration of aneurysms (Figure 7–9) in the region of the circle of Willis, arteriovenous malformation (AVM) (particularly a small dural AVM

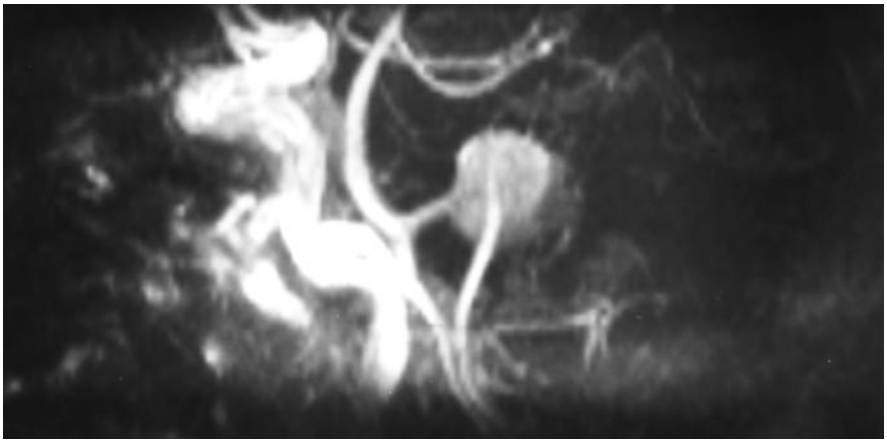


FIGURE 7–9. Magnetic resonance angiography demonstrating an aneurysm of the left vertebral artery. The aneurysm compresses the acoustic nerve and mimics an acoustic neurinoma.

that may not be visible on routine spin-echo images), and vaso-occlusive pathology, including dural sinus thrombosis. Magnetic resonance angiography of the extracranial circulation provides excellent information about the patency of the carotid and vertebral arteries. These vessels may be compressed or displaced by neck masses and their lumens stenosed or obstructed by thrombosis or atheromatous plaques.

The introduction of ultra-fast CT has opened the possibility of obtaining excellent angiographic images. In CT angiography, the continuous acquisition of the images allows the following of the transit of a rapidly intravenously injected bolus of contrast through the arteries and veins of the area under investigation. The reconstructed images can be rotated in the plane that best demonstrates the vessels.

Conventional angiography is seldom required for the diagnosis of vascular tumors or anomalies within or adjacent to the temporal bone. Arteriography is, however, mandatory for identifying the feeding vessels of lesions, usually glomus tumors, whenever embolization or surgical ligation is contemplated. Subtraction is necessary to delineate the vascular mass and feeding vessels, which are otherwise obscured by the density of the surrounding temporal bone. The injection should be performed in the common carotid artery to visualize both internal and external carotid circulation. A vertebral arteriogram may also be performed.

Retrograde jugular venography is rarely used for the diagnosis of a high jugular bulb or glomus tumor. The study is done by percutaneous puncture of the vein with a Seldinger needle.

PATHOLOGIC CONDITIONS

The pathologic conditions of the ear and temporal bone will be reviewed by anatomic location. Seven sites of involvement are considered: mastoid, external auditory canal, middle ear, inner ear, petrous pyramid, facial nerve, and cerebellopontine angle.

Mastoid

The development of the mastoids varies from person to person and, to a lesser degree, from side to side of the same individual. In some mastoids, the pneumatization is limited to a single antral cell; in others, it may extend into the mastoid tip and squama of the temporal bone and may even invade the adjacent zygoma and occipital bone. The nonpneumatized mastoid process may be made up of solid bone or contain spongy diploic spaces filled with fatty marrow. In MRI, the fatty marrow produces a high signal in the T_1 sequence that decreases in the T_2 and should not be confused with fluid or other pathologic processes that have a high signal in the T_2 sequence. One should also remember that it may be impossible by MRI to distinguish a small mastoid containing a larger aerated cell from a sclerotic mastoid since both produce an identical signal void.

Lateral Sinus The lateral or sigmoid sinus forms a shallow indentation on the posterior aspect of the mastoid. Occasionally, the sinus courses more anteriorly and produces a deep groove in the mastoid, best seen in the axial sections. In some cases, only a thin bony plate separates the sinus from the external auditory canal (Figure 7–10). Without this informa-

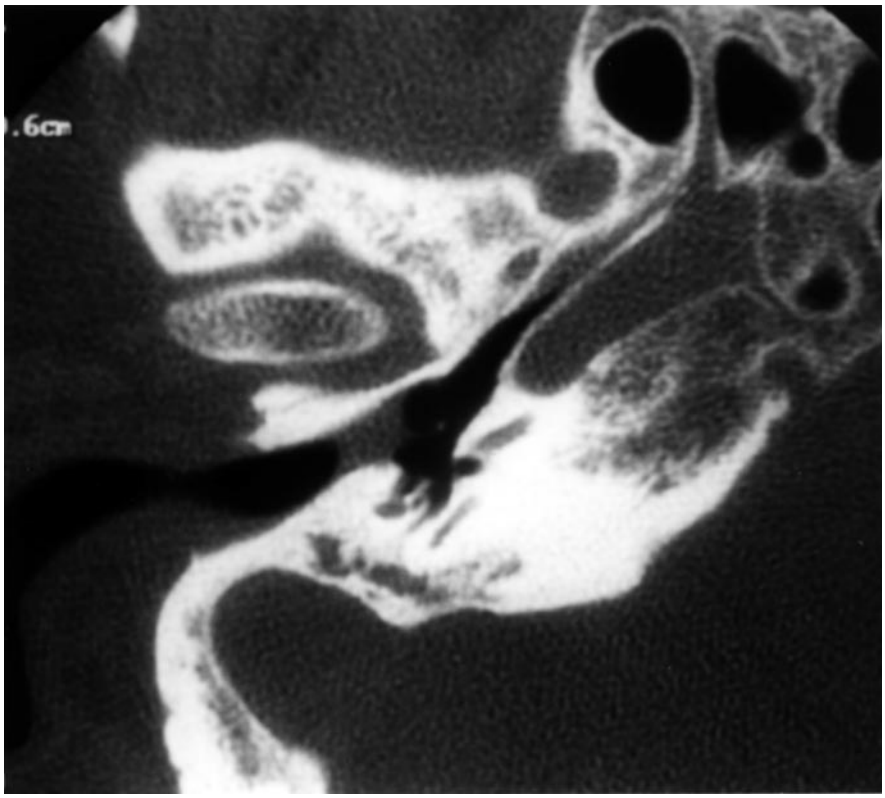


FIGURE 7–10. Anteriorly situated lateral sinus. Axial computed tomographic section. The sinus produces a deep groove in the posterior aspect of the mastoid and reaches the posterior wall of the external auditory canal.

tion, the otologist could enter the sinus with consequent severe bleeding and possible complications from sinus thrombosis.

Tegmen The tegmen of the mastoid and attic usually passes in a horizontal plane slightly lower than the arcuate eminence produced by the top of the superior semicircular canal. A depression of the tegmental plate is not unusual, particularly in patients with congenital atresia. As seen in the coronal sections, the floor of the middle cranial fossa deepens to form a groove lateral to the attic and to the labyrinth (Figure 7–11). The low-lying dura may cover the roof of the external auditory canal and, when the canal is

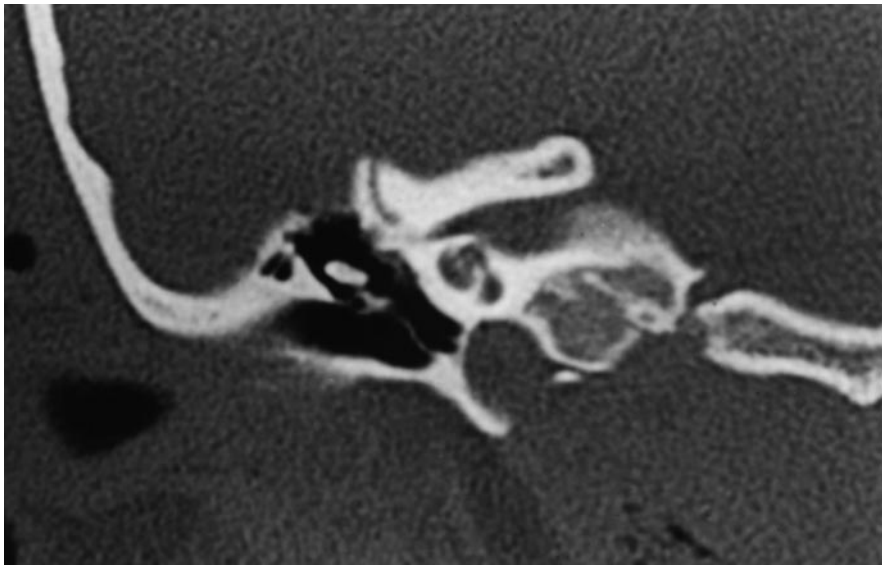


FIGURE 7–11. Low tegmen. Coronal computed tomographic section. The low-lying dura covers the roof of the external auditory canal and the middle cranial fossa deepens to form a groove lateral to the attic.

not developed, extend lateral to the mesotympanum. Unless aware of this feature, the otologist may penetrate the cranial cavity during surgery.

Acute Mastoiditis The early findings of acute mastoiditis are diffuse and homogeneous clouding of the middle ear cavity and mastoid air cells and often air-fluid levels within the air cells (Figure 7–12, A). With the progression of the process, the mastoid trabeculae first become demineralized and then destroyed, with formation of coalescent areas of suppuration.[3] Erosion of the cortex of the mastoid results in subperiosteal abscesses (Figure 7–12, B). Involvement of the posterior sinus plate often leads to thrombophlebitis of the lateral sinus and to posterior fossa abscesses (Figure 7–13), whereas erosion of the tegmen may lead to extension of the infection into the middle canal fossa with formation of extradural (epidural) or temporal lobe abscesses.

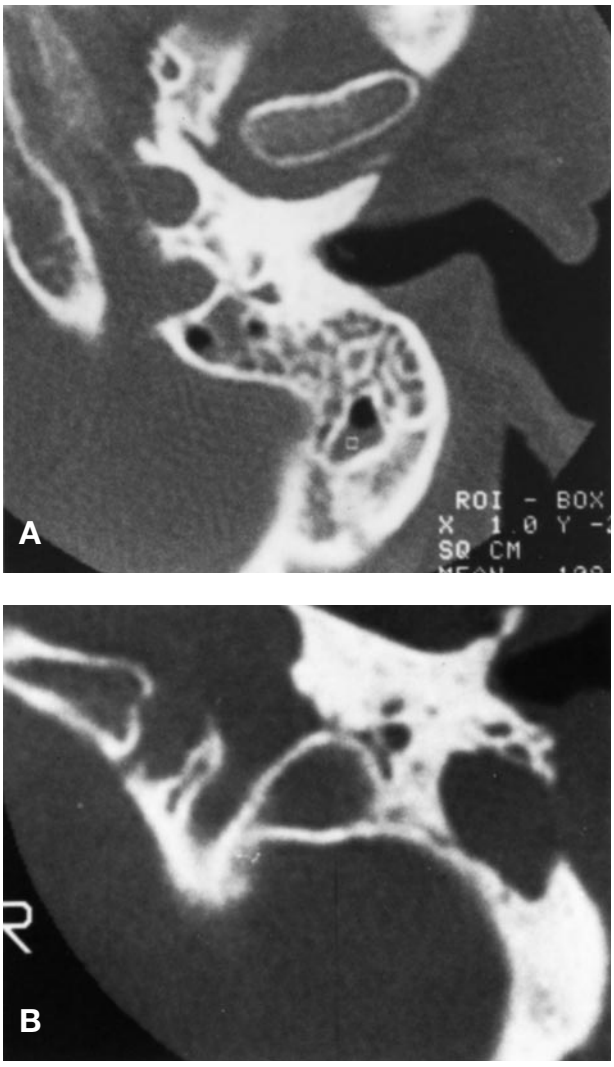


FIGURE 7–12. A, Acute mastoiditis. This axial computed tomographic (CT) section of the left mastoid shows clouding and air-fluid levels within the air cells. The trabecular pattern is intact. *B*, Coalescent mastoiditis, axial CT section of the left mastoid. A large coalescent cavity is noticed in the mastoid with erosion of the outer cortex and formation of a subperiosteal abscess.

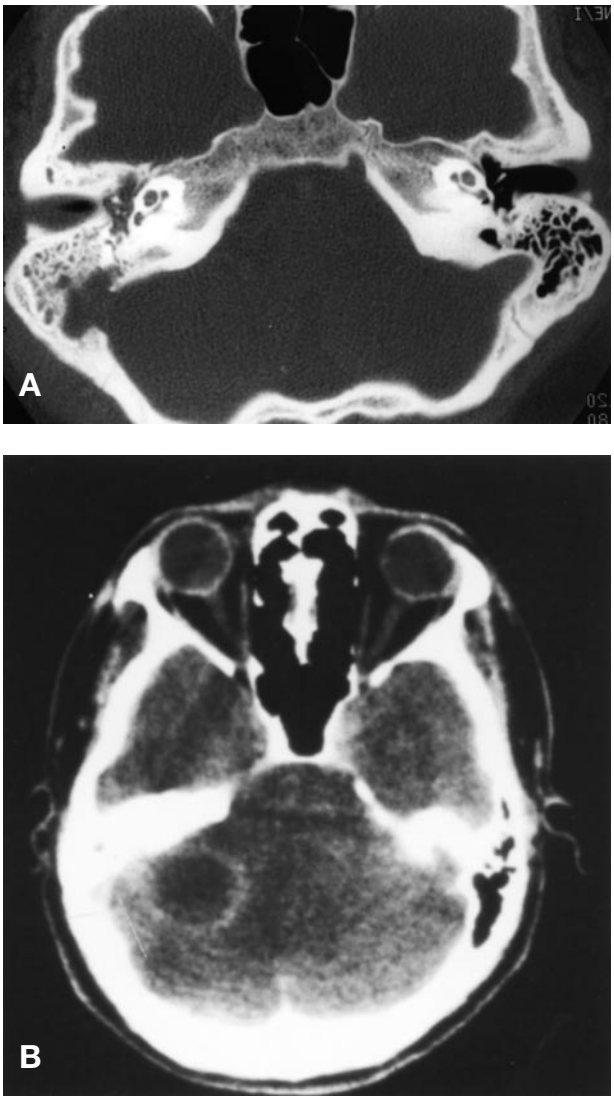


FIGURE 7–13. Acute mastoiditis with perisinus and epidural abscess. *A*, Axial computed tomographic (CT) section. There is a coalescent cavity in the posterior aspect of the mastoid with erosion of the sinus plate. *B*, Axial CT section after injection of contrast. A ring of enhancement surrounds an area of low density produced by the cerebellar abscess.

Chronic Mastoiditis The mastoid is often poorly pneumatized, and the mastoid antrum and mastoid air cells appear nonhomogeneously cloudy. Reactive new bony formation produces thickening of the trabeculae (Figure 7–14), which may lead to complete obliteration of the air cells (sclerotic mastoiditis).

Eosinophilic Granuloma Lytic lesions of variable size are observed in the mastoid and temporal squama. In the involved areas, the trabecular pattern is completely erased and the mastoid cortex may be thinned out, destroyed, or expanded (Figure 7–15).

External Auditory Canal

Computed tomography is the study of choice, but MRI may add useful information whenever the lesion extends outside the confines of the canal into the adjacent structures.

Congenital Malformations Microtia of varying degrees is often associated with dysplasia of the external auditory canal, although no direct correlation exists between the two anomalies. The degree of malformation of the external auditory canal varies from complete agenesis of the tympanic bone and canal (Figure 7–16) to stenosis of its lumen (Figure 7–17) and atresia of an otherwise well-developed canal by a bony plate, usually lateral. The small tissue tag and pit often observed in patients with atresia of the canal often have no topographic relationship to the underlying mastoid and middle ear.

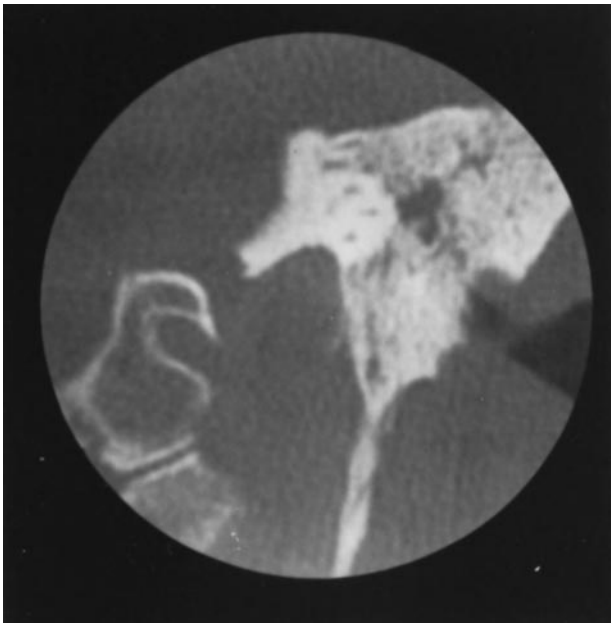


FIGURE 7–14. Chronic mastoiditis, coronal computed tomographic section. The trabeculae are thickened and the air cells cloudy and partially obliterated.

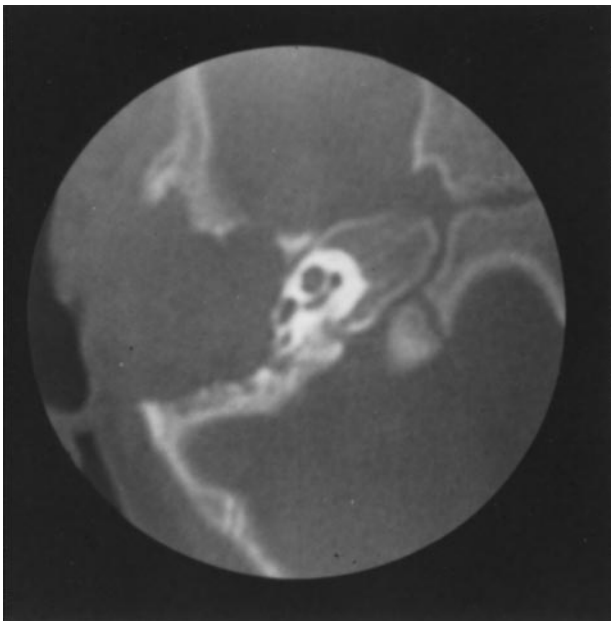


FIGURE 7–15. Eosinophilic granuloma, axial computed tomographic section of the right ear. A large lytic lesion involves the right mastoid, external auditory canal, and middle ear. A soft tissue mass fills the cavity.

Necrotizing External Otitis This condition, also known as malignant external otitis, is an acute osteomyelitis of the temporal bone that occurs in debilitated, diabetic, and immunosuppressed patients and is usually caused by the *Pseudomonas* bacterium. The infection starts in the external canal and spreads rapidly to the other portions of the temporal bone and adjacent areas. Computed tomography reveals erosion of the external canal, particularly of its floor, and stenosis of the lumen of the canal caused by soft tissue swelling (Figure 7–18). The infection may then spread inferiorly along the undersurface of the temporal bone to the facial nerve at the stylomastoid foramen, anteriorly to the temporomandibular fossa, posteriorly to the mastoid, and medially into the middle ear and petrous pyramid.

External Auditory Canal Cholesteatomas These lesions are caused either by blockage of the canal,

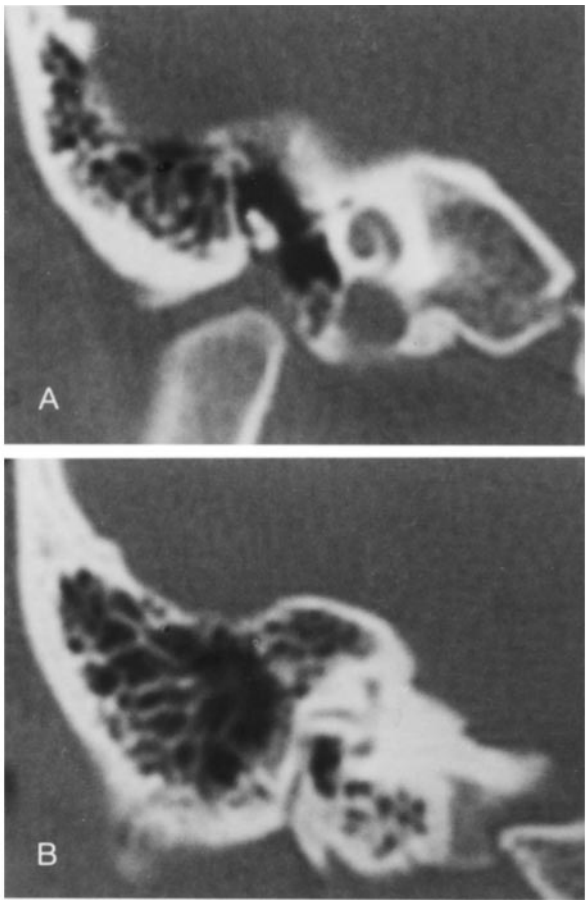


FIGURE 7–16. Agenesis of the right external auditory canal, coronal computed tomographic sections. *A*, The external auditory canal is absent and the mandibular condyle is displaced posteriorly and lies lateral to the atretic plate. The middle ear cavity is aerated and normal in size, but the head of the malleus and body of the incus appear malformed and fused. *B*, The mastoid is well pneumatized and clear. The vertical segment of the facial canal is rotated slightly outward.

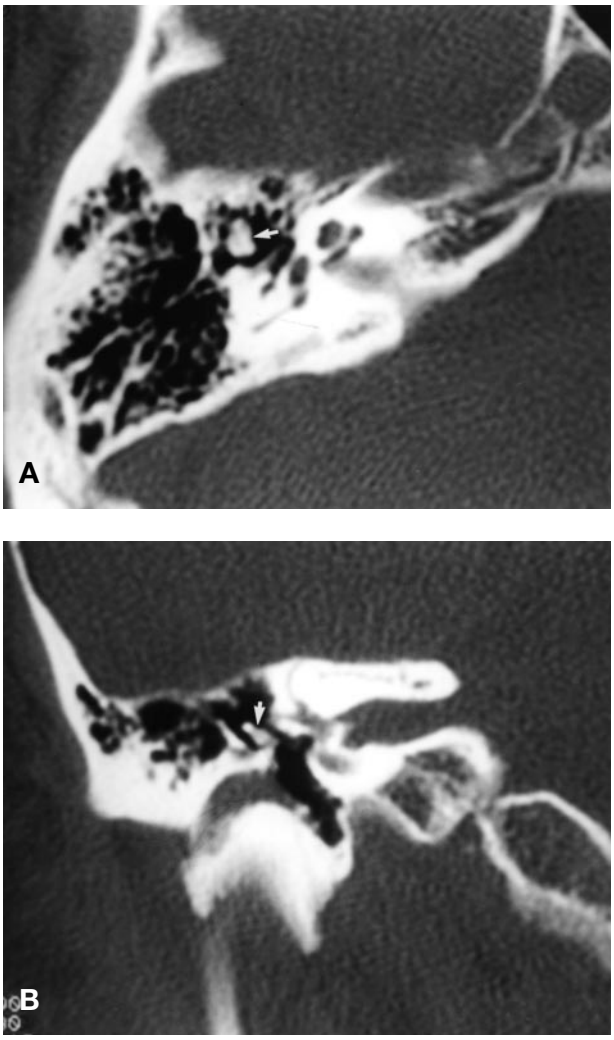


FIGURE 7–17. Stenosis of the right external auditory canal with malformation of the ossicular chain. *A*, Axial; *B*, coronal computed tomographic sections. The external auditory canal is markedly stenotic and closed laterally by a thin atretic plate. The head of the malleus and the body of the incus appear malformed and fused (*arrows*).

with consequent accumulation of epithelial debris (keratosis obliterans), or by dyskeratosis with localized accumulation of debris on the floor of the canal (invasive keratitis). In the first type, a soft tissue mass fills and expands the canal medial to the site of stenosis or obstruction (Figure 7–19). In the second type, the lumen of the canal is patent, but areas of bony erosion are demonstrated in the involved portion of the canal. When the cholesteatoma is large and erodes the annulus, it may extend into the middle ear cavity.

Carcinoma Carcinoma of the temporal bone usually arises in the external auditory canal. The CT findings vary with the extent of the lesion. In an early lesion, portions of the bony wall of the external auditory canal are eroded (Figure 7–20). When further destruction occurs, the tumor may spread anteriorly into the temporomandibular fossa; posteriorly into the mastoid, where it often reaches the facial canal; and medially into the middle ear, jugular fossa, petrous pyramid, and labyrinth. Extension into the mastoid and petrous pyramid causes a typ-

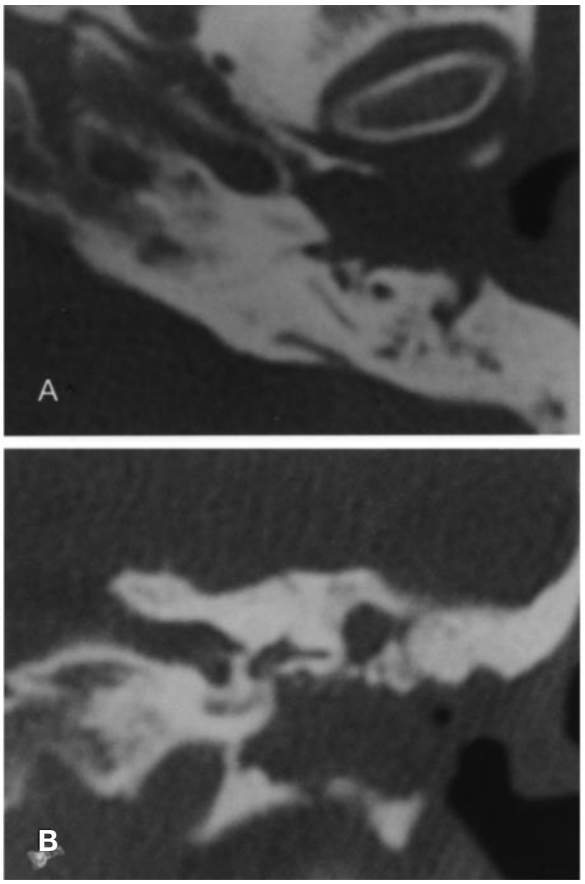


FIGURE 7–18. Necrotizing external otitis. Axial (*A*) and coronal (*B*) computed tomographic sections of the left ear. There is erosion of all walls of the external auditory canal with stenosis of the lumen of the canal caused by soft tissue swelling. The infection has eroded the lateral wall of the attic and spread into the middle ear.

ical moth-eaten appearance because of infiltration of the bone.

Middle Ear and Ossicular Chain

Computed tomography is the study for pathologic processes that arise in the middle ear and involve the ossicular chain.

Because granulation tissue enhances after the injection of paramagnetic agents, MRI with contrast is useful in some cases to differentiate granulation tissue from fluid and cholesteatoma, which have a similar density in the CT sections.

Congenital Malformations Malformations of the middle ear space vary from minor hypoplasia to

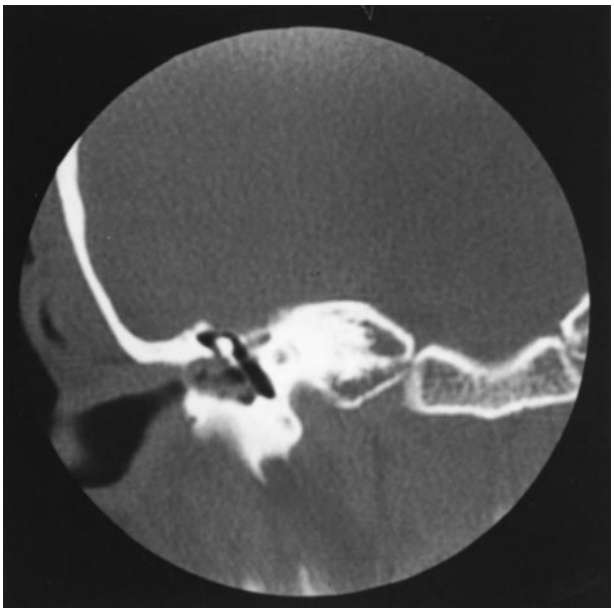


FIGURE 7–19. External auditory canal cholesteatoma, coronal section of the right ear. The cholesteatoma fills and obstructs the bony external auditory canal. The cholesteatoma erodes the inferior margin of the lateral wall of the attic and extends into the attic lateral to the ossicles.

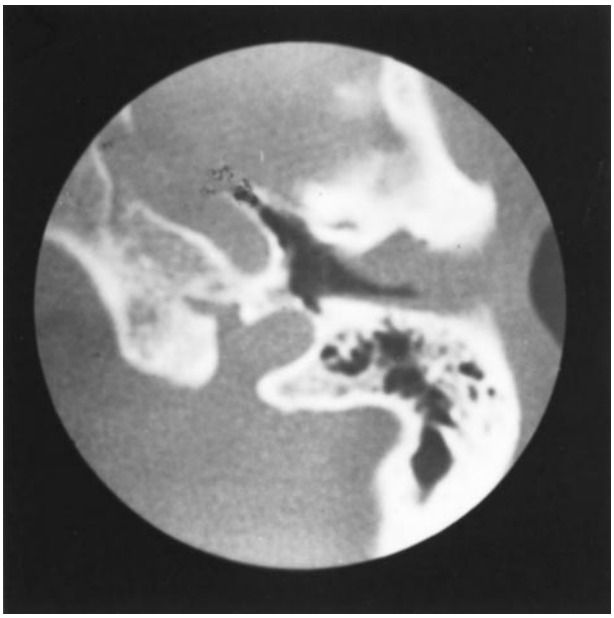


FIGURE 7–20. Carcinoma of the left external auditory canal, axial computed tomographic section. The anterior wall of the external auditory canal is eroded by an adjacent soft tissue mass. The lesion does not extend into the middle ear cavity.

complete agenesis. In the majority of cases with an atretic external auditory canal, the middle ear cavity is normal in size and aerated. The head of the malleus and body of the incus are often fused and fixed to the atretic plate at the level of the malleus neck (see Figure 7–16). If the middle ear cavity is grossly hypoplastic, the malleus and incus form a rudimentary amalgam that is often in an ectopic position (Figure 7–21). A congenital anomaly may be confined to the ossicular chain, which may be malformed or fixed.

Malformation of the stapes superstructure and fixation of the stapes footplate are not uncommon, isolated defects. Computed tomographic sections in 20-degree coronal oblique projection are useful for the study of the stapes and oval window.

The jugular bulb sometimes projects into the hypo- or mesotympanum. The bulb may be covered by a thin bony shell or may be exposed in the middle ear, often in contact with the medial surface of the tympanic membrane, thus mimicking a glomus tumor. The MRI appearance of a high jugular bulb may be misread as showing a glomus tumor because of the mixed signal within the bulb produced by turbulent flow. However, whereas a glomus tumor contains multiple punctate areas of signal void within a mass of medium or high signal, in a high jugular bulb, linear streaks of high and low signal are seen within the lumen of the bulb, usually paralleling its wall, owing to variations in flow velocity. In questionable cases, an MR venogram should be performed using a two-dimensional (2–D) time of flight technique sensitive to slow flow or a two-dimensional phase contrast using short velocity encoding.

The intratemporal segment of the carotid artery may take an ectopic course through the middle ear. In these cases, the CT examination shows a soft tissue mass extending throughout the entire length of the middle ear cavity to regain its normal position in the petrous apex (Figure 7–22). The proximal portion of the carotid artery, rather than being located underneath the cochlea, enters the temporal bone through a canal or defect in the floor of the posterior part of the hypotympanum. A carotid arteriogram, or MRA, may be used to confirm the anomalous course of the artery.

Trauma The middle ear cavity is usually involved in longitudinal fractures of the temporal bone. Computed tomography permits precise evaluation of the course of the fracture and status of the ossicular chain. A fracture line may disappear at a certain level only to reappear a few millimeters distant. This apparent gap is not caused by interruption of the fracture line but rather by a change in its course so that the fracture becomes invisible in some of the

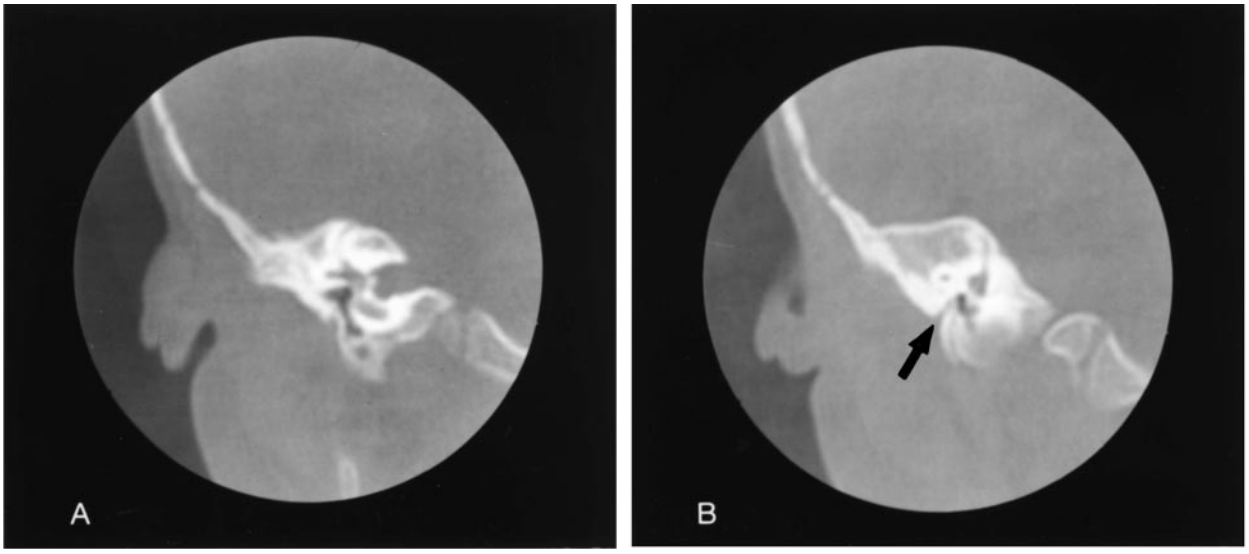


FIGURE 7–21. Congenital malformation of the right external and middle ears, coronal computed tomographic sections. *A*, Agenesis of the external auditory canal with hypoplasia of the attic and rudimentary ossicular mass. *B*, Notice the lack of development of the mastoid pneumatization and the outward rotation of the vertical segment of the facial canal (*arrow*).

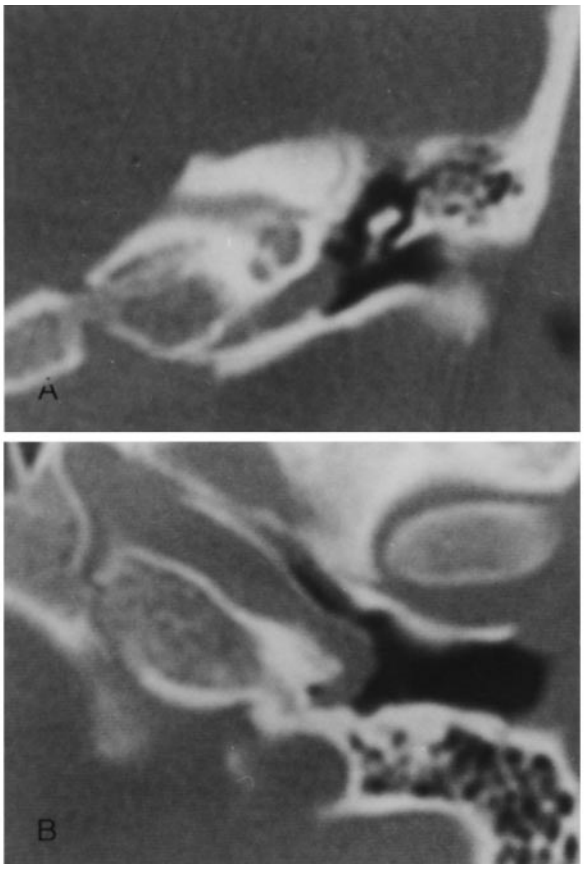


FIGURE 7–22. Ectopic left internal carotid canal. Coronal (*A*) and axial (*B*) computed tomographic sections. The ectopic carotid artery courses throughout the entire length of the middle ear cavity. Notice the absence of the proximal portion of the carotid canal normally seen underneath the cochlea and of the bony wall dividing the anterior part of the mesotympanum from the horizontal segment of the carotid canal.

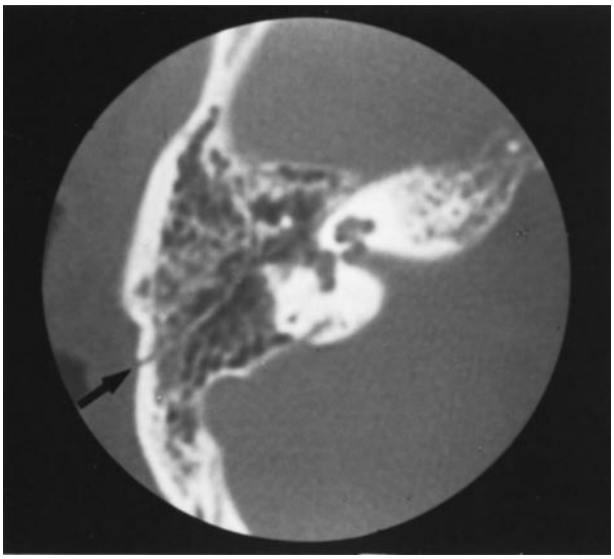


FIGURE 7–23. Longitudinal fracture of the right temporal bone, axial computed tomographic section. The longitudinal fracture extends from the cortex of the posterior portion of the mastoid to the posterosuperior canal wall (*arrow*). The fracture reaches the attic and disrupts the ossicular chain. The body of the incus is displaced laterally and posteriorly.

sections. The tegmen is usually involved, and cerebrospinal fluid otorrhea or rhinorrhea occurs whenever the dura is torn. Anterior extension of the fracture may reach the eustachian tube, which becomes obstructed. Conductive hearing loss is usually secondary to disruption of the ossicular chain. The body of the incus is usually rotated and displaced superiorly, posteriorly, and laterally (Figure 7–23). Interruption at the incudostapedial connection results from a fracture of the lenticular process of the incus or of the stapes superstructure and from a dislocation of the incudostapedial joint (Figure 7–24).

Meningoceles and Meningoencephaloceles A soft tissue mass contiguous to a defect in the tegmen of the mastoid suggests the possibility of a meningocele or meningoencephalocele. If the brain and meninges herniate into the small space of the antrum or epitympanum, the constant pulsation of the cerebrospinal fluid is transmitted through the walls of the meningocele to cause a gradual resorption of the surrounding bony walls. Computed tomography demonstrates the bony defect in the tegmen and a soft tissue mass protruding through the defect into the mastoid or middle ear cavity (Figure 7–25, A). This soft tissue lesion cannot be reliably differentiated from a recurrent cholesteatoma since the absorption coefficients of both lesions are similar. A more definite diagnosis of meningoencephalocele is reached with CT by injecting, by means of lumbar puncture, a small amount of metrizamide that will diffuse in the subarachnoid space and produce an enhancement of the herniated mass.[4] The differentiation between a meningocele and a cholesteatoma can best be made by MRI. On MRI, cholesteatoma appears as a lesion of medium signal in the T_1 and high signal intensity in T_2 images. A meningocele will have the same characteristics as cerebrospinal fluid: low signal intensity in T_1 and high intensity in the T_2 images. On MRI,

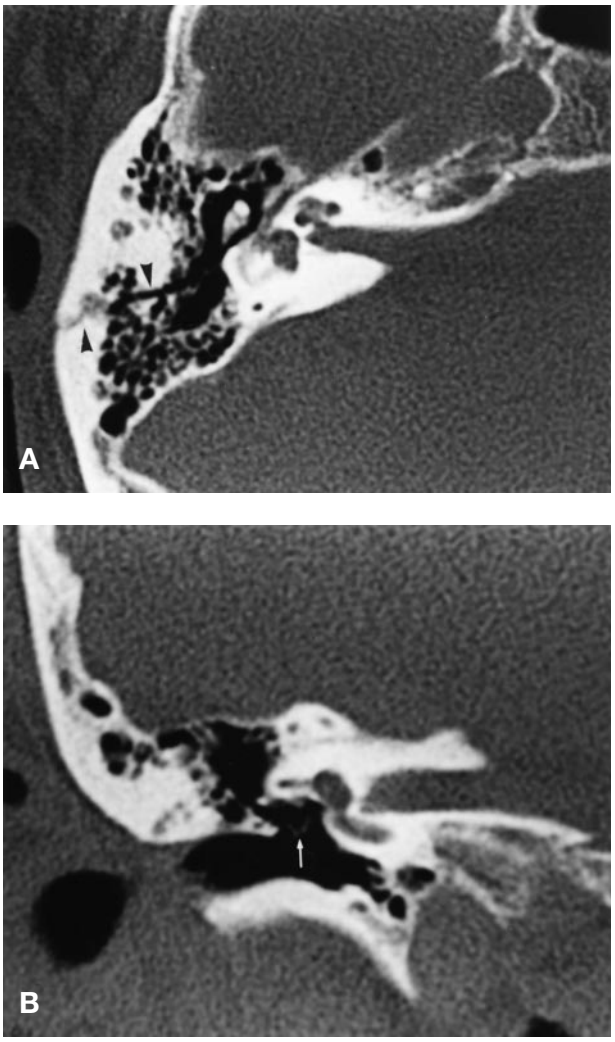


FIGURE 7–24. Longitudinal fracture of the right temporal bone. *A*, Axial; *B*, coronal computed tomographic sections. A fracture extends from the mastoid cortex to the region of the aditus (*black arrows*). The head of the malleus and the body of the incus are in normal relationship, but the incudostapedial joint is disrupted (*white arrows*).

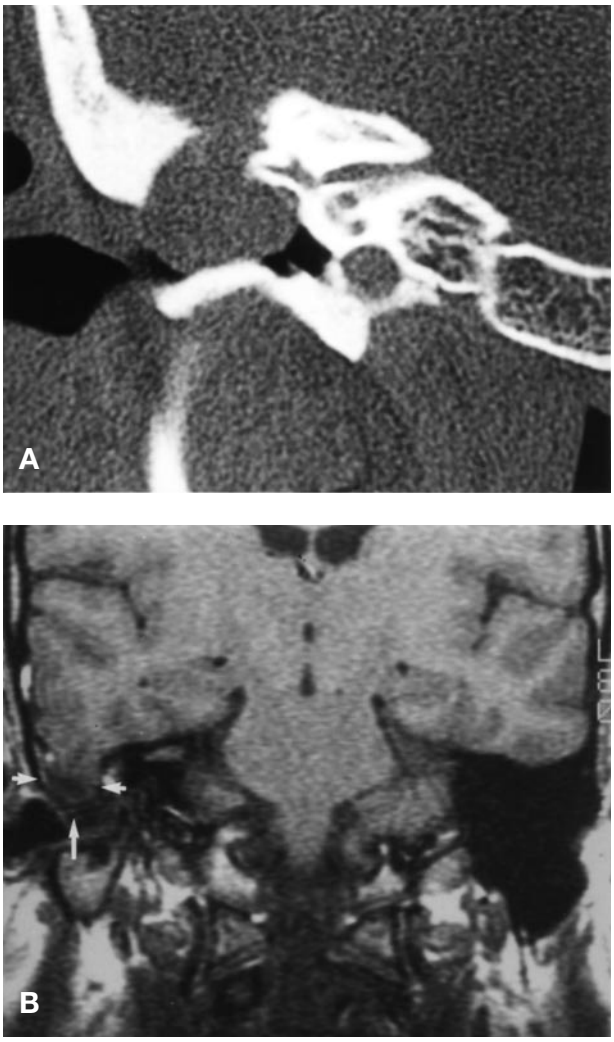


FIGURE 7–25. Meningoencephalocele, after modified radical mastoidectomy. *A*, Coronal computed tomographic section showing a large defect in the tegmen of the cavity with a soft tissue mass underneath it. *B*, T_1 coronal magnetic resonance image clearly identifying the soft tissue mass (*arrow*) as a brain herniation.

encephaloceles have the same characteristics of the adjacent brain, which are quite different from a cholesteatoma (Figure 7–25, B).

Acute Otitis Media The CT sections demonstrate a nonspecific and diffuse clouding of the middle ear cavity. The tympanic membrane is often swollen and bulges externally.

Chronic Otitis Media There are two types of chronic otitis media. In chronic suppurative otitis media, a partial, nonhomogeneous clouding of the middle ear cavity is caused by granulation tissue, polyps, and fluid or pus. Because the tympanic membrane is perforated, some aeration of the middle ear space is present. Erosion of the long process of the incus is a common finding, whereas erosion of the body of the incus and the head of the malleus is rare unless a cholesteatoma is present. In chronic adhesive otitis media, the middle ear space is contracted because of retraction of the tympanic membrane onto the promontory (Figure 7–26). The handle of

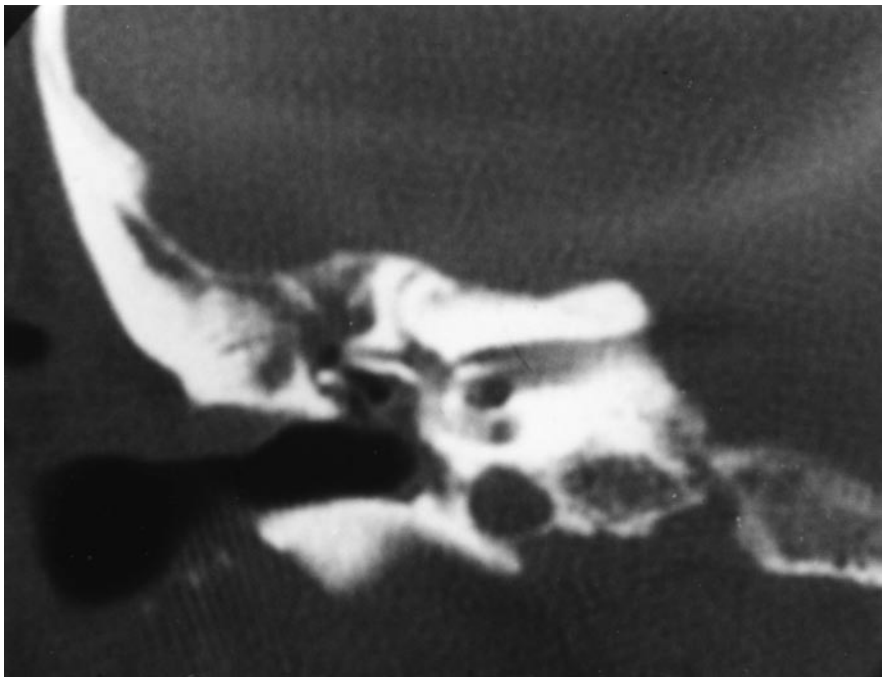


FIGURE 7–26. Right chronic otitis media, coronal computed tomographic section. The tympanic membrane is thickened and retracted. The mesotympanum is markedly contracted, and the mastoid air cells appear cloudy. The lateral wall of the attic and incudostapedial junction are intact.

the malleus is foreshortened, and the long process of the incus is often thinned out or eroded. In these cases, the retracted tympanic membrane may be attached to the head of the stapes with formation of a natural myringostapediopexy. Tympanosclerotic deposits are recognizable by CT whenever they are sufficiently large and calcified. The deposits appear as punctate or linear densities within the tympanic membrane or the mucosa covering the promontory. Large deposits of tympanosclerosis in the attic may surround and fix the ossicles.

Active and vascular granulation tissue undergoes a nonhomogeneous enhancement in the T_1 MRIs obtained after injection of contrast. This approach may occasionally be useful to differentiate granulation tissue from cholesteatoma, cholesterol granuloma, or other masses occurring in the middle ear.

Cholesteatomas Congenital cholesteatomas appear as well-defined soft tissue masses often producing an outward bulge of the intact tympanic membrane (Figure 7–27). If there is accompanying serous otitis media, the fluid may obscure the mass as the entire tympanic cavity becomes cloudy. If the lesion extends into the attic, the medial aspect of the lateral wall of the attic is eroded from within.

Acquired cholesteatoma produces a soft tissue mass in the middle ear and typical erosion of the lateral attic wall, posterosuperior canal wall, and ossicles. If the middle ear cavity is aerated, the soft tissue mass is well outlined. When fluid or inflammatory tissue surrounds the cholesteatoma, the margin of the mass is obscured because the x-ray densities of cholesteatoma, inflammatory tissue, and fluid are similar. Magnetic resonance imaging with contrast may be useful in these cases to differentiate the enhancing granulation tissue from fluid and cholesteatoma. Different patterns of x-ray findings are observed. Cholesteatomas that arise from the pars flaccida of the tympanic membrane produce erosion of the anterior portion of the lateral wall of the attic and tympanic spine (Figure 7–28). The lesion extends into the attic lateral to the ossicles, which may become medially displaced. Cholesteatomas arising from the pars tensa, usually from the posterosuperior margin of the tympanic membrane, appear as a soft tissue mass in the middle ear, eroding the long process of the incus and extending into the attic medial to the ossicles that may be displaced laterally (Figure 7–29). The lateral wall of the attic is usually intact, but the posterosuperior canal wall is often eroded. Sometimes a cholesteatoma may involve both the pars flaccida and tensa, producing a mixed x-ray pattern. In advanced cholesteatoma of all types, extensive bony destruction occurs, and no distinct pattern remains. The ossicles in the attic,

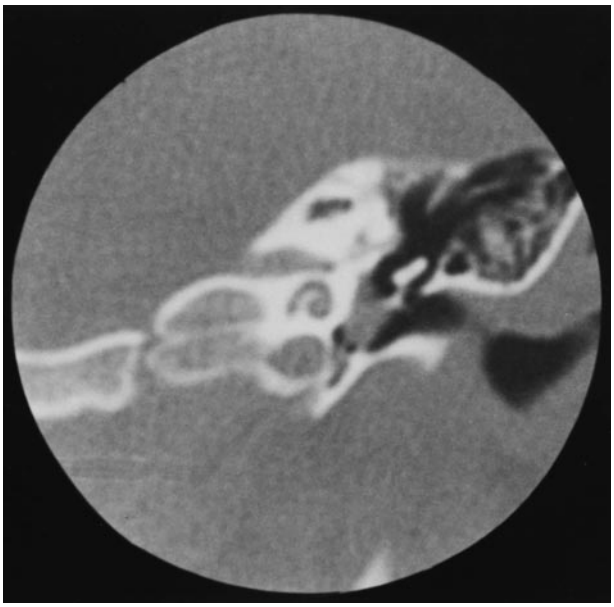


FIGURE 7–27. Left congenital cholesteatoma, coronal computed tomographic section. A well-defined soft tissue mass lies in the mesotympanum and extends from the intact tympanic membrane to the promontory. The remainder of the middle ear cavity and the mastoid are well aerated.

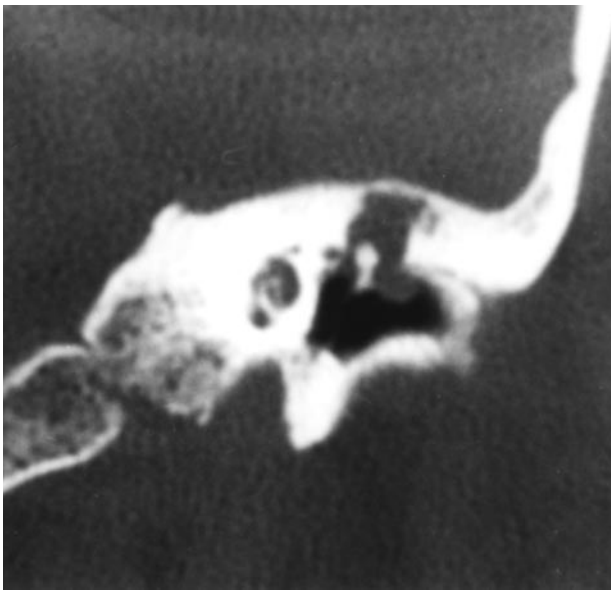


FIGURE 7–28. Left attic cholesteatoma, coronal computed tomographic section. The lateral wall of the attic is eroded by a soft tissue mass extending into the attic lateral to the ossicles. A polyp protrudes into the external auditory canal through a perforation of the pars flaccida of the tympanic membrane.

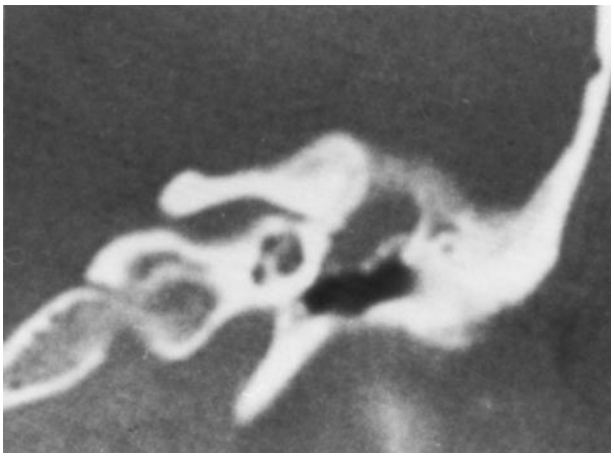


FIGURE 7–29. Left cholesteatoma, pars tensa perforation type, coronal computed tomographic section. The cholesteatoma extends into the attic medial to the ossicles and displaces the ossicles laterally.

particularly the body of the incus, are eroded, the aditus is widened, and the mastoid antrum becomes enlarged, cloudy, and smooth in outline because of erosion of the air cells lining the walls of the antrum. Further extension into the mastoid causes destruction of the trabeculae with formation of large cavities. Erosion of the tegmen may lead to meningeal and intracranial complications. Labyrinthine fistulae occur most commonly in the lateral portion of the horizontal semicircular canal (Figure 7–30). The CT sections show flattening of the normal convex

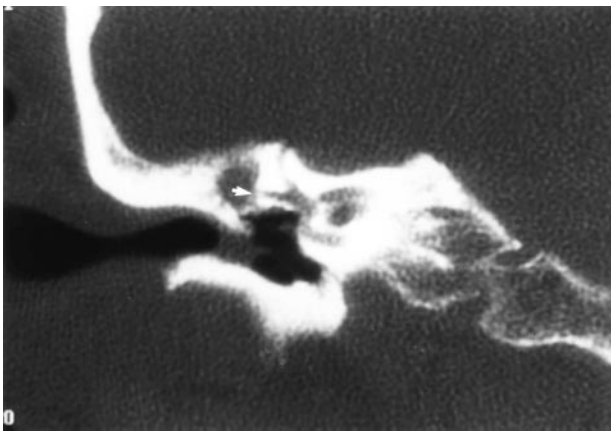


FIGURE 7–30. Right cholesteatoma with labyrinthine fistula. Coronal computed tomographic section: a soft tissue mass fills the attic, eroding the ossicles and the lateral aspect of the horizontal semicircular canal (*arrow*).

contour of the horizontal semicircular canal and erosion of the bony capsule covering the lumen of the canal. Further complications of cholesteatoma are extension into the petrous pyramid, which usually occurs in well-pneumatized bones, and erosion of the facial nerve canal, which may lead to facial paralysis.

Tumors Osteomas are frequent in the external auditory canal but also occur in the middle ear, where they may cause conductive hearing loss by impinging on the ossicular chain. Glomus tumors, also called chemodectomas or nonchromaffin paragangliomas, arise in the middle ear or jugular fossa from minute glomus bodies. Glomus tympanicum tumors arise from glomus bodies along Jacobson's nerve on the promontory. The CT sections reveal a soft tissue mass of variable size, usually in the lower portion of the middle ear cavity (Figure 7–31, A). As the lesion enlarges, it may cause a lateral bulge of the tympanic membrane, smooth erosion of the promontory, and involvement of the mastoid and hypotympanic air cells. If the lesion erodes into the jugular fossa, it becomes indistinguishable from a glomus jugulare tumor. Rapid sequential CT images of a preselected section showing the tumor mass, obtained after injection of a bolus of contrast material, generate a curve with a high quasiarterial peak rather than the high but delayed venous peak of a high jugular bulb[5] (Figure 7–31, B). Selective arteriography with subtraction is indicated to identify feeding vessels before embolization. Carcinomas extend into the middle ear from the external auditory canal. Adenocarcinoma is rare and produces a nonspecific soft tissue mass, often eroding the walls of the middle ear.

Otosclerosis Otosclerosis that involves the oval window causes fixation of the stapes and consequent conductive hearing loss. A 20-degree coronal oblique is far superior to the other projections for the study of the oval window. In active otosclerosis or otospongiosis, the margin of the oval window becomes decalcified so that the window seems larger than normal. In mature otosclerosis, the oval window becomes narrowed or closed (Figure 7–32), and in severe cases, the entire oval window niche is obliterated by calcified otosclerotic foci. The x-ray assessment is useful before surgery to confirm the clinical diagnosis and in some bilateral cases for selection of

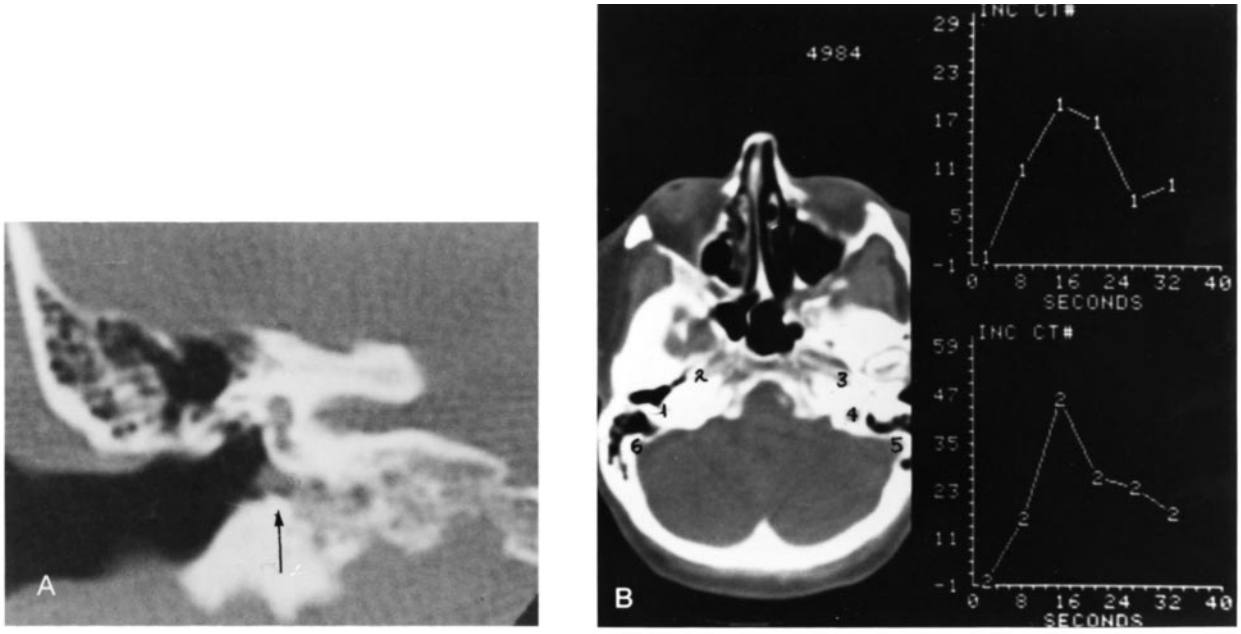


FIGURE 7–31. Right glomus tympanicum. *A*, Coronal computed tomographic (CT) section showing a small soft tissue mass in the lower portion of the middle ear cavity (*arrow*). *B*, Dynamic CT study of the preselected section showing the mass. Graphic display of the circulation in the tumor (1) and internal carotid artery (2). Notice the high peak of the tumor at quasiarterial time.

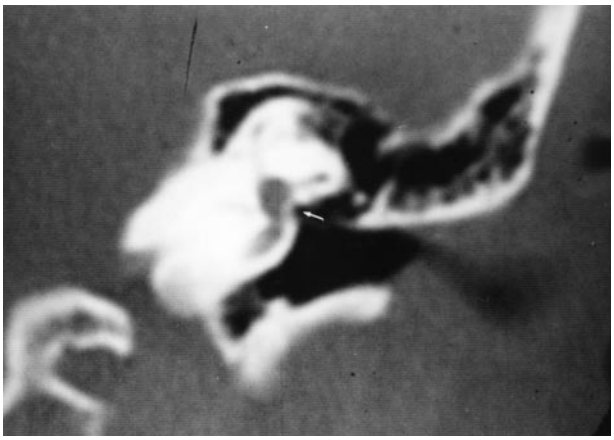


FIGURE 7–32. Stapedial otosclerosis, 20-degree coronal oblique section of the left ear. The footplate of this stapes appears thickened (*arrow*).

the ear to be operated on. More important is the study of the poststapedectomy ear in determining the cause for recurrent or persistent hearing loss and of immediate or delayed vertigo. Computed tomography can demonstrate protrusion of the prosthesis into the vestibule, reobliteration of the oval window with fixation of the strut, dislocation of the medial end of the prosthesis from the oval window, and separation of the lateral end of the strut from the incus or necrosis of the long process of the incus.

Inner Ear

Both CT and MRI are used for the study of the inner ear structures. Computed tomography is best for the study of the labyrinthine capsule; MRI is best for the assessment of the membranous labyrinth.

Congenital Anomalies Most cases of congenital sensorineural hearing loss have abnormalities limited to the membranous labyrinth and are therefore not demonstrable by CT.[6] Defects in the otic capsule are visible by CT. Anomalies may involve a single structure or the entire capsule. The most severe anomaly is the Michel type, which is characterized by hypoplasia of the petrous pyramid and almost complete lack of development of the inner ear structures (Figure 7–33). A less severe but more common malformation is the Mondini defect, which is characterized by hypoplasia or absence of the bony partitions between the cochlear coils, a short but wide vestibular aqueduct, and dilatation of the vestibule and ampullated ends of the semicircular canals (Figure 7–34). Enlargement of the vestibular aqueduct, hypoplasia of the cochlea and internal auditory canal, and deformity of the vestibule and semicircular canals may occur as isolated anomalies.[7] Dilatation of the cochlear aqueduct may be responsible for cerebrospinal fluid gush, which sometimes occurs

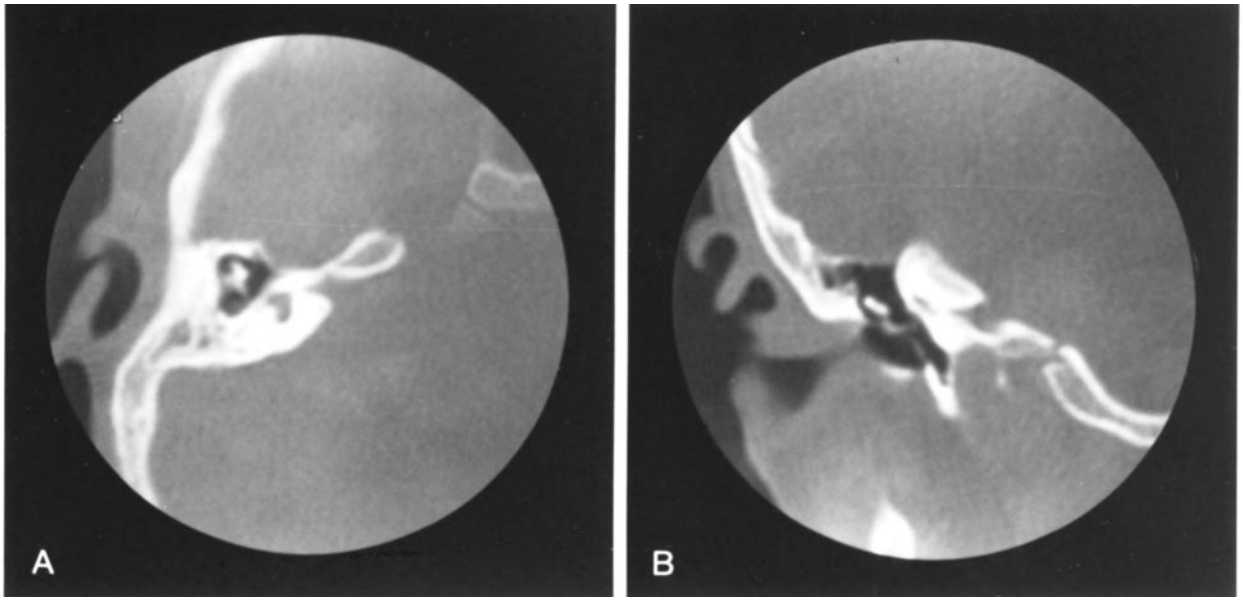


FIGURE 7–33. Michel malformation of the right inner ear structures. Axial (*A*) and coronal (*B*) computed tomographic sections. The middle ear cavity is normal, but the petrous pyramid is hypoplastic. The cochlea is absent or markedly hypoplastic. Notice the cavity in the region of the vestibule with rudimentary semicircular canals.

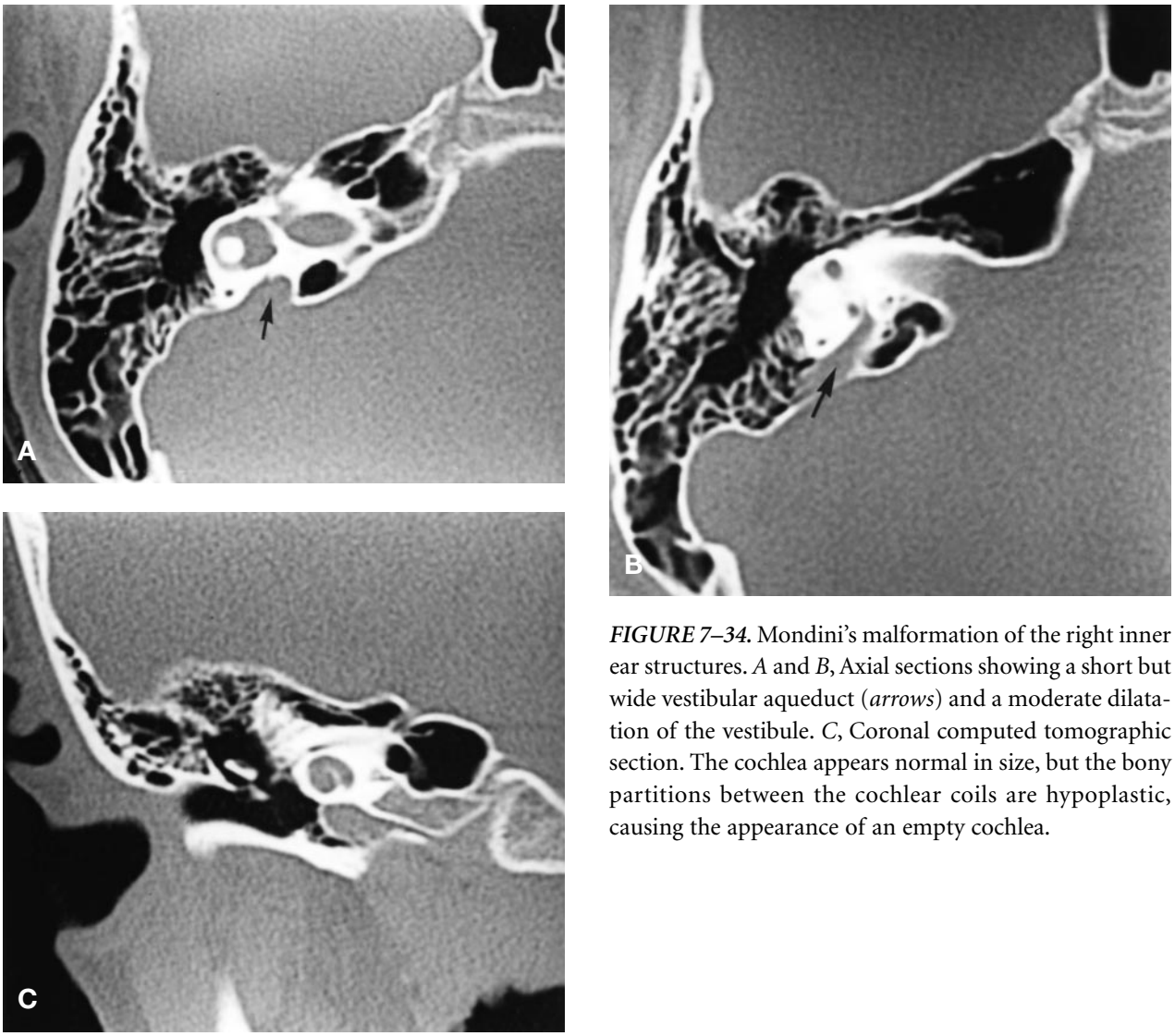


FIGURE 7–34. Mondini's malformation of the right inner ear structures. *A* and *B*, Axial sections showing a short but wide vestibular aqueduct (*arrows*) and a moderate dilatation of the vestibule. *C*, Coronal computed tomographic section. The cochlea appears normal in size, but the bony partitions between the cochlear coils are hypoplastic, causing the appearance of an empty cochlea.

during a stapedectomy for congenital footplate fixation or otosclerosis.

Trauma Bleeding within the lumen of the inner ear structures may occur after trauma. If bleeding occurs by concussion without actual fracture, MRI may be indicated to confirm the diagnosis. The study should be performed at 36 hours after the injury to allow the transformation of deoxyhemoglobin into methemoglobin, which has a bright signal in both T_1 and T_2 images (Figure 7–35).

The inner ear structures are seldom crossed by longitudinal fractures of the temporal bone but are usually involved in transverse fractures. These fractures typically cross the petrous pyramid at a right angle to the longitudinal axis of the pyramid and extend from the dome of the jugular fossa to the superior petrous ridge. Laterally placed fractures involve the promontory, vestibule, horizontal and posterior semicircular canals, and occasionally the tympanic segment of the facial nerve (Figure 7–36). Medially situated fractures involve the vestibule, cochlea, fundus of the internal auditory canal, and common crus.

Labyrinthitis Enhancement within the lumen of the bony labyrinth is often observed in MRI obtained after injection of contrast material in patients with acute bacterial or viral labyrinthitis and sudden deafness[8–10] (Figure 7–37).

Chronic labyrinthitis varies from a localized reaction caused by a fistula of the bony labyrinth to

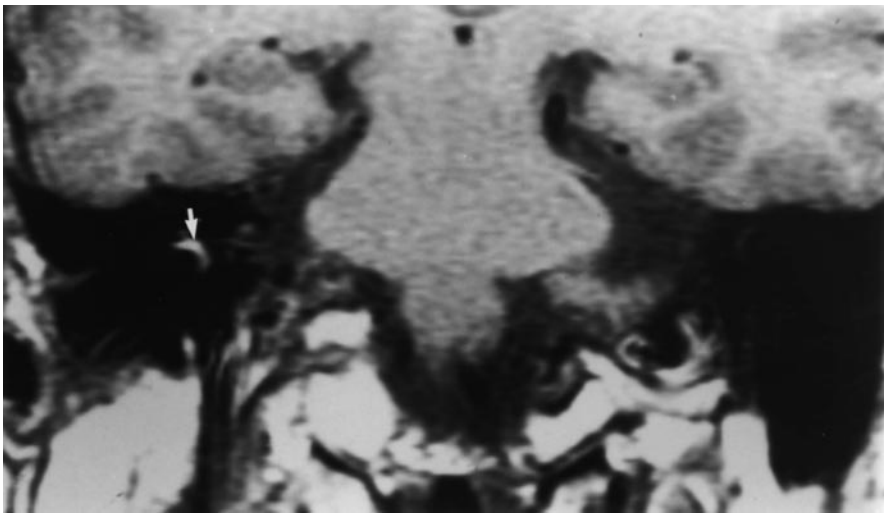


FIGURE 7–35. Labyrinthine concussion with bleeding. This T_1 coronal MR section shows an area of high signal intensity within the right vestibule (*arrow*) and adjacent portion of the semicircular canals.

a diffuse process. The lumen of the inner ear is partially or totally filled with granulation and fibrous tissues. Osteitis of the bony labyrinth occurs, which may lead to a partial or complete bony obliteration of its lumen (Figure 7–38). Whereas bony obliteration of the inner ear is readily identified by CT, fibrous obliteration is recognizable only by MRI. In the T_2 images, the high signal seen within the normal inner ear structures is absent, therefore making involved structures no longer recognizable.

Labyrinthine Schwannomas In the past, small schwannomas have been found within the vestibule and cochlea during postmortem dissection of the temporal bone. These lesions are usually not recognizable by CT but are well demonstrated as small, enhanced masses in MRI examinations performed after injection of contrast material (Figure 7–39).

Endolymphatic Sac Tumors Endolymphatic sac tumors are locally aggressive papillary adenomatous

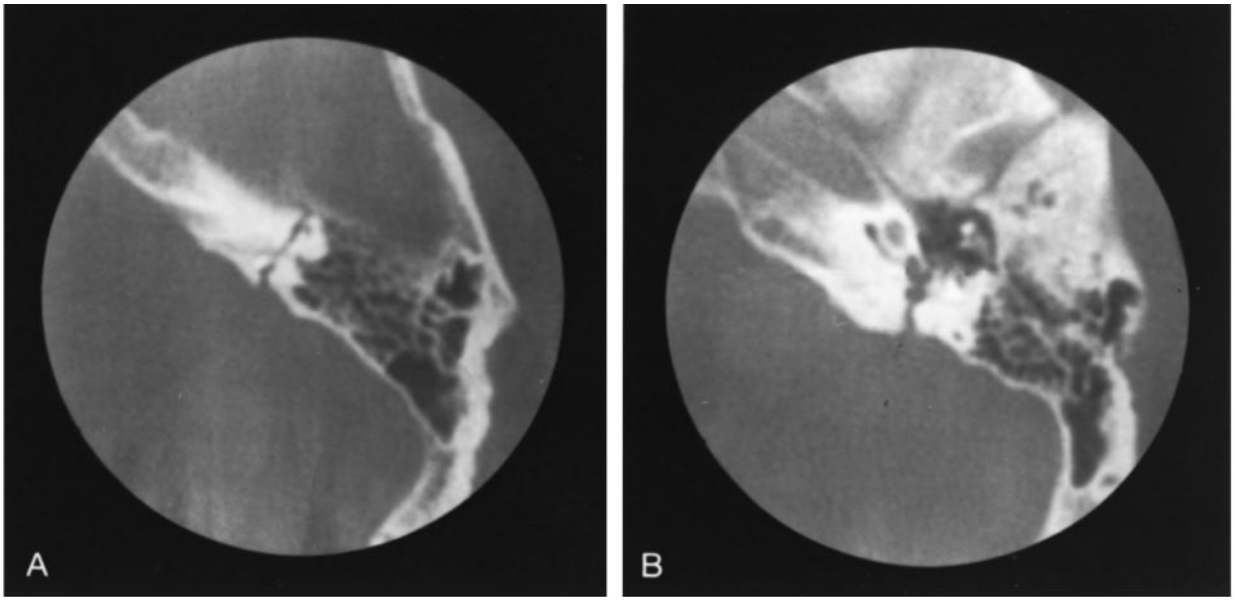


FIGURE 7–36. Transverse fracture of the left petrous pyramid. *A* and *B*, Axial computed tomographic sections. The fracture extends from the superior part of the petrous ridge to the undersurface of the temporal bone crossing the superior semicircular canal, vestibule, and promontory of the cochlea.

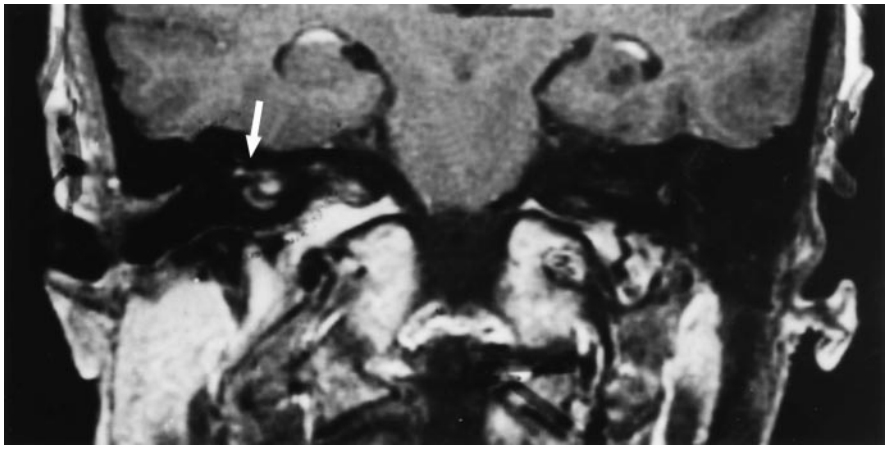


FIGURE 7–37. Acute labyrinthitis, presumably viral. This T_1 coronal MR image obtained after injection of gadolinium diethylenetriamine pentaacetic acid reveals enhancement within the lumen of the right cochlea (*arrow*).

tumors.[11] They are often associated with von Hippel-Lindau disease, a genetic multisystem neoplastic disorder. An early finding in endolymphatic sac tumors is a sensorineural hearing loss. At first, endolymphatic sac tumors involve the adjacent dura and endolymphatic duct. From there the lesion extends to the vestibule, semicircular canals, mastoid, and middle ear cavity, where it appears through an intact tympanic membrane as a bluish mass, often confused with a glomus tumor. Continuous growth leads to complete replacement of the mastoid and petrous pyramid by tumor. Axial CT images initially show a localized area of erosion of the posterior aspect of the petrous pyramid in the region of the endolymphatic sac (Figure 7–40, A). As the lesion enlarges, destruction of the petrous pyramid is observed with involvement of the inner ear structures. In the T_1 MRI, the tumor has a heterogeneous appearance with areas of high signal owing to cysts filled with blood or high proteinaceous fluid and multiple small areas of signal void owing to calcifications and blood vessels. Following administration of contrast, the solid portion of the mass undergoes a nonhomogeneous enhancement (Figure 7–40, B).

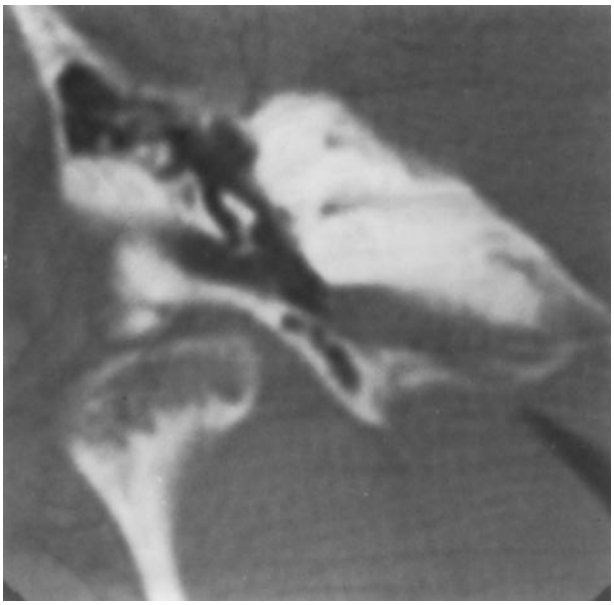


FIGURE 7–38. Obliterative labyrinthitis, coronal computed tomographic section of the right ear. Notice the complete bony obliteration of the cochlear lumen.

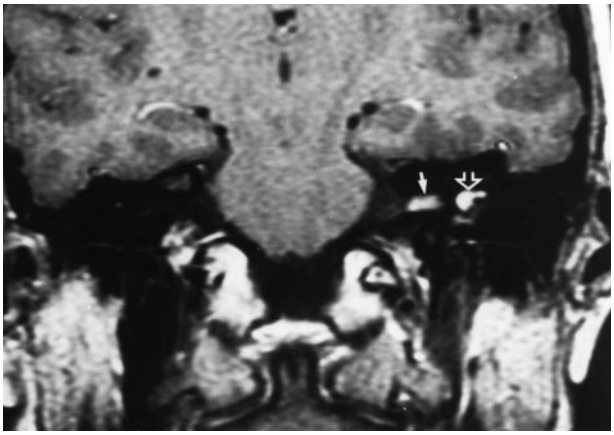


FIGURE 7–39. Left labyrinthine schwannoma. T_2 coronal magnetic resonance image after injection of contrast material. An enhancing soft tissue mass fills the vestibule and extends into the horizontal semicircular canal (*open arrow*). A second, seemingly not connected, mass is seen within the left internal auditory canal (*solid arrow*).

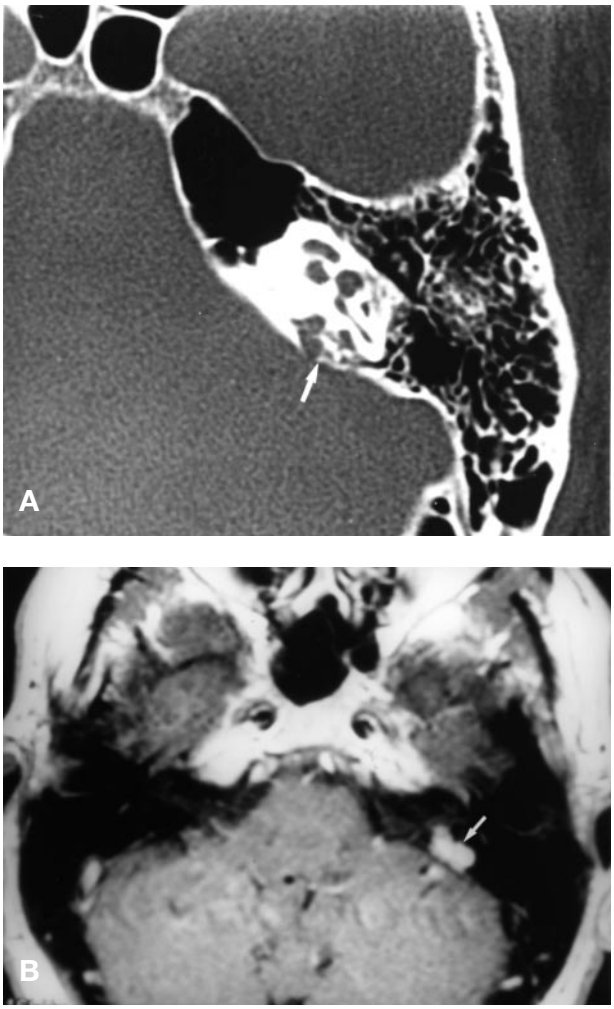


FIGURE 7–40. Endolymphatic sac tumor. *A*, Axial computed tomographic section of the left petrous pyramid. There is erosion of the posterior aspect of the petrous pyramid in the region of the endolymphatic sac with formation of an irregular cavity (*arrow*). *B*, T_1 axial magnetic resonance image. The tumor mass enhances after injection of contrast (*arrow*).

Otosclerosis Otosclerosis involving the cochlear capsule is often responsible for sensorineural hearing loss by an unknown mechanism. Cochlear otosclerosis is caused by progressive enlargement of the perifenestral foci or by single or multiple foci in other locations in the cochlear capsule. The CT findings vary with the stage of maturation of the process. In the active or spongiotic phase, small areas of demineralization are first observed in the normally sharp contour of the capsule. These foci may enlarge and become confluent, producing large areas of demineralization and finally complete dissolution of the capsule. A typical sign of active cochlear otosclerosis is the "double ring" caused by confluent spongiotic foci within the thickness of the capsule (Figure 7–41). In the mature or sclerotic stage, localized or diffuse areas of thickening of the capsule are present. A precise and quantitative assessment of the involvement of the cochlear capsule is accomplished by CT densitometric readings. The contour of the cochlear capsule is scanned with the smallest cursor (0.25 × 0.25 mm), and 31 or more densitometric readings are obtained. A profile of the density of the capsule is obtained by plotting densitometric values versus the points where the readings were made. The obtained curve is then compared with the densitometric profile of the normal capsule, which was previously determined[12] (Figure 7–42, A). Variations in density exceeding the standard deviation of 10 to

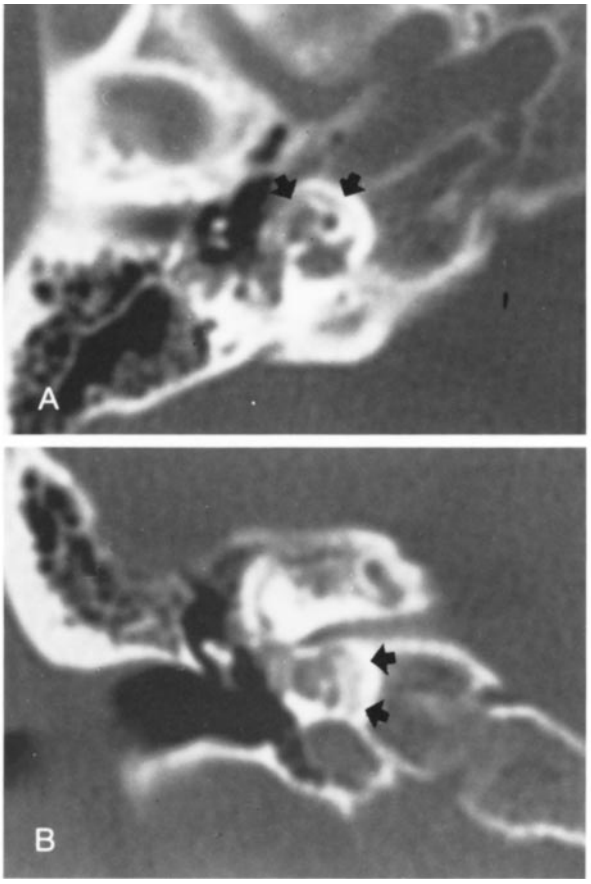


FIGURE 7–41. Cochlear otosclerosis. Axial (*A*) and coronal (*B*) computed tomographic sections. Multiple spongiotic foci are noted in the cochlear capsule with formation of a double ring (*arrows*).

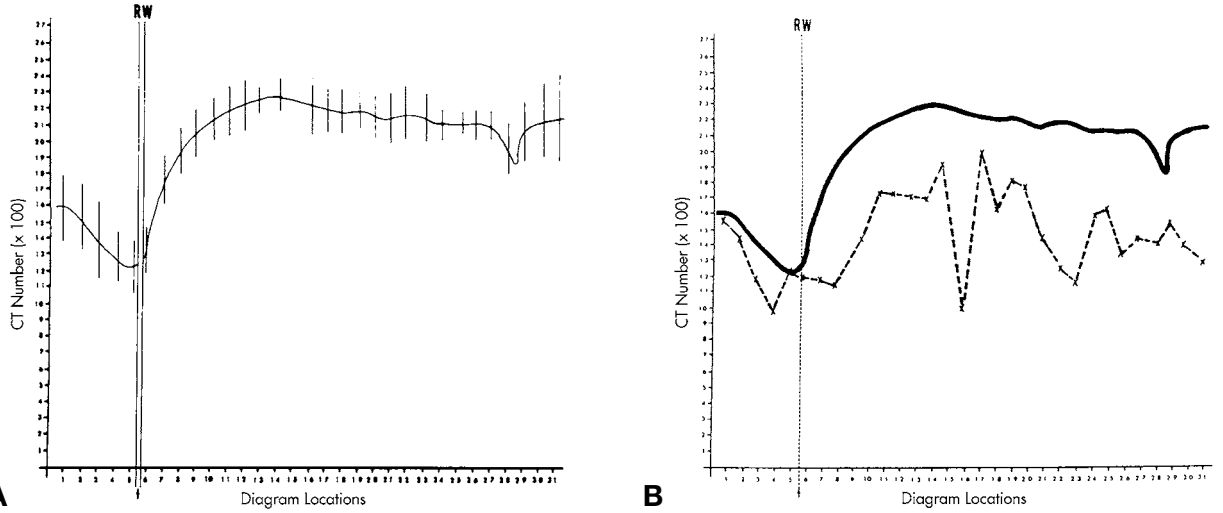


FIGURE 7–42. A, Densitometric profile of the normal cochlear capsule with standard deviations. *B*, Densitometric profile of the patient shown in Figure 7–41 (*interrupted line*). Comparison with the densitometric profile of the normal capsule (*solid line*) shows a diffuse demineralization of the cochlear capsule that is more severe in the region of the middle and apical coils. RW = round window.

15% are considered significant (Figure 7–42, B). With comparison of the original densitometric profile with follow-ups, it is possible to assess the evolution of the otosclerotic process and, when surgical or medical treatment was performed, to determine if any change has occurred in the maturation and extent of the disease.

Petrous Pyramid

Petrous Apex The petrous apex may be extensively pneumatized or made up of compact or diploic bone. In an MRI study, the signal intensity of the apex varies with the bony texture: high or bright in the T_1 images when diploic, low or dark when highly pneumatized or compact. Often the bony texture of the two petrous apices of the same person is different, resulting in one apex being brighter than the other (Figure 7–43).

Petrositis Extension of an acute suppurative mastoiditis and otitis media into a pneumatized petrous apex leads to petrositis. The air cells of the petrous pyramid become cloudy, and the cells' walls are first ill defined and later destroyed with formation of coalescent cavities. The infection may spread intracranially with consequent serious complications.

Glomus Jugulare Tumors These tumors arise from minute glomus bodies found in the jugular

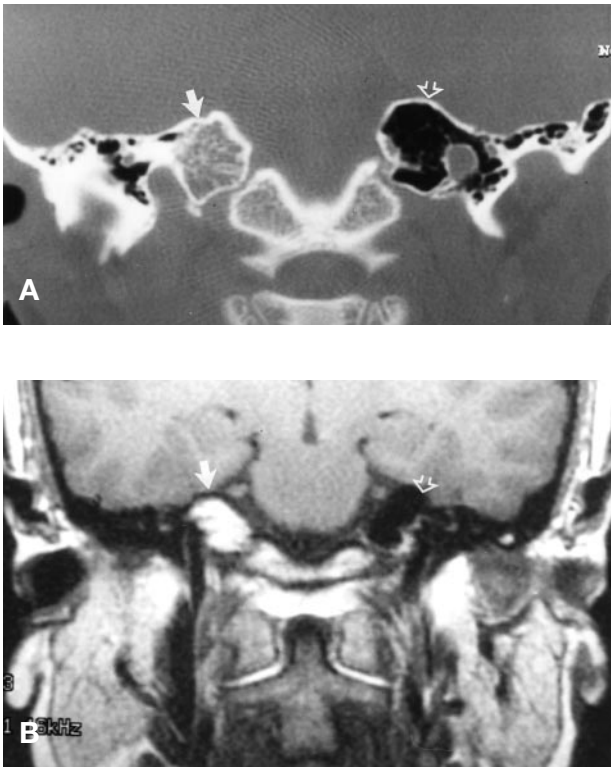


FIGURE 7–43. Petrous apex. *A*, Coronal computed tomographic section; *B*, coronal T_1-weighted magnetic resonance image (MRI). The right petrous apex (*solid arrow*) is diploic, whereas the left (*open arrow*) is pneumatized. The right diploic apex is bright in the MRI, whereas the left pneumatized apex appears as an area of signal void.

fossa. The typical CT findings include enlargement of the jugular fossa with erosion of the cortical outline, the septum dividing the jugular fossa from the outer opening of the carotid canal, and the hypotympanic floor with extension of the tumor into the middle ear cavity. As the lesion enlarges, the posteroinferior aspect of the entire petrous pyramid becomes eroded (Figure 7–44), as well as the adjacent aspect of the occipital bone, including the hypoglossal canal. Large tumors protrude extradurally in the posterior cranial fossa and inferiorly below the base of the skull along the jugular vein. These extensions are better demonstrated by MRI, in which the tumor appears in both T_1- and T_2-weighted images as a mass of medium signal intensity containing several small areas of signal void produced by blood vessels. After the injection of contrast material, the tumor undergoes a moderate enhancement (Figure 7–45). The signal intensity of the glomus tumor is differentiated easily from the surrounding intra- and extracranial structures. In addition, MRI allows determination of displacement, encroachment, narrowing, or obstruction of the jugular vein and internal carotid artery because these large vessels are well visualized, with no need for invasive vascular procedures. Jugular venography and carotid arteriography are seldom needed, and carotid arteriography is limited to preembolization.

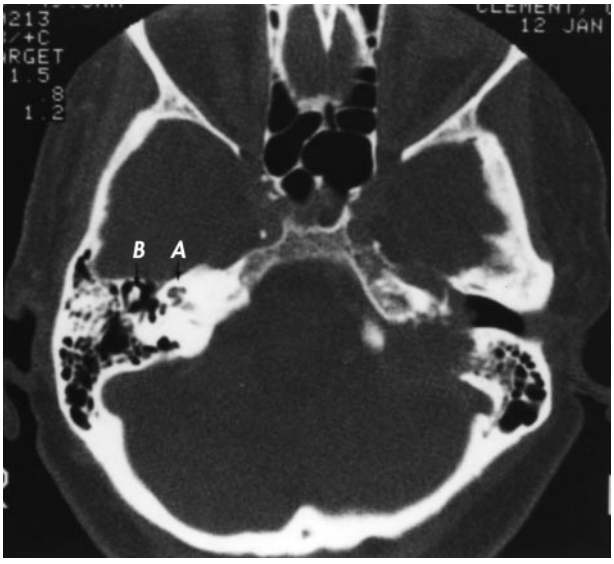


FIGURE 7–44. Left glomus jugulare tumor, axial computed tomographic section. The posterior aspect of the left temporal bone in the region of the jugular fossa is eroded. A soft tissue mass extends into the middle ear cavity through a large defect in its posterior wall. A = cochlea; B = ossicles, on the right side.

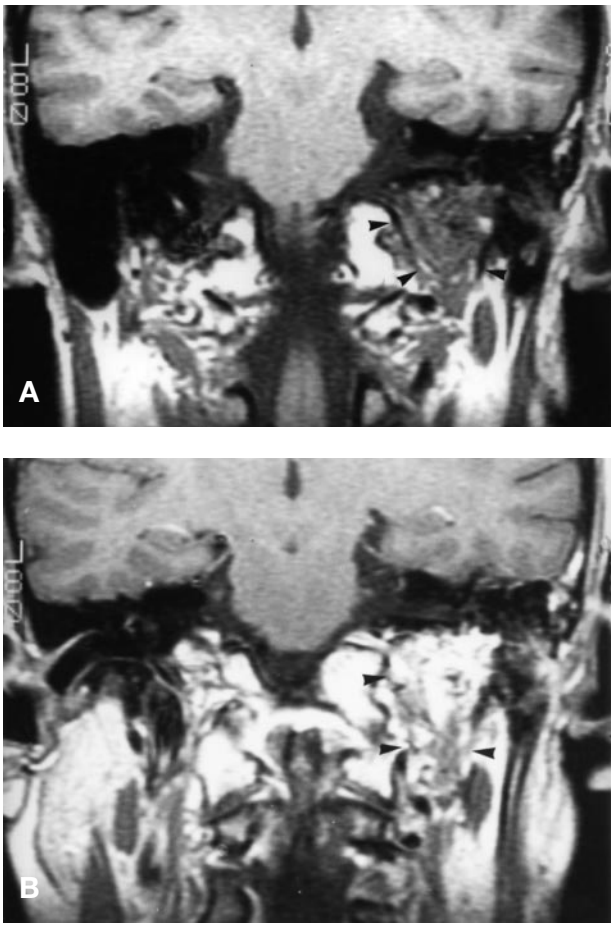


FIGURE 7–45. Glomus jugulare tumor. Coronal T_1-weighted magnetic resonance image (*A*) prior to and (*B*) after injection of contrast. A large and enhancing soft tissue mass (*arrowheads*) is present in the region of the left jugular fossa. The tumor extends into the lower portion of the middle ear cavity and erodes the undersurface of the petrous pyramid. Multiple areas of signal void, produced by blood vessels, are seen within the tumor mass, producing the classic "salt and pepper" appearance.

Congenital Cholesteatomas and Cholesterol Granulomas Congenital cholesteatomas may arise in the petrous pyramid. In the CT images, the involved area of the pyramid appears expanded by a cyst-like low-density lesion that often reaches and erodes the internal auditory canal and labyrinth. A CT study with infusion shows no enhancement of

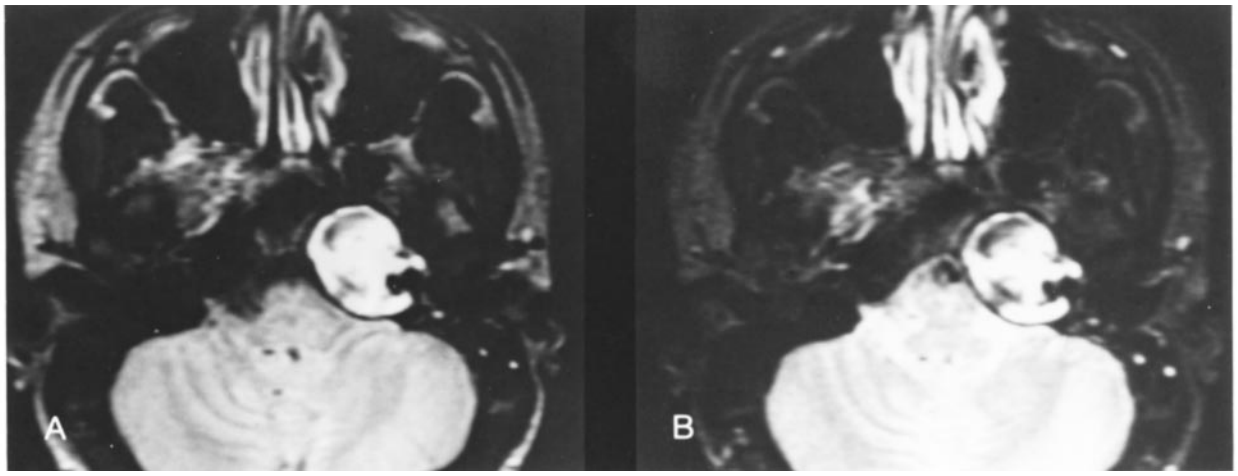


FIGURE 7–46. Cholesterol granuloma. Axial spin density (*A*) and axial T_2-weighted (*B*) magnetic resonance images. A mass of high signal intensity in both sequences is demonstrated in the left petrous apex. The areas of signal void within the mass are produced by deposits of hemosiderin.

the mass except for its capsule. With CT, a congenital cholesteatoma of the petrous apex is difficult to differentiate from a cholesterol granuloma cyst that occurs in extensively pneumatized petrous pyramids. The two lesions can be differentiated by MRI because congenital cholesteatomas produce a signal of medium intensity in the T_1 images and high intensity in T_2, whereas cholesterol granulomas have a similar high signal in both T_1 and T_2 sequences. Areas of signal void are often observed in cholesterol granulomas produced by deposits of hemosiderin (Figure 7–46).

Meningiomas These tumors arise from the meningeal covering of the temporal bone and from the meningeal extension into the internal auditory canal. The latter mimics, both clinically and in imaging, the appearance of an acoustic schwannoma. Meningiomas that arise from the petrous ridge are usually recognizable in the CT images as highly enhancing masses producing hyperostotic or lytic changes in the adjacent petrous pyramid. In meningiomas en plaque, only the bony changes are recognizable. Magnetic resonance images obtained after injection of paramagnetic contrast demonstrate a strong and usually homogeneous enhancement of the tumor. En plaque lesions are recognizable as areas of enhancing meningeal thickening (Figure 7–47).

Hemangiomas These uncommon tumors of the petrous pyramid usually arise in the region of the geniculate ganglion and extend into the internal auditory canal. Facial paralysis is often the first symptom. Computed tomography shows a poorly defined lytic lesion containing an almost pathognomonic bony spiculation[13] (Figure 7–48, A). With MRI, the tumor is of medium signal intensity in T_1 images, undergoes a strong enhancement after injection of contrast, and contains linear areas of signal void owing to bony spiculae (Figure 7–48, B).

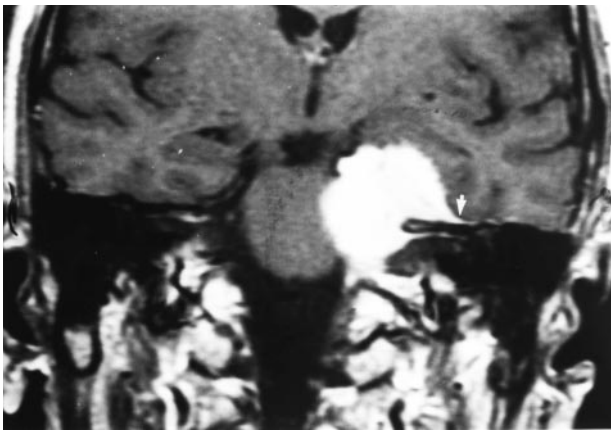
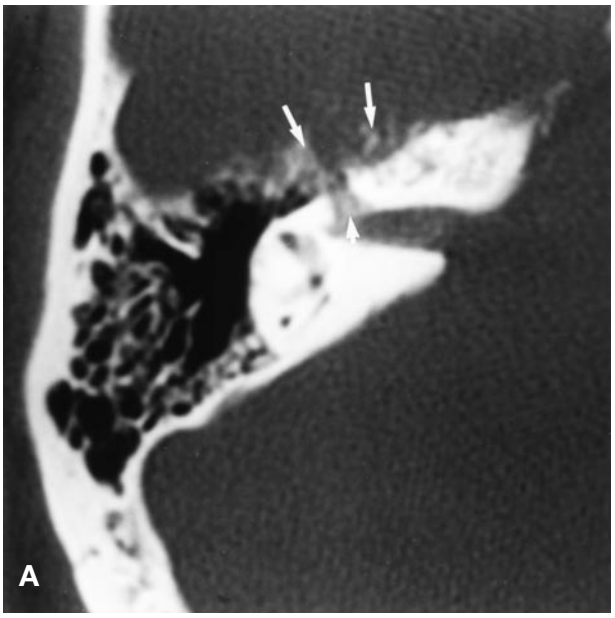


FIGURE 7–47. Meningioma coronal T_1-weighted magnetic resonance image section after injection of contrast material. A large enhancing mass extends above and below the left side of the tentorial notch. The tumor spares the internal auditory canal. Notice the lateral en plaque extension of the lesion (*arrow*).



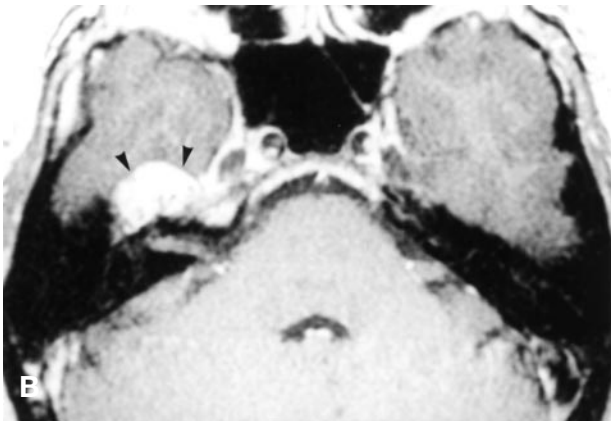


FIGURE 7–48. Hemangioma, right petrous pyramid. *A*, Axial computed tomographic section; *B*, axial T_1-weighted magnetic resonance image after contrast. The anterior aspect of the right petrous pyramid is eroded by an enhancing soft tissue mass (*arrowheads*) extending into the attic, labyrinthine segment of the facial nerve canal, and fundus of the internal auditory canal. Note the characteristic bony spiculation within the tumor mass (*arrows*).

Paget's Disease Paget's disease often affects the calvarium and the base of the skull, including the petrous pyramids. The disease usually spreads from the petrous apex laterally and produces a typical washed-out appearance of the involved pyramid caused by extensive demineralization (Figure 7–49). The internal auditory canal is usually involved first, followed by the otic capsule, which first becomes thinned out and is then completely erased. In the late stage of the disease, deposition of irregularly mineralized bone occurs and results in thickening of the petrosa, narrowing of the internal auditory canal, and fixation of the footplate of the stapes.

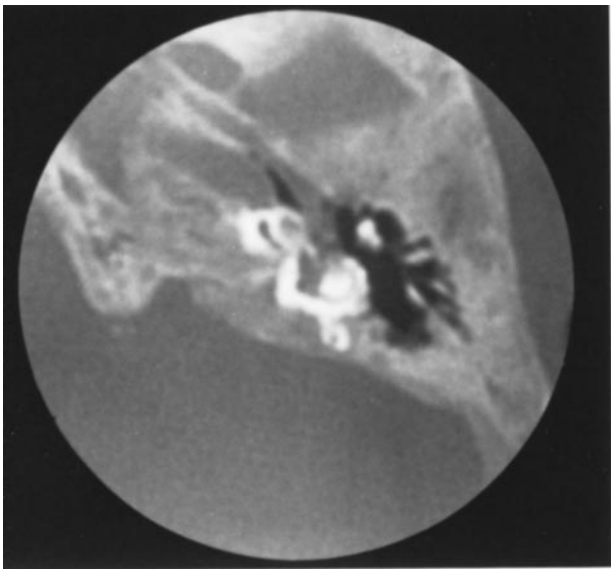


FIGURE 7–49. Paget's disease, axial section of the left temporal bone. Notice the severe demineralization of the petrous pyramid and mastoid with thinning of the otic capsule, particularly of the cochlea.

Facial Nerve

Computed tomography is the study of choice of the facial nerve canal. The examination should be performed in two or three planes to visualize the various segments of the canal. Magnetic resonance imaging visualizes the facial nerve itself, particularly when the nerve is thickened.

Congenital Anomalies Congenital anomalies involve the size and course of the facial canal. The canal may be partially or completely absent, hypoplastic, or unusually narrow. Minor variations of the course of the facial nerve are common and of no clinical significance. More severe anomalies should be identified to avoid serious damage to the nerve during surgery. The horizontal segment may be displaced inferiorly to cover the oval window or lie exposed on the promontory. In congenital atresia of the external auditory canal, the mastoid segment of the facial canal is rotated laterally. The rotation varies from minimal obliquity to a true horizontal course.

Trauma Traumatic lesions of the intratemporal portion of the facial nerve occur in approximately 20% of longitudinal fractures of the temporal bone and 50% of transverse fractures. The most common site of involvement in longitudinal fractures is the anterior genu (Figures 7–50 and 7–51) and in transverse fractures the labyrinthine or intracanalicular segments. The facial nerve may be transected by the fracture, compressed, sheared by a depressed fragment of the canal wall, or simply contused by the violent shock.

Facial Neuritis Moderate bilateral enhancement of the normal facial nerve, particularly in the region of its anterior genu, is often observed in MRI obtained after injection of contrast material.

Asymmetric enhancement of the facial nerve more prominent on the paralyzed side is common in patients with Bell's palsy and Ramsay Hunt syndrome. In Bell's palsy, the involvement is segmental and usually confined to the anterior genu and adjacent labyrinthine and tympanic segments (Figure 7–52). In Ramsay Hunt syndrome, the involvement by the herpes zoster virus is more continuous and often extends to the nerve within the internal auditory canal.

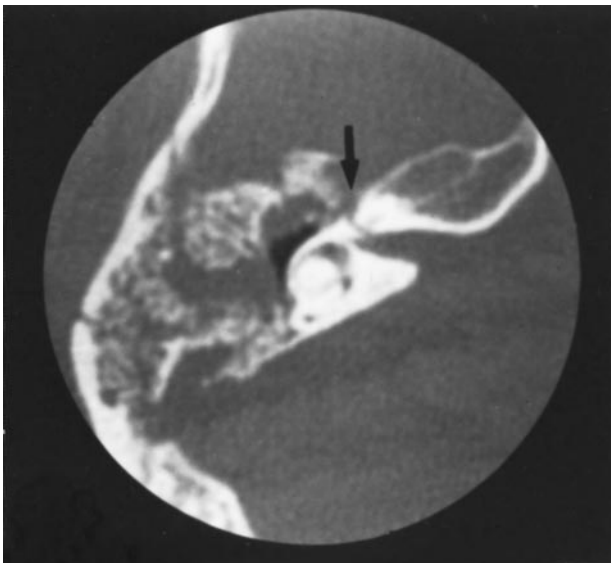


FIGURE 7–50. Longitudinal fracture of the right temporal bone, axial computed tomographic section. The fracture passes from the mastoid through the attic into the petrous pyramid anterior to the labyrinth. The fracture reaches and involves the facial canal at its anterior genu (*arrow*).

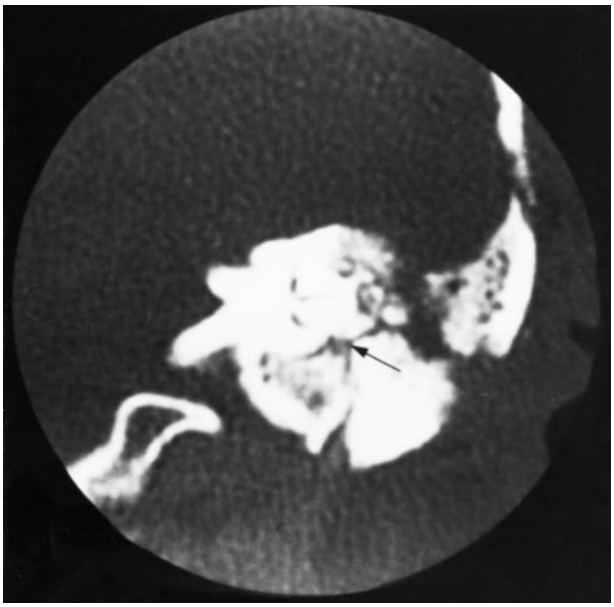


FIGURE 7–51. Comminuted fracture of the left mastoid, coronal computed tomographic section. Multiple fractures are noticed in the mastoid. One of the fractures passes underneath the posterior semicircular canal and transects the vertical segment of the facial canal (*arrow*).

In viral neuritis, the nerve is usually not thickened. In sarcoidosis, the involvement is similar to Bell's palsy, but the nerve is moderately thickened.

Tumors Primary tumors of the facial nerve are rare. Facial neurinomas cause thickening of the nerve and expansion of the canal (Figure 7–53, A). Further enlargement of the lesion results in erosion of the bony canal, extension into the middle ear space, and involvement of the mastoid and petrous pyramid. Magnetic resonance imaging is useful to differentiate large facial nerve neurinomas from other lesions. The mass has a nonspecific low signal in T_1-weighted images and a high signal in T_2-weighted images but changes to bright in T_1 images obtained after injection of gadolinium DTPA (Figure 7–53, B).

The facial canal, particularly its vertical segment, is often involved by carcinoma arising in the external auditory canal and malignancies arising in the parotid gland and extending into the temporal bone.

Internal Auditory Canal

Acoustic Schwannomas Acoustic schwannomas account for 90% of the tumors of the cerebellopon-

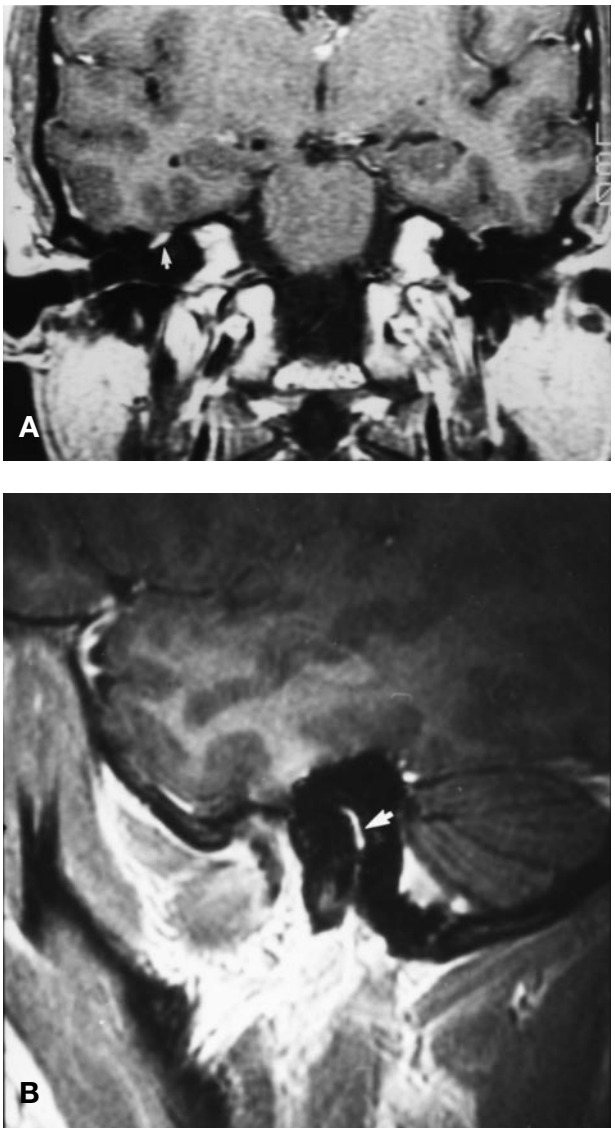


FIGURE 7–52. Facial neuritis. Coronal (*A*) and sagittal (*B*) T_1-weighted magnetic resonance images after injection of contrast material. There is enhancement of the right facial nerve in the regions of the anterior genu and proximal portion of the mastoid segment (*arrows*).

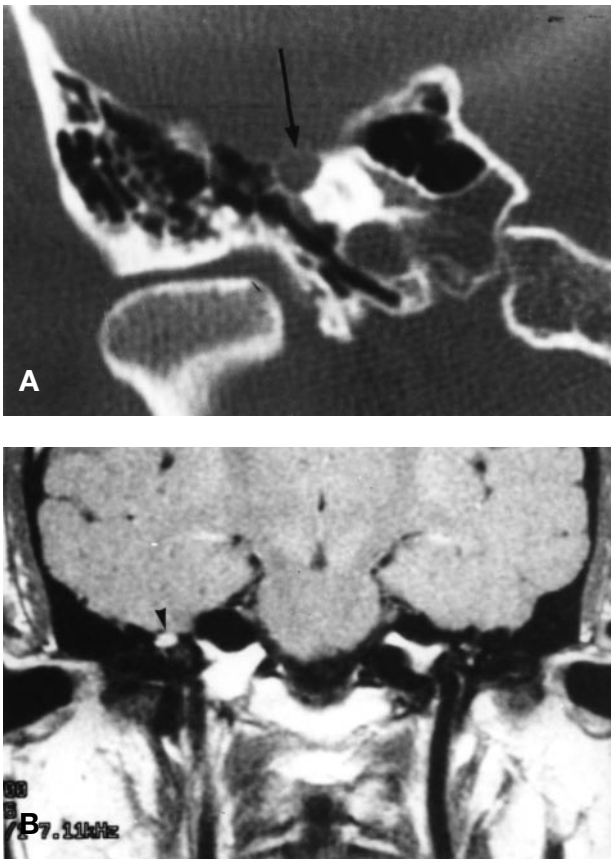


FIGURE 7–53. Facial neurinoma. *A*, Coronal computed tomographic section revealing expansion of the right facial canal at the anterior genu (*arrow*). *B*, The coronal T_1-weighted magnetic resonance image section obtained after injection of gadolinium diethylenetriamine pentaacetic acid shows the actual enhancing tumor mass (*arrowhead*).

tine angle. These tumors usually arise within the internal auditory canal, which becomes enlarged. Expansion of the internal auditory canal, shortening of its posterior wall, and erosion of the crista falciformis are well visualized in high-definition CT images targeted for bone. Both sides should always be examined to compare significant differences. Acoustic schwannomas are not visualized in plain CT scans because the tumor is isodense to the surrounding brain and is not surrounded by edema. After infusion of iodinated contrast material, the mass enhances and becomes visible (Figure 7–54). Intracanalicular lesions and cisternal masses smaller than 0.8 cm usually are not visualized by the infusion technique and, if MRI is not available, a CT pneumocisternogram is indicated. This examination requires a spinal puncture for the injection in the subarachnoid space of air, carbon dioxide, or oxygen. With proper positioning, the gas is moved into the cerebellopontine angle under investigation, and several thin sections are obtained. In normal cases, the gas fills the cerebellopontine cistern and the internal auditory canal outlining the seventh and eighth cranial nerves. If a small tumor is present, the gas outlines the localized swelling of the nerve; if the

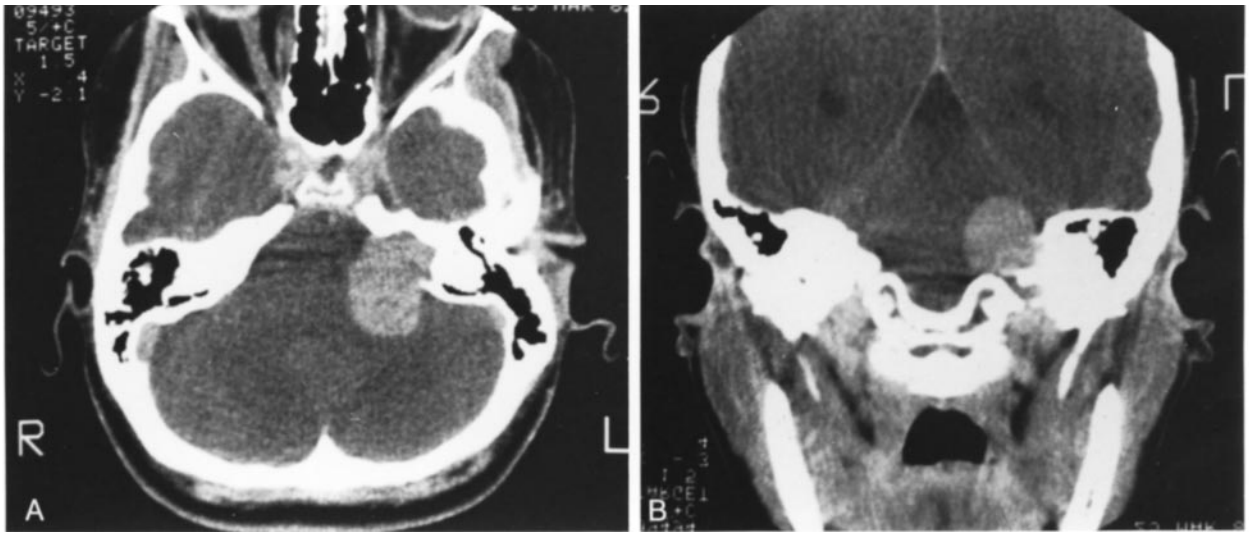


FIGURE 7–54. Left acoustic schwannoma, postinfusion study. *A*, Horizontal computed tomographic (CT) section. *B*, Coronal CT section. The left internal auditory canal appears grossly expanded and eroded. A large tumor mass fills the canal and the cerebellopontine cistern. Notice the displacement to the right of the brainstem and fourth ventricle.

tumor is large, it outlines the convex medial aspect of the mass, obstructing the canal (Figure 7–55).

Magnetic resonance imaging is the study of choice for diagnosis of acoustic schwannomas without exposing the patient to ionizing radiation and without the necessity for spinal puncture.[14,15]

Thin MRIs are obtained in the axial or coronal plane prior to and after injection of paramagnetic agents. In the precontrast image, the tumor is brighter than cerebrospinal fluid and isointense to gray matter. In the postcontrast study, the mass undergoes a marked enhancement as the contrast

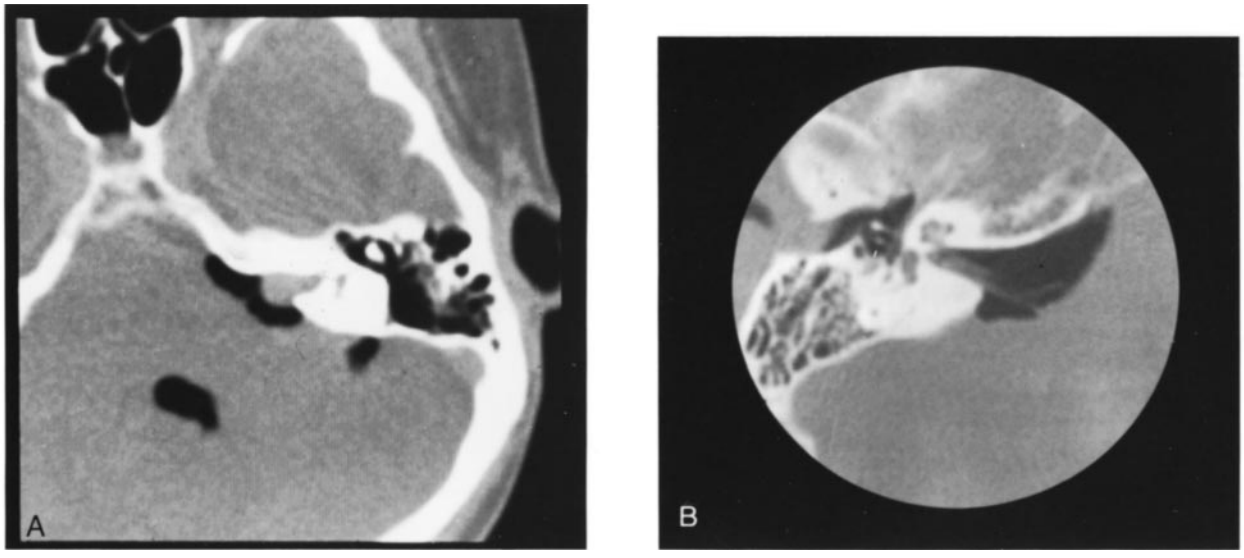


FIGURE 7–55. Left acoustic schwannoma. *A*, The computed tomographic pneumocisternogram demonstrates a tumor mass filling the left internal auditory canal and slightly protruding into the cerebellopontine cistern. Notice the uninvolved portion of the eighth cranial nerve extending from the medial aspect of the tumor to the brainstem. *B*, The right cerebellopontine cistern and internal auditory canal are well filled by air. The eighth cranial nerve courses from the brainstem through the cistern to the internal auditory canal.

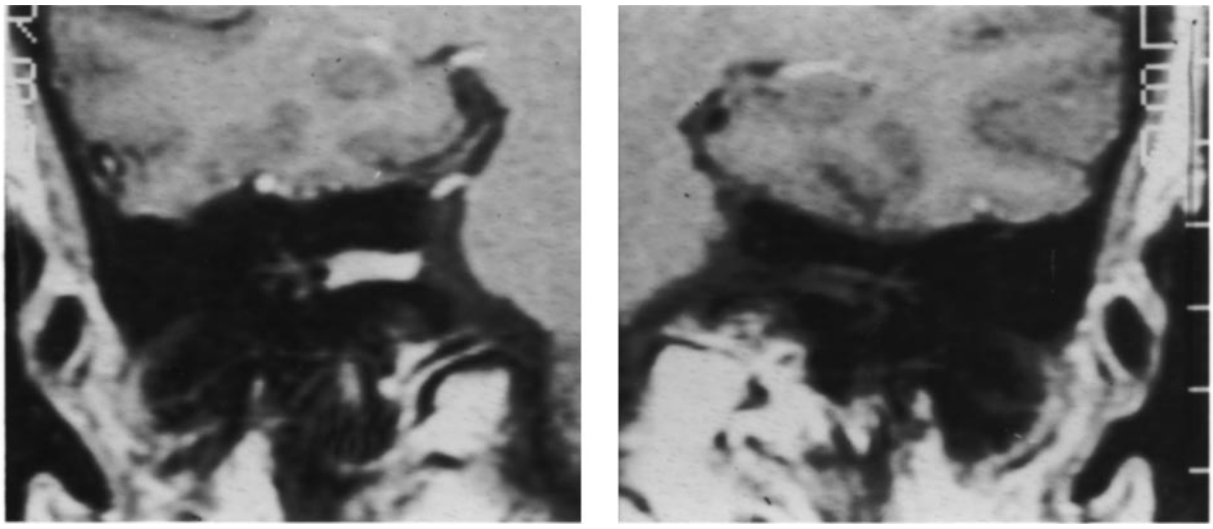


FIGURE 7–56. Right vestibular schwannoma. Right and left coronal T_1-weighted magnetic resonance images after contrast. An enhancing soft tissue mass fills the right internal auditory canal.

concentrates within the tumor with consequent shortening of the T_1 relaxation time (Figures 7–56 and 7–57). These images clearly show the size of the tumor within the internal auditory canal and cerebellopontine cistern, as well as possible extension into the modiolus, which carries a poor prognostic value unless the cochlea is sacrificed at surgery.[16] Extension of the tumor within the facial canal is usually indicative of a schwannoma arising from the facial nerve.

Meningiomas Meningiomas account for approximately 3 to 4% of the cerebellopontine angle tumors. The lesion has a signal intensity less than the brain in the T_1 sequences. In the T_2 images, the lesion has variable characteristics, either a pathognomonic further decrease in signal intensity or a nonspecific increase in brightness. After intravenous injection of gadolinium DTPA, in the T_1 images meningiomas show a marked increase in signal intensity similar to that seen in acoustic schwannomas. Unlike the latter tumor, meningiomas usually arise in the cerebellopontine cistern and spare the fundus or the entire internal auditory canal (see Figure 7–47).

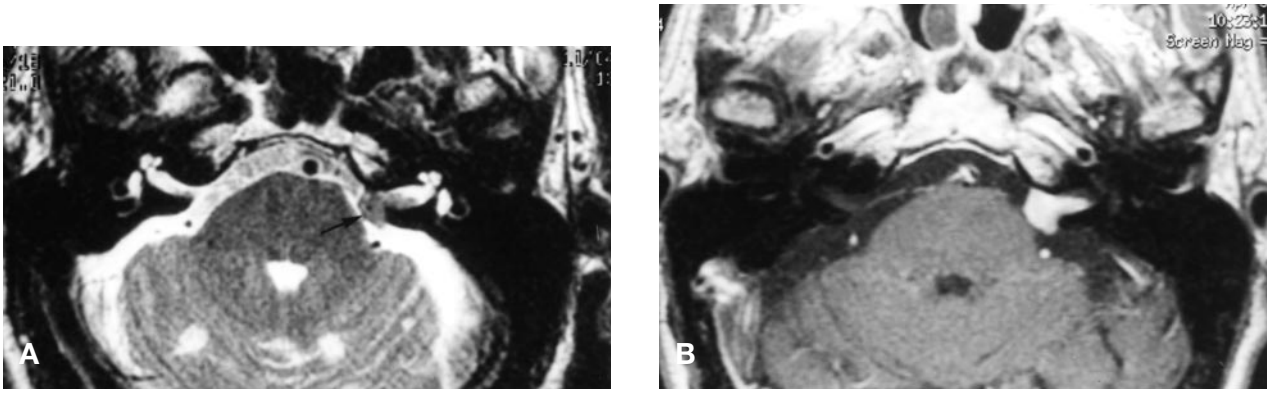


FIGURE 7–57. Left acoustic schwannoma. Axial magnetic resonance images. *A*, T_2 fast spin echo. *B*, T_1 postcontrast. An enhancing soft tissue mass partially fills the left internal auditory canal and protrudes into the cerebellopontine cistern. Note that the tumor is well seen in the fast spin echo image (*A*) as a filling defect within the bright signal of the cerebrospinal fluid (*arrow*).

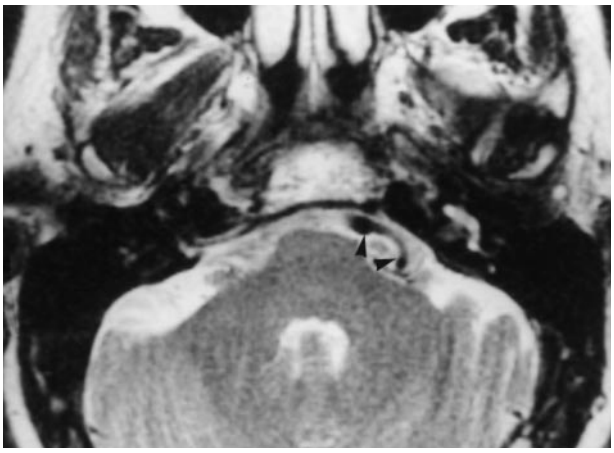


FIGURE 7–58. Vascular loop. Axial T_2 fast spin echo image. A prominent left vertebral artery loops within the left cerebellopontine cistern (*arrowheads*) crossing and presumably compressing the acoustic nerve.

Vascular Abnormalities The anterior inferior cerebellar artery often loops within the cerebellopontine cisterns, and in over 20% of cases, it actually enters the internal auditory canal. In addition, tortuous vertebral or basilar arteries often form prominent loops within the cerbellopontine cistern (Figure 7–58). Cross-compression of the acoustic nerve usually occurs at its takeoff from the brainstem owing to absence of slack in the nerve at this point or in a narrow internal auditory canal. In the past, vascular loops were identified by CT pneumocisternography. At present, we use T_2 FSE images in both axial and coronal planes.

Aneurysms Aneurysms in the internal auditory canal are extremely rare. An aneurysm appears on T_1- and T_2-weighted MRIs as a small mass of high signal presumably owing to a thrombosis or slow flow. Following the injection of contrast, the lesion becomes slightly larger. An aneurysm within the cerebellopontine cistern may compress the acoustic or facial nerve and mimic the symptomatology of a schwannoma. The MRIs obtained prior to and after contrast reveal a nonhomogeneous high signal produced by the blood clotting within the aneurysm (see Figure 7–9). If the lumen of the aneurysm is partially patent, the flowing blood will appear as an area of signal void.

Hemangiomas Small hemangiomas or AVMs limited to the lumen of the internal auditory canal are rare. They appear in preconstrast images as areas of high signal intensity owing to slow flow, which become larger following injection of contrast. The mass has nonhomogeneous intensity and may contain signal void areas owing to calcifications.

Malignant Neoplasms Leukemic and lymphomatous infiltrations have been described as occurring within the internal auditory canal and cerebellopontine cistern. The enhancing lesions are usually poorly defined and extend into the leptomeningeal space and meninges.

Metastatic deposits may occur within the internal auditory canal. In the author's experience, the most frequently encountered primaries are breast or lung carcinoma and melanoma. Whenever there is no bone involvement, the diagnosis of metastasis is difficult, although it should be considered whenever the patient has a history of neoplasia (Figure 7–59). Magnetic resonance imaging is the study of choice. The lesion has a signal of medium intensity in the T_1-weighted images, which becomes brighter in the T_2 images. Following the injection of contrast, the mass undergoes a homogeneous or nonhomogeneous enhancement.

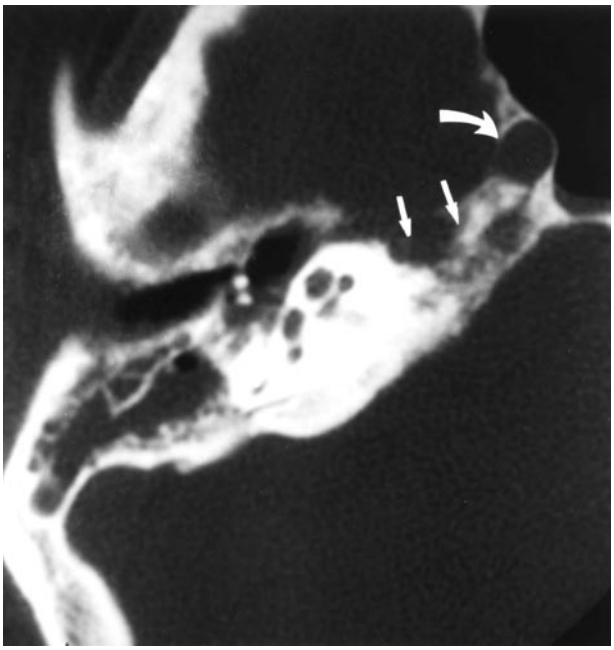


FIGURE 7–59. Metastatic lesion from carcinoma of the breast. Axial computed tomographic section showing a mottled area of bone erosion in the right petrous apex (*arrows*). The carotid canal (*curved arrow*) is intact.

REFERENCES

1. Valvassori GE, Buckinghm RA. Tomography and cross sections of the ear. Philadelphia: WB Saunders; 1975.
2. Valvassori GE, Carter BL, Mafee MF. Imaging of the head and neck. Stuttgart (Germany): Georg Thieme Verlag; 1995.
3. Antonelli PJ, Garside JA, Mancuso AA, et al. Computed tomography and the diagnosis of coalescent mastoiditis. Otolaryngol Head Neck Surg 1999;120: 350–4.
4. Stone JA, Castillo M, Neelon B, Mukherji SK. Evaluation of CSF leaks: high resolution CT compared with contrast-enhanced CT and radionuclide cisternography. AJNR Am J Neuroradiol 1999;20: 706–12.
5. Rodgers GK, Applegate L, De la Cruz A, Lo W. Magnetic resonance angiography: analysis of vascular lesions of the temporal bone and skull base. Am J Otol 1993;14:56–62.
6. Bamiou DE, Phelps P, Sirimanna T. Temporal bone computed tomography findings in bilateral sensorineural hearing loss. Arch Dis Child 2000;82: 257–60.
7. Harnsberger HR, Dahlen RT, Shelton C, et al. Advanced techniques in magnetic resonance imaging in the evaluation of the large endolymphatic duct and sac syndrome. Laryngoscope 1995;105: 1037–42.
8. Seltzer S, Mark AS. Contrast enhancement of the labyrinth on MR scans in patients with sudden hearing loss and vertigo: evidence of labyrinthine disease. AJNR Am J Neuroradiol 1991;12:13–6.
9. Casselman JW, Kuhweide R, Ampe W, et al. Pathology of the membranous labyrinth: comparison of T_1- and T_2-weighted and gadolinium-enhanced spin-echo and 3DFT-CISS imaging. AJNR Am J Neuroradiol 1993;14:59–69.
10. Fitzgerald DC, Mark AS. Sudden hearing loss: frequency of abnormal findings on contrast-enhanced MR studies. AJNR Am J Neuroradiol 1998;19: 1433–6.
11. Mukherji SK, Albernaz VS, Lo WW, et al. Papillary endolymphatic sac tumors: CT, MR imaging and angiographic findings in 20 patients. Radiology 1997;202:801–8.
12. Valvassori GE, Dobben GD. CT densitometry of the cochlear capsule in otosclerosis. AJNR Am J Neuroradiol 1985;6:661–7.
13. Gjuric M, Koester M, Paulus W. Cavernous hemangioma of the internal auditory canal arising from the inferior vestibular nerve: case report and review of the literature. Am J Otol 2000;21:110–4.
14. Shelton C, Harnsberger HR, Allen R, King B. Fast spin echo magnetic resonance imaging: clinical application in screening for acoustic neuroma. Otolaryngol Head Neck Surg 1996;114:71–6.
15. Casselman JW, Kuhweide R, Deimling M, et al. Constructive interference in steady state-3DFT MR imaging of the inner ear and cerebellopontine angle. AJNR Am J Neuroradiol 1993;14:47–57.
16. Dubrulle F, Ernst O, Vincent C, et al. Cochlear fossa enhancement at MR evaluation of vestibular schwannoma: correlation with success at hearing-preservation surgery. Radiology 2000;215:458–62.

CHAPTER 8

Diseases of the External Ear

Timothy T. K. Jung, MD, PhD, Tae Hoon Jinn, MD

The external ear is composed of the auricle, the external auditory canal (EAC), and the epithelial surface of the tympanic membrane. Diseases of the external ear include trauma, infections, and neoplasms.

ANATOMY

The auricle is composed of fibroelastic cartilage to which the skin and a small amount of subcutaneous tissue are closely attached. The skin on the external (anterior) surface of the auricle is attached firmly to the underlying cartilage, with the connective tissue of the dermis condensing to form perichondrium. The skin on the undersurface or posterior surface of the auricle, by contrast, has a true subcutaneous layer. This feature of the auricular integument, combined with the exposed position of the auricle, is responsible for the majority of clinical problems that involve the auricle: trauma, exposure, and infection. Fluid accumulation consequent to these processes results in separation of perichondrium from the cartilage. Unless this separation is promptly relieved, necrosis of cartilage will result because of interference with its perfusion from the vessels of the perichondrium. The topography of the auricle is determined by the contour of its underlying cartilaginous frame (Figure 8–1). The auricle is attached to the skull by a series of ligaments, muscles, and skin.

The EAC is approximately 2.5 cm in length and is divided into bony and cartilaginous parts. The medial two-thirds is osseous, and the lateral one-third is cartilaginous (Figure 8–2). Because of the oblique position of the tympanic membrane, the posterosuperior part of the canal is about 6 mm shorter than the anteroinferior portion. The osseous-cartilaginous junction forms an isthmus that is the narrowest segment of the EAC. Foreign bodies that make their way medial to this point are more difficult to remove.

The EAC serves as a channel for sound transmission to the middle ear and also protects the middle and inner ear from foreign bodies and fluctuation

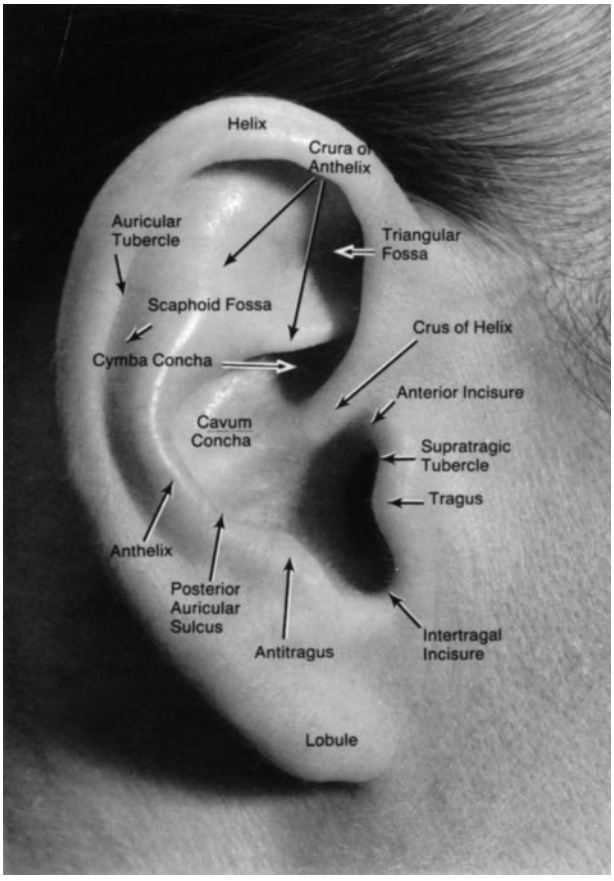


FIGURE 8–1. The right auricle showing the principal anatomic features of its lateral surface. Reproduced with permission from Schuknecht HF, Gulya AJ. Anatomy of the temporal bone with surgical implications. Philadelphia: Lea & Febiger; 1986.

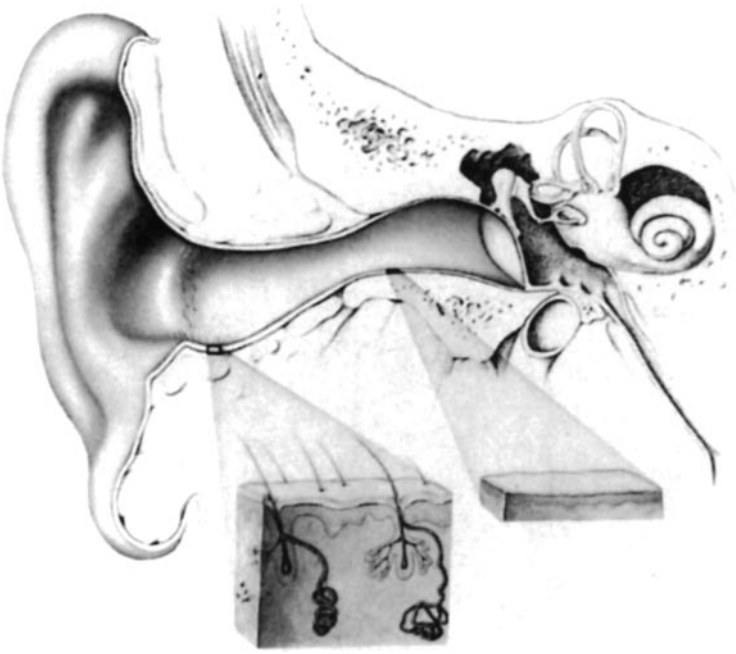


FIGURE 8–2. Coronal section of ear canal with magnification of the skin of the cartilaginous and osseous canals. Reproduced with permission from Senturia BH, Marcus MD, Lucente EF. Diseases of the external ear: an otologic-dermatologic manual. 2nd ed. New York: Grune & Stratton; 1980.

of temperature. The dehiscences in the anterior wall of the cartilaginous portion of the canal are known as the fissures of Santorini, which may allow spread of tumor or infection from the EAC into the parotid gland or temporomandibular joint. The epithelial lining of the EAC is continuous with the epithelial covering of the auricle and the outer layer of the tympanic membrane. The exfoliated squamous cell epithelium and earwax migrate in a lateral direction, serving a self-cleansing function.

The skin of the bony canal is much thinner than that of the cartilaginous portion, about 0.2 mm in thickness, and is continuous with the epithelial layer of the tympanic membrane (see Figure 8–2). There are no glands or hair follicles in the subcutaneous layer. Because of the thinness of the skin in the bony EAC, it can be easily traumatized, for example, during the removal of cerumen.

The skin of the cartilaginous part of the EAC is thicker, averaging from 0.5 to 1 mm, with four layers of epidermis and a true subcutaneous layer. It contains hair follicles and sebaceous and apocrine glands (Figure 8–3).

The external ear is innervated by contributions from the trigeminal (cranial nerve V), facial (cranial nerve VII), glossopharyngeal (cranial nerve IX), and vagal (cranial nerve X) nerves, as well as from the cervical plexus through the greater auricular nerve (C2–3). The extrinsic muscles of the ear are supplied by the facial nerve (cranial nerve VII).

The lymphatic drainage of the EAC is an important channel for the spread of infections or neoplasms. The anterior and superior parts of the EAC drain to the preauricular lymphatics in the parotid gland and the superior deep cervical lymph nodes. The inferior portion of the EAC drains into the infra-auricular lymph nodes near the angle of the mandible. The posterior part of the EAC drains into the postauricular lymph nodes and the superior deep cervical lymph nodes.

TRAUMA TO THE EXTERNAL EAR

Trauma to the external ear is common in all age groups. The unprotected auricle is at risk for all kinds of trauma including cold or hot thermal injury and blunt or sharp injury resulting in ecchymosis, hematoma, laceration, or fracture.

Auricular Hematoma

Hematoma of the auricle usually develops after blunt trauma and is common among wrestlers and boxers. The mechanism usually involves traumatic disruption of a perichondrial blood vessel. Blood accumulation in the subperichondrial space results in separation of

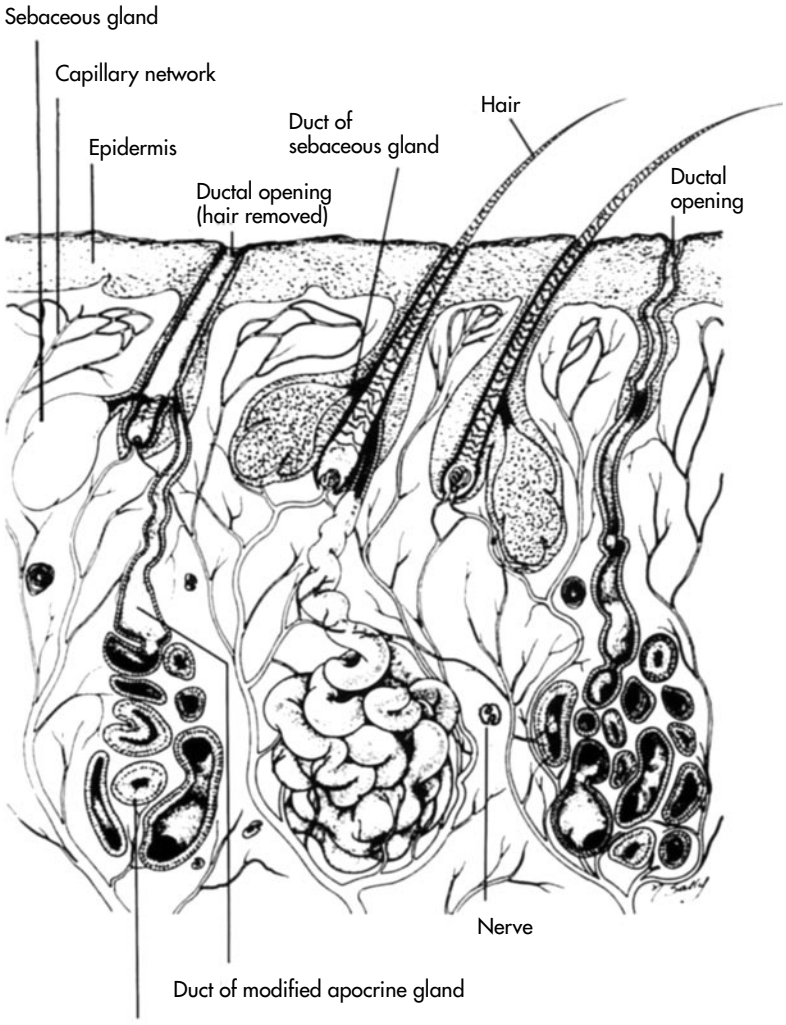


FIGURE 8–3. Schematic drawing illustrating the adnexae and secretory system of the skin of the external auditory canal. Reproduced with permission from Main T, Lim D. The human external auditory canal, secretory system—an ultrastructural study. Laryngoscope 1976;86:1164–76.

perichondrium from the cartilage. If the cartilage is fractured, blood seeps through the fracture line and extends to the subperichondrial plane on both sides.[1] This creates a bluish swelling, usually involving the entire auricle, although it may be confined to the upper half. If the lesion is not treated early, the blood organizes into a fibrous mass, causing necrosis of the cartilage because of interference with its circulation. This mass forms into a twisted scar, especially after repeated trauma, creating the deformity known as "cauliflower ear" (Figure 8–4).

Treatment is based on evacuation of the hematoma and application of pressure to prevent reaccumulation of blood. Simple needle aspiration is inadequate treatment and frequently results in fibrosis and organized hematoma. The most effective treatment for auricular hematoma is adequate incision and drainage with through-and-through suture-secured bolsters (Figure 8–5). The incision should be placed in the scapha, paralleling the helix. Sufficient exposure should be obtained to remove the entire hematoma and to inspect the cavity. If delay has resulted in some organization, sharp ring curettes may be used to remove the clot. Dental rolls are cut to proper size, applied to both sides of the auricle, and tied using through-and-through nylon or silk sutures. An antibiotic ointment is applied over the incision. The dental rolls are left in place for 7 to 14 days.

Lacerations

Auricular lacerations with or without loss of parts of the auricle are common from sharp trauma. Excellent results are possible if sound surgical principles are applied. An attempt should be made to repair, preserving all remaining viable tissue. When

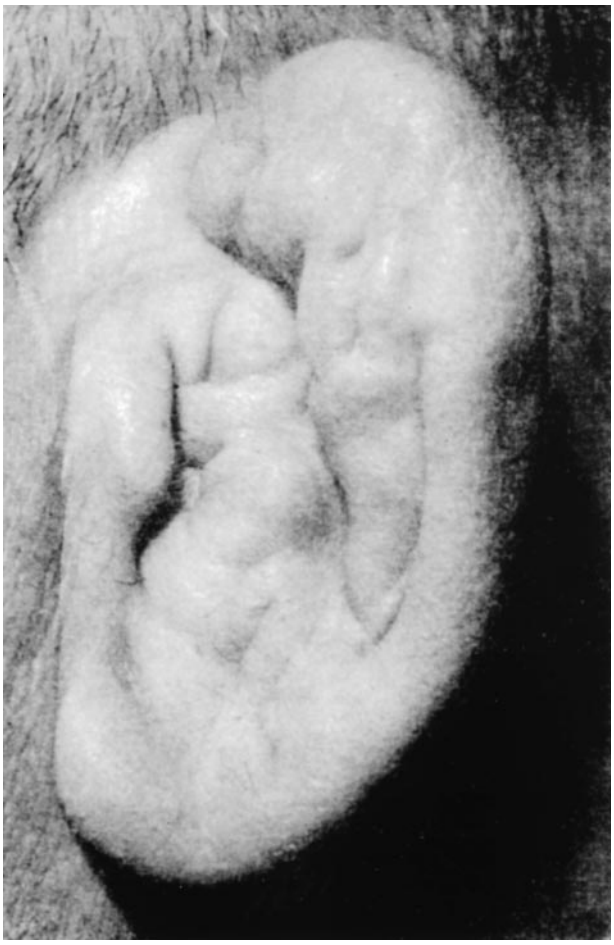


FIGURE 8–4. Cauliflower ear resulting from auricular hematoma.

the auricle is not totally severed, it can be reattached most of the time.

Frostbite

The auricle is particularly susceptible to frostbite because of its exposed location and lack of subcutaneous or adipose tissue to insulate the blood vessels. The anesthesia that develops in the area exposed to severe cold deprives the patient of any warning of impending danger. Initially, there is vasoconstriction, leaving the ear, especially the edges of the helix, blanched and cold to the touch. Hyperemia and edema follow and are caused by a marked increase in capillary permeability. Ice crystallization of the intracellular fluid may be primarily responsible for this, as well as cellular necrosis in the surrounding tissues. The ear becomes swollen, red, and tender, and bullae may form under the skin, resembling a first-degree burn.

The frostbitten ear should be rapidly warmed. Wet sterile cotton pledgets at 38 to 42°C are applied until the ear is thawed. The ear should be treated gently owing to the risk of further damage to the already traumatized and devitalized tissue. Analgesics and prophylactic antibiotics may be necessary. Necrotic tissue is débrided, the topical thromboxane inhibitor aloe vera is applied, and antiprostaglandin drugs such as ibuprofen may be beneficial.[2]

Burns

Burns are traditionally classified in three degrees of severity: erythema (first degree), blistering (second degree), and full-thickness destruction (third degree). Burns caused by scalding liquids or fire are often full thickness. Untreated, they may lead to perichondritis. It is important to avoid pressure on the ear, and gentle cleansing and topical antibiotic applications are used. Prophylactic use of antipseudomonal antibiotics is advocated. The antibiotic may be injected subperichondrially at several different injection sites over the anterior and posterior surfaces of the auricle.[3] Application of mafenide acetate (Sulfamylon) cream after cleaning the wound is recommended. In the late stage, débridement and skin grafting may be necessary. Perichondritis and chondritis should be treated with antibiotic iontophoresis, early débridement, and grafting.

Cerumen

Cerumen is a combination of the secretions produced by sebaceous (lipid-producing) and apocrine (ceruminous) glands admixed with desquamated epithelial debris. This combination forms an acidic coat that aids in the prevention of EAC infection. The pH of the cerumen is high in diabetic patients compared with 6.5 to 6.8 in the normal EAC.[4] There are genetically and racially determined differences in the physical characteristics of cerumen that vary its appearance and consistency and may be associated with immunoglobulin and lysozyme content.[5] Some individuals have a scanty amount, and others tend to form obstructive masses. The geriatric and mentally retarded populations have a tendency to accumulate excess cerumen. Accumulation of cerumen represents the most common and routine otologic problem. It may interfere with the clinician's view of the tympanic membrane, cause hearing loss and discomfort, or

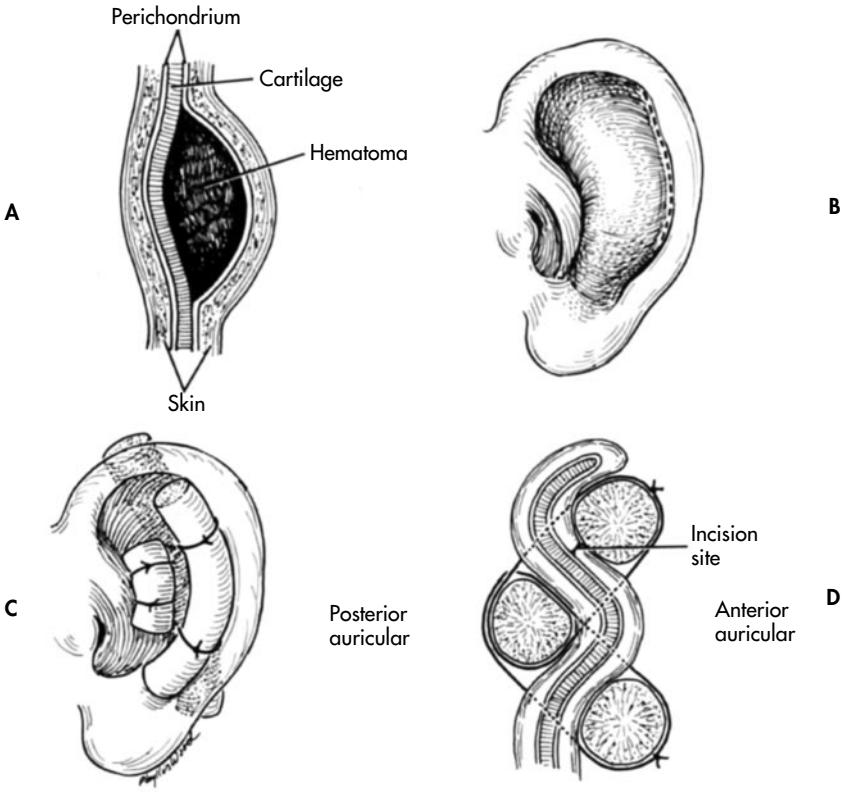


FIGURE 8–5. Otohematoma. *A*, Hematoma of the auricle. *B*, Hematoma incised and evacuated. *C*, Anterior dental rolls tied to posterior dental roll on the surface of the ear. *D*, Side view, showing how bolsters are secured. Reproduced with permission from Clemons J, Severeid LR. Trauma. In: Cummings CW, Fredrickson JM, Harker LA, et al, editors. Otolaryngology-head and neck surgery. Vol 4. St. Louis: Mosby Co.; 1986. p. 2912.

become a source of infection. Some patients make routine attempts to remove cerumen with cotton swabs, making it worse by pushing cerumen medially.

Several techniques are available for the removal of cerumen and may be used in a variety of ways. Before starting to remove cerumen, one should make sure that the patient does not have a history of tympanic membrane perforation. If perforation is suspected, the irrigation method should not be used. The irrigation method works best for soft and greasy cerumen. The canal may be irrigated with warm water, either with a syringe or with a pressure-driven irrigating bottle (Figure 8–6). The canal is straightened by pulling the auricle up and back. The water stream is directed along the superior canal wall, and outflow is caught in a basin held below the ear. Remaining irrigating solution or residual cerumen can be suctioned out using a Frazier No. 5 or 7 suction catheter. An alternative method is the use of a cerumen curette to dislodge and remove the cerumen. First the cerumen and desquamated layer of epithelium are gently separated from the canal wall with a ring curette, and the loosened cerumen is grasped with an alligator forceps and teased out. The use of an operating microscope with the patient in a supine position helps to prevent head movement and to perform this procedure painlessly.

If impaction of hard cerumen persists or is too painful to remove, the patient may be sent home with instructions for using an agent to soften the cerumen. Such agents include a variety of common corticosteroid and antibiotic otic drops, ceruminolytic solutions (Cerumenex), or hydrogen peroxide. Following its use for a few days, the patient is re-examined and the softened remaining cerumen can be removed with irrigation or suction. This approach is particularly useful in pediatric patients.

Foreign Bodies

Foreign bodies in the EAC are found most frequently in the pediatric age group or in mentally retarded institutionalized patients. Any objects small enough to enter the EAC can become prospective foreign bodies. These include animate, inanimate, or mineral objects. They may cause symptoms of irritation, pain, and hearing loss.

The removal of a foreign body can be safely done under direct visualization, preferably under an operating microscope with the patient in a supine position. Instruments helpful for this task include

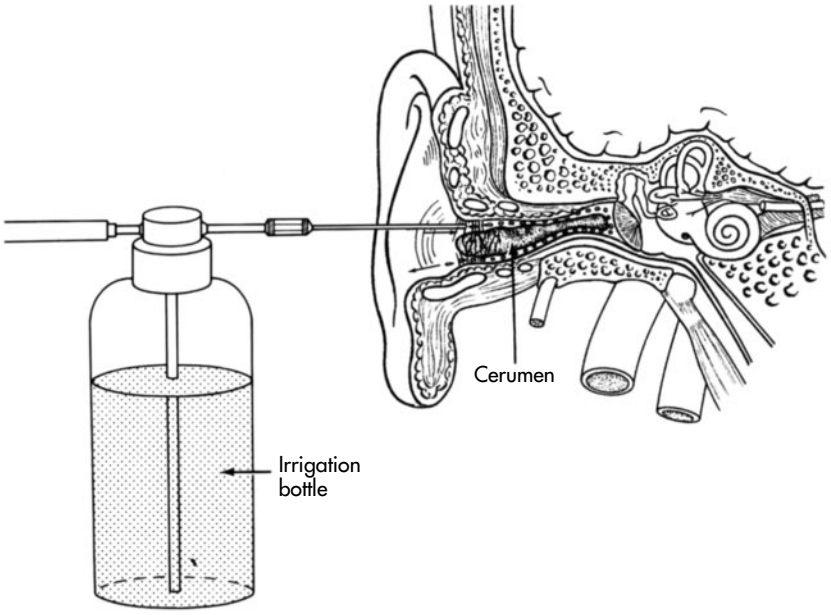


FIGURE 8–6. Cerumen irrigation with compressed air irrigation bottle system. Reproduced with permission from Goodhill V. Ear diseases, deafness, and dizziness. Hagerstown (MD): Harper & Row; 1979.

the alligator forceps, ring curettes, and hooks. Inanimate objects located lateral to the isthmus of the canal are removed with an alligator forceps or by placing a hook or ring curette behind it and pulling it out. Suctioning with Frazier suction catheters is useful in removing an object with a smooth surface that is hard to grasp. Irrigation can be used in certain instances. Objects located medial to the isthmus of the canal are more difficult to remove and may require local or general anesthesia.

Some foreign bodies may incite more inflammatory reaction and be damaging. Vegetable and plant debris can adhere to the skin of the EAC or tympanic membrane. Miniature round batteries used for cameras and hearing aids, once placed in the canal, can cause reaction and damage canal skin or even the tympanic membrane. A live insect in the EAC should be immobilized before removal by instilling mineral oil or alcohol into the canal.

The best chance for removal of a foreign body in the EAC is the first attempt. When it fails, the ear may become extremely painful, and proper anesthesia may be necessary. Four-quadrant canal skin injection with a local anesthetic agent is sufficient in adults. For young children, it is best to use general anesthesia.

Fractures of the External Auditory Canal

A strong blow to the mandible can drive the mandibular condyle into the ear canal, resulting in fracture of the anterior canal wall. The patient is treated for this fracture by repositioning lacerated or avulsed tissue and bone in the canal and packing the canal with antibiotic-saturated gauze. Fractures of the canal can be a part of temporal bone fractures. Longitudinal temporal bone fractures may extend into the bony ear canal, usually passing through the bony tympanic ring at the junction of the scutum and the tympanomastoid suture. Blood with cerebrospinal fluid may drain for a while. There may be an area of ecchymosis over the mastoid (Battle's sign). These fractures usually heal spontaneously with an occasional stenosis remaining in the bony annulus. If the patient develops a conductive hearing loss owing to ossicular damage or cholesteatoma caused by entrapped skin in the fracture line, tympanoplasty and ossiculoplasty or tympanomastoidectomy may be required later.

INFECTION AND INFLAMMATION OF THE EXTERNAL EAR

Infection and inflammation may involve skin or cartilage of the auricle, EAC, or epithelial layer of the tympanic membrane. It may be acute or chronic. The infectious agent may be bacterial, fungal, viral, or mixed.

Auricle

Cellulitis of the Auricle Cellulitis is a bacterial infection that usually follows abrasion, laceration, or

ear piercing. The auricle is red, swollen, painful, and tender to manipulation. It is usually caused by gram-positive cocci such as *Staphylococcus* or *Streptococcus* and rarely other microorganisms such as *Pseudomonas*. In the absence of a history of trauma, a topical allergic reaction or relapsing polychondritis should be considered. Treatment includes oral or intravenous antibiotic and wound care. Erysipelas is a cellulitis caused by group A β-hemolytic *Streptococcus* and may involve the auricle. It is marked by systemic toxicity with fever and chills, erythema, pain, and swelling. It is contagious. Treatment of choice is oral or intravenous penicillin G and wound care.

Allergic Dermatitis of the Auricle Allergic dermatitis of the auricle is characterized by localized erythema, swelling, and itching in the area of allergen exposure. A patient with neomycin allergy, who has been using eardrops containing neomycin, will present with swelling and redness in the area exposed to the drops, such as the meatus, EAC, and inferior part of the auricle. A patient with metal allergy will present with a red swollen ear lobule owing to contact with the earring. Treatment includes removal of the allergen, topical corticosteroid cream, and oral antihistamines.

Perichondritis and Chondritis Perichondritis or chondritis is a bacterial infection of perichondrium or cartilage of the auricle. This condition may follow inadequately treated auricular cellulitis, acute otitis externa, accidental or surgical trauma, or multiple ear piercing in the scapha. The affected ear is painful, red, and swollen and drains serous or purulent exudates. The surrounding soft tissues of the face and neck may be affected. The most common pathogen is *Pseudomonas*.

In the early stage, oral fluoroquinolone antibiotics (ciprofloxacin [Cipro], levofloxacin [Levaquin]), local antibiotic drops, and débridement are sufficient. In the advanced stage with involvement of regional lymph nodes and surrounding soft tissue, the patient may need to be hospitalized with aggressive intravenous antibiotics using ceftazidime or fluoroquinolones and local treatment. Topical antibiotic irrigation with indwelling catheters may be tried.

Relapsing Polychondritis Relapsing polychondritis is an autoimmune disease manifested by intermittent episodes of inflammation of cartilage throughout the body. Type II collagen antibodies have been found in these patients.[6] Auricular cartilage is most commonly involved, whereas nasal and laryngeal cartilages are less frequently involved.[7] The typical patient presents with a red, swollen tender auricle. Recurrent episodes may result in a floppy and distorted auricle. The disease may involve both auricles simultaneously, or there may be alternate involvement. The patient may develop a saddle deformity with destruction of nasal septum and hoarseness and subglottic stenosis owing to involvement of the larynx or trachea.

Treatment of relapsing polychondritis includes corticosteroid, salicylate, or indomethacin for acute episodes. Dapsone (Avlosulfon), 100 mg once or twice daily after an initial trial of 50 mg/day, has been successfully used for chronic disease with systemic manifestations.[8]

External Auditory Canal

Although the EAC is a well-protected and self-cleansing structure, various forms of infection may develop in the EAC. Otitis externa is one of the most common diseases in clinical practice.

Acute Localized Otitis Externa (Furuncle) Acute localized otitis externa is an infection of a hair follicle, beginning as a folliculitis but usually extending to form a small abscess or furuncle. The infecting microorganism is usually *Staphylococcus aureus*. It involves the lateral cartilaginous portion of the EAC, usually at the meatus. Pain is severe in this condition, and examination is difficult owing to pain and swelling. If the abscess occludes the canal, hearing loss may develop. Discharge is not usually present until the abscess ruptures.

Acute localized otitis externa should be treated the same as an abscess in any part of the body. If it is seen before suppuration has taken place, resolution may occur with the use of topical and systemic antibiotics. If a localized abscess has formed, it should be treated by incision and drainage and topical antibiotic ointment with or without oral antibiotics.

Acute Diffuse Otitis Externa (Swimmer's Ear) *Etiology/Predisposing Factors.* Acute diffuse otitis externa is a bacterial infection of the EAC and the most common form of otitis externa. It is caused by the removal of the protective lipid film from the canal, allowing bacteria to enter. It usually begins with itching in the canal and skin maceration and

local trauma from scratching the canal with a cotton swab, bobby pin, fingernail, or other object. Predisposing factors include frequent swimming; a warm and humid climate; a narrow and hairy ear canal; presence of exostosis in the canal; trauma or foreign body in the canal; impacted or absent cerumen; use of hearing aids or earplugs; diabetes or an immunocompromised state; skin conditions such as eczema, seborrhea, and psoriasis; and excessive sweating. Absence of cerumen may be a predisposing factor because the act of cerumen removal may be traumatic and lead to breaks in the fragile EAC skin, and cerumen serves an antimicrobial role through physically protecting the EAC skin; establishing a low-pH, inhospitable environment for pathogens; and containing compounds such as lysozyme.

The usual pathogens for acute diffuse otitis externa are *Pseudomonas aeruginosa, Proteus mirabilis,* or *S. aureus*. Culture of the canal is usually not done except for recurrent or recalcitrant infections.

Diagnosis/Stage of Disease. The typical patient presents with one or all of the symptoms including pain, itching, fullness, and hearing loss. On physical examination, there are various degrees of tenderness and narrowing of the ear canal with red swollen skin. Clear to purulent exudates may be present. One of the classic signs is pain elicited by pulling the auricle upward and backward.

Senturia and colleagues proposed three clinical stages of otitis externa: preinflammatory, acute inflammatory, and chronic inflammatory.[9,10] The acute inflammatory stage may be mild, moderate, or severe. The preinflammatory stage begins when the stratum corneum becomes edematous owing to the removal of the protective lipid layer and acid mantle from the canal, resulting in itching, edema, and the sensation of fullness. The mild acute inflammatory stage is characterized by increased itching and pain. Mild erythema and edema are present on examination. As inflammation increases to the moderate phase, the patient complains of intermediate pain and itching. The lumen of the EAC is decreased by edema and debris from the thickened, irritated skin (Figure 8–7). Secretions are exudative and more profuse. In the severe inflammatory stage, the canal lumen becomes completely obstructed because of the increasing hyperemia, edema, and purulent otorrhea. The patient complains of intense pain, especially on chewing or tragal manipulation. In the severe stage, the clinician often sees evidence of extension of infection beyond the EAC to involve adjacent soft tissues and cervical lymph nodes.

Treatment Certain principles of management must be observed in every case of otitis externa. These are frequent inspection and cleansing of the canal, control of pain, use of specific medication appropriate to the type and severity of the disease, acidification of the canal, and control of predisposing causes. Frequent inspection with cleaning using suction-débridement under an operating microscope and drying of the ear canal is the single most important step in obtaining resolution of all forms of otitis externa.

The preinflammatory and mild stage of inflammation can be managed with thorough cleaning and débridement of the canal, preferably under a microscope. Acidifying/antiseptic agents such as gentian violet may be applied. Antibiotic-hydrocortisone drops can be used for a few days. Patients are advised to avoid water and resist using cotton swabs or digital manipulation of the ear canal.

In the moderate stage of inflammation, treatment may include gentle cleaning-débridement of the canal and application of acidifying/antiseptic/ antibiotic agents. If the canal is severely swollen, a cotton or Pope wick may be inserted and otic drops instilled on it. A variety of eardrops are available for treatment. Most of the eardrops contain a combination of antipseudomonal antibiotics with or without corticosteroids (neomycin sulfate, colistin sulfate [Cortisporin], ofloxacin [Floxin], ciprofloxacin hydrochloride [Cipro HC]). Most of these agents are acidic to inhibit the proliferation of bacteria and fungi. The major adverse reaction to the use of an acidic agent is burning on application. Ophthalmic preparations tobramycin and dexamethasone (gentamicin, [Tebradex], ciprofloxacin [Ciloxan]) are pH neutral and may be tolerated better than otic drops. An additional advantage of ophthalmic drops is a low viscosity, allowing improved penetration. Fluoroquinolone antibiotics, such as ciprofloxacin and ofloxacin, may be a better choice because of an appropriate antimicrobial spectrum and low ototoxicity, especially for patients with neomycin sensitivity.[11–13] Instillation of otic powder containing a combination of ciprofloxacin/chloramphenicol, amphotericin B (Fungizone), and hydrocortisone is helpful to deliver high doses of medications locally. An oral analgesic is often

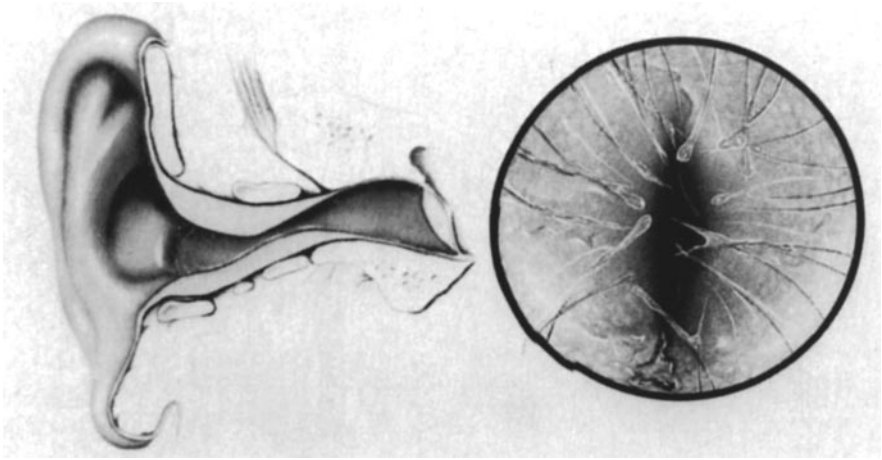


FIGURE 8–7. Moderate inflammatory phase of acute diffuse external otitis. Canal skin has become more erythematous and edematous. Greenish secretion coats skin. Reproduced with permission from Senturia BH et al.[10]

needed. System antibiotics are usually not necessary. Patients are advised to avoid any predisposing factors.

In the severe stage, the lumen of the EAC may be swollen shut by edema and debris to such a degree that antibiotic drops cannot be instilled into the canal. As discussed above, a cotton or Pope wick must be gently inserted into the canal to carry the topical medication to the affected canal skin. After 2 or 3 days, the canal usually opens, permitting cleaning and instillation of the medication. If infection extends beyond the limit of the EAC, oral antipseudomonal quinolone antibiotics (eg, ciprofloxacin, levofloxacin) should be used in adults. For pediatric patients, intravenous ceftazidime should be considered.

Chronic Otitis Externa Chronic otitis externa is a low-grade, diffuse infection and inflammation of the EAC that persist for months or years. It is characterized by pruritus and dry hypertrophic skin of the EAC. It results in a thickening of the EAC skin and progressive narrowing of the lumen of the EAC (postinflammatory stenosis). Although bacterial or fungal infection is the main cause of chronic otitis externa, skin conditions, such as seborrheic dermatitis, psoriasis, neurodermatitis, and sensitization to an otic drop, can result in chronic otitis externa.

The goal of treatment is to prevent the stenosis and restore the EAC skin to its normal healthy state. Antibiotics and corticosteroid otic drops can decrease the inflammation and edema of the EAC. The routine use of antibiotic otic drops such as neomycin is not recommended because of the high risk of sensitization. The most important treatment is frequent inspection and thorough débridement of the EAC. The canal can be painted with gentian violet, and otic powder may be applied. Painting the canal with triamcinolone/nystatin (Mycolog) cream helps to relieve pruritus and clear chronic otitis externa. Patients are instructed not to touch the EAC or use a cotton swab.

Surgical Treatment. In case of medical treatment failure, especially with canal stenosis, surgical procedures are indicated to restore canal patency and hearing. Canalplasty with skin graft can be used to enlarge the canal (Figures 8–8 and 8–9).[14,15] The abnormal canal skin is removed entirely. The denuded canal is enlarged, using diamond burs, to an optimum size, increasing gradually in diameter. A split-thickness skin graft is harvested from the upper medial surface of the arm with a dermatome. The graft is placed in the canal to cover all exposed surfaces. A "rosebud" type of packing with Owen's silk strips and cotton balls is placed over the skin graft for 2 weeks (see Figure 8–9). Crusting may occur for several weeks and requires meticulous removal and cleansing until complete healing takes place.

Prevention. Preventive measures are recommended in patients who demonstrate a propensity for recurrent episodes of otitis externa. Patients are instructed not to touch or place any objects such as cotton swabs, paper clips, or any other objects into the canal. Swimmers are instructed to use earplugs and are advised to use alcohol-vinegar (1:1) drops after swimming.

Necrotizing (Malignant) External Otitis Malignant external otitis is a progressive, potentially lethal

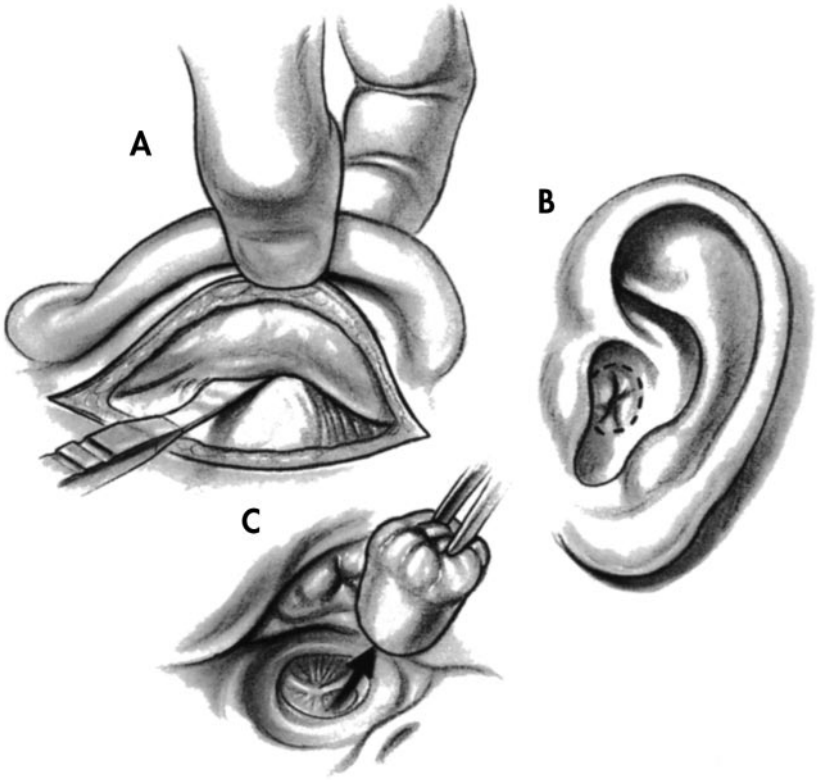


FIGURE 8–8. Canalplasty for refractory otitis externa. *A*, A postauricular incision is made and the cartilaginous canal is sectioned. *B*, A circumferential through-and-through incision is made at the meatus. *C*, A diseased stenotic plug of canal skin is removed along with the infected meatus and outer squamous cell layer of the tympanic membrane. The bony canal is widened by drilling. Reproduced with permission from Jung TT.[15]

infection of the EAC, surrounding tissue, and skull base typically seen in elderly diabetic or other immunocompromised patients. The term malignant external otitis was first coined and presented in a case series by Chandler in 1968.[16] Diabetic patients were almost exclusively the population at risk. However, a number of cases of malignant external otitis without diabetes have been documented.[17,18] Most of these patients had an underlying disease that resulted in neutropenia or immunosuppression, such as leukemia, treatment with bone marrow–suppressive drugs, and acquired immune deficiency syndrome. *Pseudomonas aeruginosa* is the most common pathogen in malignant otitis externa.

Pathophysiology. The infection begins in the EAC. The local infection in the EAC progresses to cellulitis, chondritis, osteitis, and, finally, osteomyelitis. The disease may gain access to the osseous auditory canal and skull base through Santorini's fissures. From there, the infection usually progresses along the base of the skull. Once access to the bony skull base has occurred, there is progressive replacement of compact bone with granulation tissue. Facial nerve paralysis occurs when there is involvement of the stylomastoid foramen, and cranial nerve IX, X, or XI palsies occur when the jugular foramen becomes involved. The jugular vein may become thrombosed, progressing to a lateral sinus thrombosis, which usually ends in death.

Diagnosis. A high index of suspicion for malignant external otitis is needed in diabetic or immunocompromised patients with otitis externa. Granulation tissue in the floor of the EAC near the bony-cartilaginous junction is a typical otoscopic finding. Patients usually complain of severe otalgia. Facial nerve paralysis and jugular foramen syndrome are poor prognostic signs.

A culture of purulent discharge usually reveals *P. aeruginosa*. Sensitivity testing to all antipseudomonas antibiotics should be sought. The erythrocyte sedimentation rate is a nonspecific test but is useful in following the response to antibiotics. A biopsy of the granulation tissue in the EAC is necessary to exclude carcinoma of the ear canal or *Aspergillus* skull base osteomyelitis since they have clinical and radiologic findings similar to those in necrotizing external otitis.

Radiologic examinations are required to determine the extent of disease. High-resolution com-

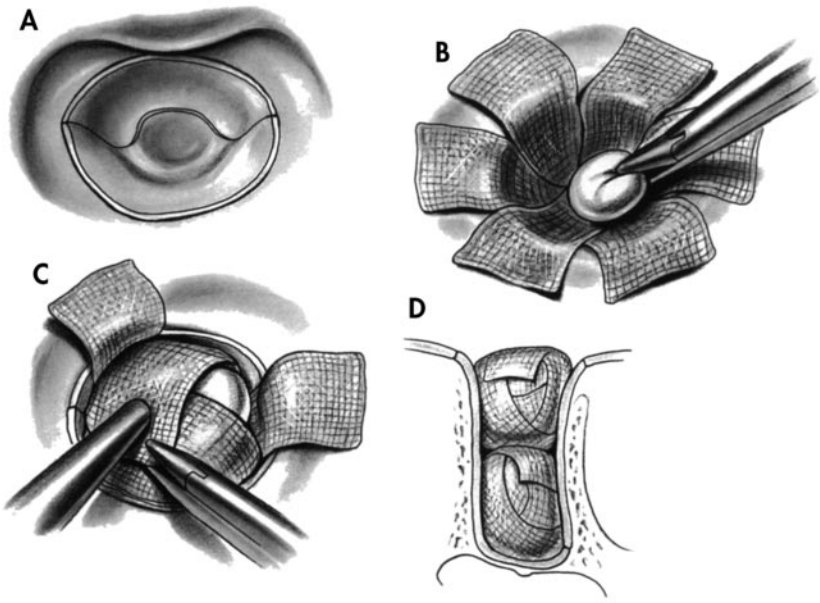


FIGURE 8–9. Canalplasty for refractory otitis externa, skin graft. *A*, The longer convex piece of skin is laid out on the drumhead and posterior half of the canal. The shorter concave piece covers the anterior half of the canal. *B* and *C*, A "rosebud" packing is placed with strips of Owen's silk and pieces of cotton balls saturated in an antibiotic-corticosteroid suspension. *D*, A second pack may be placed lateral to the first pack. Reproduced with permission from Jung TT.[15]

puted tomography (CT) is recommended to assess the extent of disease at initial evaluation.[19] It is of little use in monitoring the resolution of infection because remineralization is required for the bony changes to return to normal. Magnetic resonance imaging (MRI) is of no use in detecting bony changes, but it is better for evaluation and follow-up of soft tissue involvement, especially meningeal enhancement and changes within the osseous medullary cavity.[19]

Bony involvement in patients with early malignant external otitis can be detected with technetium scan (bone scan).[20] Although it is nonspecific, bony involvement can be detected when there is absence of destruction on CT scans.[21] But a technetium scan may remain positive even after the resolution of osteomyelitis. Thus, it is not useful for therapeutic monitoring.

A gallium scan is a sensitive indicator of infection. In the presence of infection, there is increased uptake of gallium, which may occur in soft tissue or bone. The gallium scan becomes negative as the infection clears. Therefore, the gallium scan is best for monitoring resolution of disease and is useful in determining duration of antimicrobial therapy.[20,22,23]

Treatment. The standard treatment is hospitalization of the patient and treatment of diabetes, if present, together with the use of high doses of antibiotics specific for *Pseudomonas* for an extended period of time. Daily débridement of the EAC is performed, and culture of débrided material and sensitivity testing should be repeated during the initial phase of management.

Antipseudomonal otic drops are instilled with placement of a wick when necessary. Standard antibiotic therapy has been aminoglycosides combined with either an antipseudomonal penicillin or cephalosporin for dual-drug therapy as primary intervention. Special precaution must be taken during aminoglycoside treatment because of nephrotoxicity and ototoxicity. Blood levels must be obtained regularly to ensure adequate dosage. Creatinine levels should be obtained three times a week to measure renal function. Electrolyte disturbances such as hypokalemia may occur; therefore, regular determination of electrolytes should be performed. Periodic hearing tests should be obtained if possible, particularly of the uninvolved ear. Because of potential nephrotoxicity/ototoxicity of aminoglycosides, oral quinolones (eg, ciprofloxacin) have been used successfully as an alternative to treat malignant external otitis.[24–26] Quinolone antibiotics have the advantage of being equally effective whether administered parenterally or orally. The duration of antimicrobial therapy

depends on serial gallium scans performed at 4-week intervals.[21] A strict control of diabetes is essential for the treatment of malignant external otitis. Without the control of diabetes, malignant external otitis cannot be controlled.

Pain can be severe and worsens as the infection fails to respond. A mastoidectomy with possible facial nerve decompression or subtotal petrosectomy or even partial temporal bone resection may be necessary. Hyperbaric oxygen therapy has been used as an adjunctive measure. Lessening of pain is the first indication of a favorable therapeutic response.

Fungal Otitis Externa (Otomycosis) Fungal infection of the EAC is usually caused by *Aspergillus* and/or *Candida*.[27] Otomycosis develops most commonly in a patient with chronic otitis externa that has been treated with a variety of topical or oral antibiotics. The most common presenting symptom is pruritus. Other common symptoms include hearing loss, drainage, discomfort, and feeling something in the ear. Otomicroscopy reveals erythematous canal skin with black, gray, or white fungal growth on the canal skin or on the debris in the canal. It may progress to a complete occlusion of the canal lumen with a fungal ball. Treatment consists of frequent and meticulous cleansing and topical medication. A topical agent, such as gentian violet or thimerosal, can achieve acidification and drying of the ear canal. Topical antifungal agents may be effective in powder form and/or in a cream (Mycolog). Oral antifungal agents can be used in the rare case refractory to topical therapy.

Herpes Zoster Herpes zoster is the most frequent virus to affect the external ear. The virus stays dormant in the sensory ganglia and reactivates under conditions of decreased immune competence. The virus causes blisters on the auricle, the EAC, and even on the lateral surface of the tympanic membrane. Typically, the blisters are short-lived, rupture, dry, crust, and heal. Some patients with herpes zoster infection develop involvement of the facial and cochleovestibular nerves. This clinical syndrome with facial palsy, with or without hearing loss and dizziness, owing to herpes zoster is called *herpes zoster oticus* or *Ramsay Hunt syndrome*. This infection is mostly self-limiting, and treatment is primarily supportive. Patients with full-blown herpes zoster oticus are treated with acyclovir (Zovirax) and corticosteroid.

Tympanic Membrane

Bullous Myringitis Bullous myringitis is an infection of the tympanic membrane characterized by rapid onset, severe pain, and varying sizes of blister formation on the tympanic membrane and adjacent bony ear canal. Although virus, *Mycoplasma*, and other bacteria have been suspected, definite causative microorganisms have not been established. The bullae may be filled with serous or hemorrhagic fluid. Treatment includes analgesia, topical antibiotics, and corticosteroid drops. Rupturing the blisters and packing or irrigation of the canal should be avoided.

Granular Myringitis Granular myringitis is an inflammation of the tympanic membrane characterized by persistent granulation tissue on the lateral surface of the pars tensa with an intact tympanic membrane. The eardrum is thickened, inflamed, and moist with foul-smelling drainage. The patient complains of mild otalgia and drainage. The etiology is not clear, but granular myringitis is believed to be caused by local trauma or infection that denudes the epithelial layer of the tympanic membrane. Treatment includes cauterization with silver nitrate sticks and topical antibiotic and corticosteroid drops. Most cases respond to local treatment, but some progress to inflammatory thickening of the tympanic membrane and even to obliteration of the medial part of the EAC.

TUMORS OF THE EXTERNAL EAR

Skin and its appendages, bone, cartilage, and other connective tissues that make up the external ear may give rise to a variety of benign and malignant tumors.

Benign Tumors of the External Ear

Vascular Tumors *Angiomas* are congenital tumors and are one of the most common tumors of childhood. They may involve the auricle together with other areas of the face and neck. These tumors occur in various forms. *Capillary hemangioma* consists of masses of capillary-sized vessels and may form a large flat mass. A central large vessel feeds the "port-wine stain" or spider nevus, which is a branching network of capillaries. The *spider nevus* is not a major problem, being small and fixed in size. Treatment, when necessary, usually consists of needle coagulation of the central vessel. The port-wine stain

is much more of a problem, increasing in size gradually until adolescence, and generally is disfiguring.

Cavernous hemangioma is the most alarming of these lesions, consisting of raised masses of blood-filled endothelial spaces. Often termed a "strawberry tumor," it increases rapidly in size during the first year of life but usually regresses thereafter. Much less common is the *lymphangioma*. It has the appearance of multiple pale circumscribed lesions, like a cluster of fish or frog roe.

The major problem in these tumors is cosmetic. In general, the lesion should be allowed to regress maximally and the residual tumor treated. Various modalities have been recommended, including cryosurgery, surgical excision and skin grafting, radiation, electrolysis, and tattooing for port-wine staining. Therapy should only be undertaken with caution and after the best available consultation has been sought.

Cysts Sebaceous cysts are common around the ear. They usually occur on the posterior surface of the lobule, in the skin over the mastoid process, and in the skin of the inferior or posterior parts of the cartilaginous portion of the canal (Figure 8–10). These soft, nontender swellings may become infected and may be confused with furuncles. Treatment for a sebaceous cyst is excision. Infection, if present, is treated first. The cyst is removed by sharp dissection, with care being taken to keep the walls of the cyst intact to ensure complete removal. The ductal tissue leading to the cyst as well as its external opening are removed, including a small segment of the overlying skin.

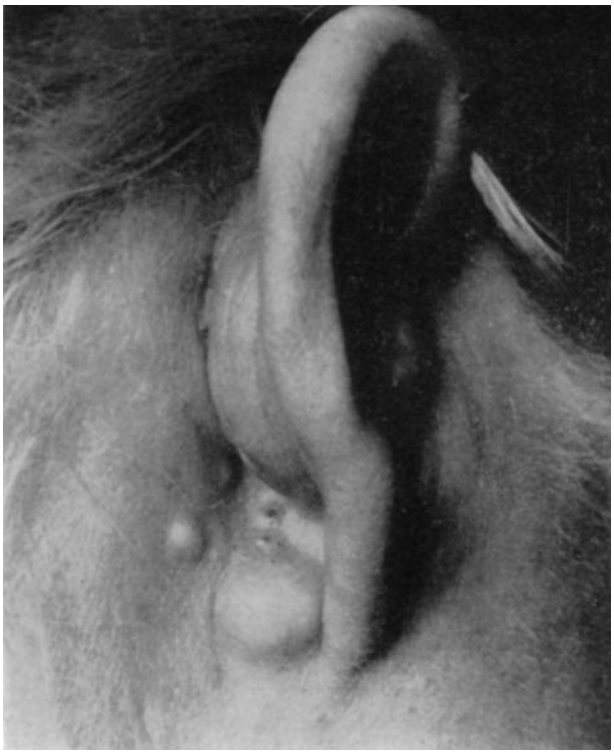


FIGURE 8–10. Sebaceous cysts in the postauricular area.

Preauricular Pits and Sinuses Preauricular pits and sinuses are of congenital origin, arising from faulty developmental closure of the hillocks of first and second branchial arches that form the auricle. They present as small openings in the skin just anterior to the crus of the helix (Figure 8–11). From this opening, a long branched tract may run under the skin between the helix and tragus and anterior to the tragus. The tract, which is lined with squamous cell epithelium, is often cystic, and the patient is frequently seen initially because of infection of the cyst.

Treatment is not necessary unless recurrent infection occurs. Treatment includes complete removal of the cyst along with the fistula tract. Incomplete removal is associated with the formation of draining sinuses, requiring even more difficult and radical surgery for their elimination. The difficulty of the surgery is caused by the branching of the fistula, which makes it hard to define the complete extent of the tract. One suggestion to aid in their removal is to inject the tract before the operation with methylene blue so that the stained tissue may be used as a rough guide to the extent of the fistula.

Keloids Keloids represent an unchecked healing response in individuals with a genetic predisposition to their development. The formation of a keloid is stimulated by trauma to the skin. A keloid is composed of massive collections of collagen interspersed with active fibroblasts and normal thin collagenous strands. Keloid formation is seen most frequently in dark-skinned races, particularly in blacks. Keloids around the ear are most frequently seen as pedunculated tumors on the lobule following ear piercing. They may also occur in mastoidectomy and endaural scars, resulting in disfigurement or stenosis of the canal. Treatment is excision followed by periodic injection of a small amount of corticosteroids (eg, triamcinolone) into the surgical site to prevent keloid reformation.

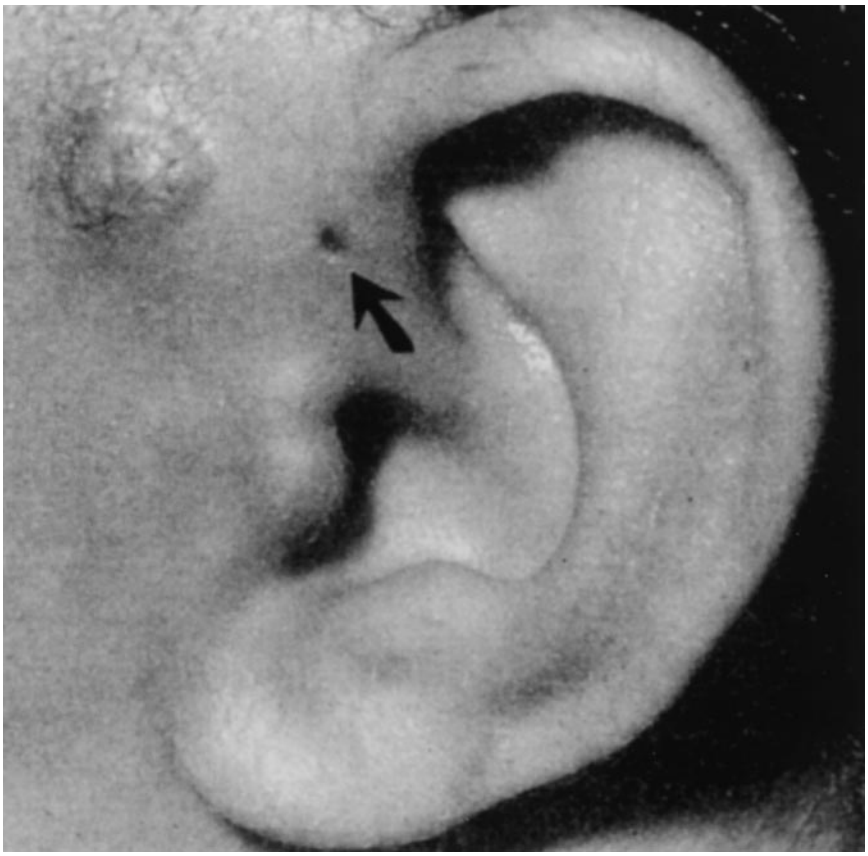


FIGURE 8–11. Preauricular cyst and fistula. *Arrow* indicates the fistula opening. Anterior to this is an inflamed swelling owing to an infected preauricular cyst.

Winkler's Nodule (Chondrodermatitis Nodularis Chronica Helicis) Chondrodermatitis nodularis chronica helicis is a benign nodular growth usually occurring on the rim of the helix in older men (Figure 8–12). It appears as a firm elevated nodular lesion with a grayish crust on the surface. It is characterized by exquisite tenderness with digital compression, out of proportion to its size. It must be differentiated from other lesions such as basal cell carcinomas. The cause of chondrodermatitis nodularis chronica helicis is unknown. It can be treated with injection of a corticosteroid for pain relief. Definitive treatment requires full-thickness excision, including a wedge of cartilage.

Keratoacanthoma Keratoacanthoma is a benign tumor resembling carcinoma and is believed to be related to actinic exposure. The common location of the tumor is anterior to the tragus. It is characterized by a central crater that contains a keratin plug. The lesion tends to grow rapidly after its initial appearance and then slowly regresses. Although the disease is self-limiting, excisional biopsy is required to rule out a malignant tumor.

Keratosis Obturans Keratosis obturans is characterized by the accumulation of large plugs of desquamated keratin in the bony portion of the EAC. This condition is usually asymptomatic and bilateral and is discovered incidental to examination for conductive hearing loss. Removal of the keratin plug reveals diffuse hyperemia of the canal skin. There is marked inflammation in the subepithelial tissue but no bone erosion. The cause is unknown. It has been associated with chronic pulmonary disease and sinusitis. Bronchiectasis has been present in a high percentage of these patients, especially those with an onset before the age of 20 years. The disease may be controlled in most cases by regular cleaning.

Cholesteatoma of the EAC and *keratosis obturans* are different clinical and pathologic processes.[28] External auditory canal cholesteatoma consists of a localized erosion of the bone by squamous cell tissue. It is usually unilateral and is not associated with systemic disease. Treatment consists of periodic cleaning and surgery.[29]

Exostosis and Osteoma Exostoses are the most common tumors of the EAC. A benign growth of

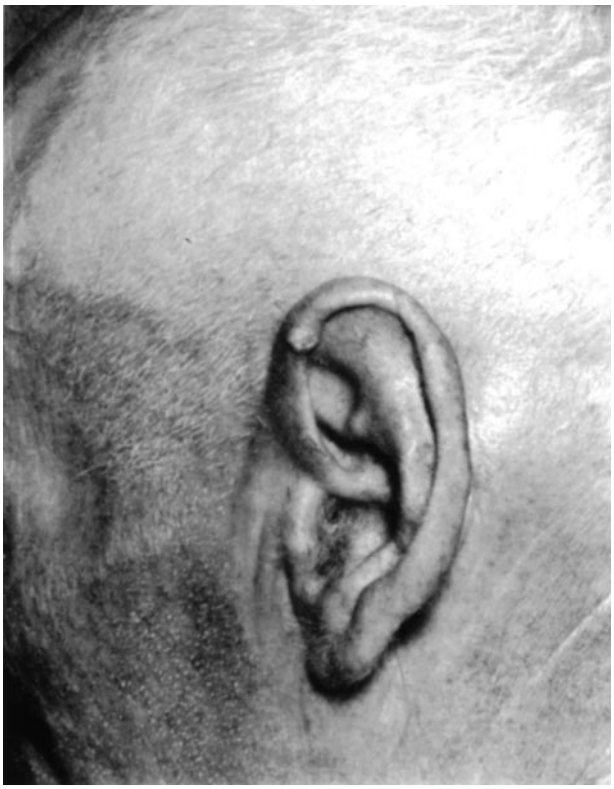


FIGURE 8–12. Typical raised, tender, firm nodule of chondrodermatitis nodularis chronica helicis.

bone from periosteum gives rise to the onionskin histologic appearance (lamellar bone) (Figure 8–13). Exostoses are not traditionally considered neoplastic. Large exostoses can cause retention of cerumen, recurrent inflammation of the canal skin, and even conductive hearing loss. Exostoses are common in patients who have histories of swimming or bathing frequently in cold water. They are usually bilateral and appear as two or three smooth sessile protrusions on opposing surfaces of the bony canal medial to the isthmus.

An osteoma occurs as a single larger cancellous (trabecular) bone formation near the lateral end of the bony canal. It is a benign tumor of bone and is usually pedunculated. It invariably occurs unilaterally and may resemble a foreign body or cyst.

These bony growths are benign and may be ignored unless of sufficient size to obstruct the canal or cause repeated infection from retention of debris. Canalplasty is frequently used for the removal of a large exostosis or osteoma.[15] Skin over the lesion is elevated and preserved. After the lesion is exposed, it is drilled out with a cutting or diamond bur using continuous suction and irrigation. After all of the lesions are removed and the rest of the canal wall is smoothed, the skin flap is laid back. The EAC is packed with Gelfoam saturated with an antibiotic solution or Owen's silk strips and pieces of cotton packing (rosebud packing). In some instances, an exostosis interferes with middle ear surgery, in which case the canalplasty becomes a part of the tympanoplasty, with the skin flap becoming a tympanomeatal flap.

Fibrous Dysplasia The temporal bone may be the site of monostotic or, less commonly, polyostotic fibrous dysplasia. The mastoid and petrous portions of temporal bone are most often affected. In comparison with other craniofacial sites, the temporal bone is less commonly involved and is often asymptomatic. The most common presentation of temporal bone fibrous dysplasia is a progressive conductive hearing loss secondary to occlusion of the EAC or impingement on the ossicular chain. Progressive narrowing of the ear canal may lead to the development of a cholesteatoma. Diagnosis of fibrous dysplasia is based on radiographic and histologic findings. On plain radiographs, a fibrous dysplastic lesion consists of varying degrees of radiolucency and sclerosis, giving a "ground-glass" appearance. Surgery is indicated for diagnostic biopsy, presence of cholesteatoma, hearing loss, and correction of significant cosmetic deformity.[30]

Adenoma Adenomas of various sorts occur in the EAC. These are derived from sweat glands, sebaceous glands, or aberrant salivary gland tissue. Differentiation of these tumors depends on microscopic examination because each appears as a smooth, skin-covered, polypoid mass arising from the canal wall. Symptoms are minimal unless the growth completely occludes the canal. Pain suggests malignancy.

Ceruminous glands are modified apocrine sweat glands emptying into hair follicles. They are present only in the dermis of the cartilaginous portion of the EAC.[31] A *ceruminoma* is an adenoma of sweat gland origin. Because it is difficult to distinguish clinically a benign ceruminous gland tumor from malignant counterparts, biopsy is essential. There remains controversy regarding the behavior of ceruminous gland tumors.[31]

The treatment is surgical excision. Some of these lesions occasionally become malignant, and

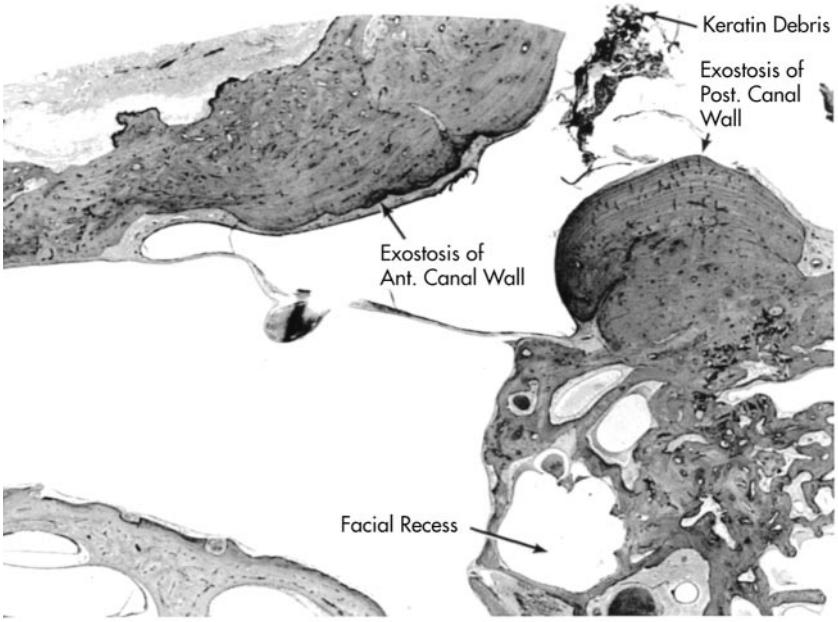


FIGURE 8–13. Photomicrograph showing occult (asymptomatic) exostoses of the anterior and posterior walls of the external auditory canal (male, age 75 years). Reproduced with permission from Schuknecht HF, Gulya AJ. Anatomy of the temporal bone with surgical implications. Philadelphia: Lea & Febiger; 1986.

pathologic examination is required in each case. Most of these adenomas may be removed through a transmeatal approach, although with large growths, an endaural incision may be needed. Any mass occurring within the EAC should be removed for histologic examination.[32] Lesions thought to be polyps, granulomas, or other benign tumors have proved to be malignant on occasion, usually squamous cell carcinoma. Because these tumors are often associated with chronic discharge from the ear, the patient with such discharge, especially associated with pain and bleeding, should be considered to have a malignant tumor.

Malignant Tumors of the External Ear

The most common malignant tumors of the external ear are squamous cell and basal cell carcinomas. Other malignant tumors include melanomas, adenoid cystic carcinomas, adenocarcinomas, and sarcomas.

Squamous Cell Carcinoma of the Auricle Cancer of the auricle accounts for about 6% of all skin cancers. Prolonged exposure to the sun is a risk factor for the development of skin malignancy. The degree of pigmentation plays a major role. Malignancy of the auricle is common in fair-skinned whites and rare in blacks. Squamous cell carcinoma represents about 55% of all malignant tumors of the auricle.[33]

Squamous cell carcinomas present as ulcerated lesions with an area of surrounding erythema and induration (Figure 8–14). They occur more often on the helix. The treatment of choice for squamous cell

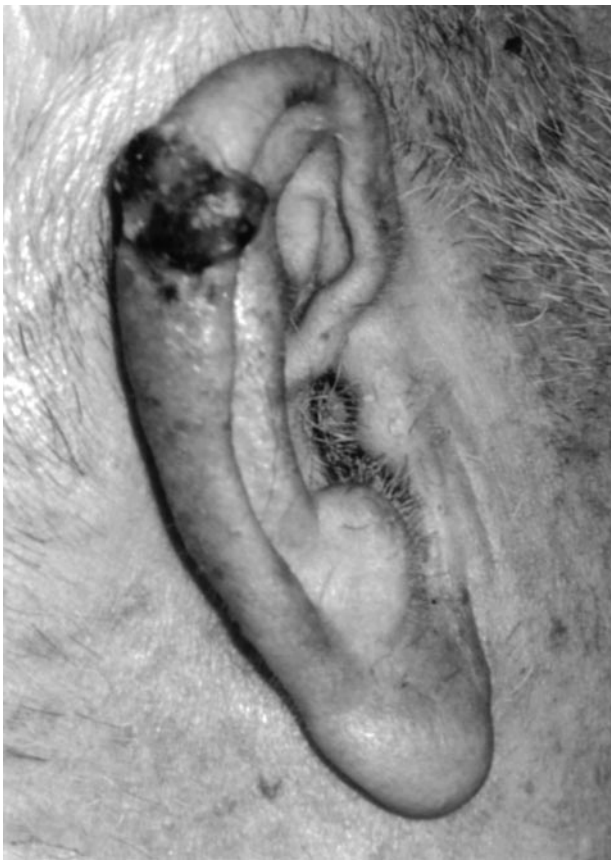


FIGURE 8–14. Squamous cell carcinoma of the helix.

carcinoma of the auricle is wide excision and reconstruction (Figure 8–15). Frozen sections to check the margins of excision should be done at the time of resection. A full-thickness skin graft from the neck to the defect is an excellent method of reconstruction, especially for the central area farther away from the helix.

Basal Cell Carcinoma of the Auricle Basal cell carcinoma presents as painless, well-circumscribed ulcers with raised margins. The tumors are often found in the preauricular or postauricular areas as well as on the helix and anterior surface of the auricle. Treatment is excision, as described above. The surgical margin need not be as large as that for squamous cell carcinoma.

Squamous Cell Carcinoma of the External Auditory Canal Malignant tumors can arise from the EAC, middle ear, and temporal bone. The tumor type, clinical presentation, prognosis, and treatment are determined by the anatomic location of the tumor. When patients present with advanced neoplasms, it may not be possible to define the sites of origin. The majority of cancers of the middle ear and mastoid originate in the EAC. Squamous cell carcinoma is the most common malignant tumor found in the EAC and middle ear space.

Carcinoma arising from the EAC often mimics chronic suppurative ear disease. Common symptoms include pain, otorrhea, bleeding, fullness, and decreased hearing. The most common finding is a growth in the EAC, which is clinically difficult to differentiate from an aural polyp, granulation tissue, and otitis externa. When a tumor is found in the canal, the location, size, and extent of the tumor should be thoroughly evaluated under an operating microscope. Malignancy should be suspected whenever otalgia or bleeding is associated with the tumor and a biopsy should be taken with local anesthesia using a cupped biopsy forceps. A difficult diagnostic challenge is the differentiation of carcinoma from malignant external otitis. The presentation of malignant external otitis may be indistinguishable from that of carcinoma of the EAC.

Either CT or MRI is helpful in defining the extent of a carcinoma. Computed tomography is unsurpassed for assessing the integrity of osseous structures of the temporal bone and middle ear. Magnetic resonance imaging is superior to CT in the evaluation of soft tissues and may be used to assess dural, intracranial, and extracranial soft tissue involvement.

Treatment consists of wide surgical excision and postoperative radiation therapy. Anterior canal tumors may spread through Santorini's fissures into a preauricular lymph node.[34] If the tumor has extended anteriorly and medially, the excision should include the cartilaginous and bony anterior canal wall, the superficial lobe of the parotid gland, and possibly the condyle of the mandible. When a malignant tumor is small, involving the posterior part of the EAC, and has not extended to the drumhead, a complete modified radical mastoidectomy is

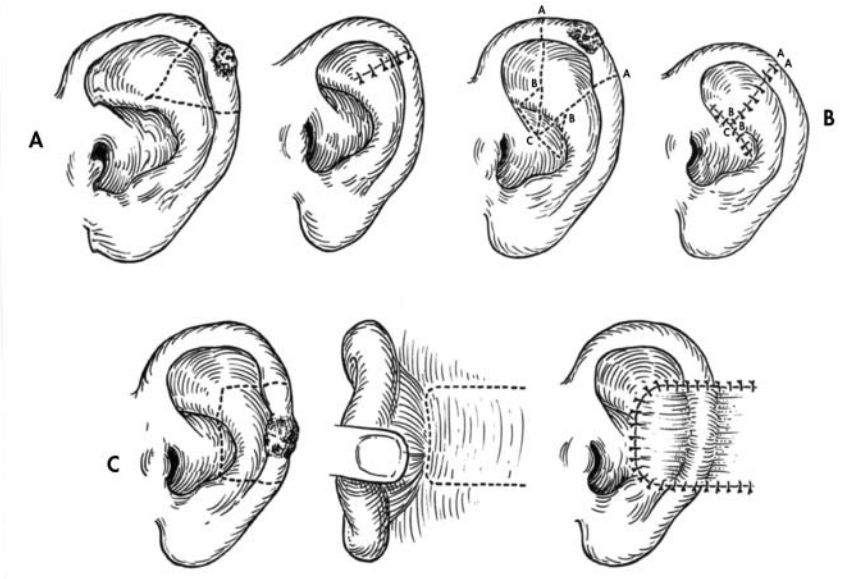


FIGURE 8–15. Surgical techniques for removal of small carcinomas of the auricle. *A*, Wedge excision; *B*, stellate excision; *C*, advancement flap. Reproduced with permission from Jenkins HA, Alford BR. Neoplasms of the auricle. In: Cummings CW, Fredrickson JM, Harker LA, et al, editors. Otolaryngology—head and neck surgery. Vol 4. St. Louis: Mosby Co.; 1986. p. 2927.

done. If the tumor is close to the drumhead, a radical mastoidectomy is performed. If a malignant tumor is extensive and involves the middle ear and mastoid, subtotal or total temporal bone resection may be needed.

Malignant Melanoma Melanomas of the external ear are rare. Melanoma may occur either on the auricle or in the EAC, more commonly in the former location. Diagnosis should be suspected when a pigmented lesion begins to increase in size or change in color. Lymphatic spread, treatment, and prognosis are greatly influenced by location. Tumors on the helical rim carry a better prognosis than those in the central areas, tragus or retroauricular regions. Treatment is wide excision.

ACKNOWLEDGMENT

The material presented in this chapter is based on the following: Austin DF. Diseases of the external ear. In: Ballenger JJ, Snow JB, editors. Otorhinolaryngology head and neck surgery. 15th ed. Philadelphia: Williams & Wilkins; 1996. p. 974–88.

REFERENCES

1. Jahn AF. Traumatic disorders of the external ear. In: Ludman H, Wright T, editors. Diseases of the ear. 6th ed. London: Arnold; 1998. p. 299–305.
2. Heggers JP, Robson MC, Manavalen K, et al. Experimental and clinical observations on frostbite. Ann Emerg Med 1987;16:1056–62.
3. Baers HA. Otologic aspects of ear burns. Am J Otol 1981;2:235–42.
4. Driscoll PV, Ramachandrula A, Drezner DA, et al. Characteristics of cerumen in diabetic patients: a key to understanding malignant external otitis? Otolaryngol Head Neck Surg 1993;109:676–9.
5. Bojrab DI, Bruderly T, Abdulrazzak Y. Otitis externa. Otolaryngol Clin North Am 1996;29:761–82.
6. Foidart JM, Abe S, Martin GR, et al. Antibodies to type II collagen in relapsing polychondritis. N Engl J Med 1978;299:1203–7.
7. McAdam LP, O'Hanlan MA, Bluestone R. Relapsing polychondritis, prospective study of 23 patients and a review of the literature. Medicine 1976;55:193–215.
8. Barranco VP, Minor DB, Soloman H. Treatment of relapsing polychondritis with dapsone. Arch Dermatol 1976;112:1286–8.
9. Senturia BH. Etiology of external otitis. Laryngoscope 1945;55:227–93.
10. Senturia BH, Marcus MD, Lucente FE. Diseases of the external ear: an otologic-dermatologic manual. 2nd ed. New York: Grune & Stratton; 1980.
11. Brownlee RE, Hulka GF, Prazma J, Pillsbury HC. Ciprofloxacin. Use as a topical otic preparation. Arch Otolaryngol Head Neck Surg 1992;118:392–6.
12. Jones RN, Milazzo J, Seidlin M. Ofloxacin otic solution for treatment of otitis externa in children and adults. Arch Otolaryngol Head Neck Surg 1997;123:1193–200.
13. Simpson KL, Markham A. Ofloxacin otic solution. A review of its use in the management of ear infections. Drugs 1999;58:509–31.
14. Paparella MM, Meyehoff WL, Morris MS, DaCosta S. Surgery of the external ear. In: Paparella MM, Shumrick DA, editors. Otolaryngology. 3rd ed. Philadelphia: WB Saunders; 1991 p. 1259–70.
15. Jung TT. Canalplasty. Oper Tech Otolaryngol Head Neck Surg 1996;7:27–33.
16. Chandler JR. Malignant external otitis. Laryngoscope 1968;78:1257–94.
17. Ress BD, Luntz M, Telischi FF, et al. Necrotizing external otitis in patients with AIDS. Laryngoscope 1997;107:456–60.
18. Gordon G, Giddings NA. Invasive otitis externa due to *Aspergillus* species: case report and review. Clin Infect Dis 1994;19:866–70.
19. Grandis JR, Curtin HD, Yu VL. Necrotizing (malignant) external otitis: prospective comparison of CT and MR imaging in diagnosis and follow-up. Radiology 1995;196:499–504.
20. Parisier SC, Lucente FE, Som PM, et al. Nuclear scanning in necrotizing progressive "malignant" external otitis. Laryngoscope 1982;92:1016–9.
21. Slattery WH, Brackmann DE. Skull base osteomyelitis. Malignant external otitis. Otolaryngol Clin North Am 1996;29:795–806.
22. Stokkel MP, Takes RP, van Eck-Smit BL, Baatenburg de Jong RJ. The value of quantitative gallium-67 single-photon emission tomography in the clinical management of malignant external otitis. Eur J Nucl Med 1997;24:1429–32.
23. Stokkel MP, Boot CN, van Eck-Smit BL. SPECT gallium scintigraphy in malignant external otitis: initial staging and follow-up. Case reports. Laryngoscope 1996;106:338–40.
24. Levenson MJ, Parisier SC, Dolitsky J, Bindra G. Ciprofloxacin: drug of choice in the treatment of

malignant external otitis (MEO). Laryngoscope 1991;101:821–4.
25. Sade J, Lang R, Goshen S, Kitzes-Cohen R. Ciprofloxacin treatment of malignant external otitis. Am J Med 1989;87:138S–41S.
26. Gehanno P. Ciprofloxacin in the treatment of malignant external otitis. Chemotherapy 1994;40 Suppl 1:35–40.
27. Lucente FE. Fungal infections of the external ear. Otolaryngol Clin North Am 1993;26:995–1006.
28. Piepergerdes MC, Kramer BM, Behnke EE. Keratosis obturans and external auditory canal cholesteatoma. Laryngoscope 1980;90:383–91.
29. Pierre G, Jean-Christophe D, Monique D. External auditory canal cholesteatoma. Arch Otolaryngol Head Neck Surg 1997;123:62–5.
30. Tran LP, Grundfast KM, Selesnick SH. Benign lesions of the external auditory canal. Otolaryngol Clin North Am 1996;29:807–25.
31. Headington JT. Epidermal carcinomas of the integument of the nose and ear. In: Batsakis JG, editor. Tumors of the head and neck. 2nd ed. Baltimore: William & Wilkins; 1979. p. 420–30.
32. Friedman I. Pathologic lesions of the external auditory meatus. J R Soc Med 1990;83:34–7.
33. Shockley WW, Stucker FJ. Squamous cell carcinoma of the external ear: a review of 75 cases. Otolaryngol Head Neck Surg 1987;97:308–12.
34. Kuhel WI, Hume CR, Selesnick SH. Cancer of the external auditory canal and temporal bone. Otolaryngol Clin North Am 1996;29:827–52.

CHAPTER 9

Otitis Media and Middle Ear Effusions

Gerald B. Healy, MD, Kristina W. Rosbe, MD

Infections of the middle ear space and their sequelae have plagued mankind from the beginning of time. First described by Hippocrates in 450 BC, this universally observed process continues to present one of the most perplexing medical problems of infancy and childhood, while being the leading cause of hearing loss in this age group. It is estimated that 70% of children will have had one or more episodes of otitis media (OM) by their third birthday.[1] This disease process knows no age boundaries but occurs mainly in children from the newborn period through approximately age 7 years, when the incidence begins to decrease. It occurs equally in males and females. A racial prevalence exists, with a higher incidence occurring in specific groups such as Native Americans, Alaskan and Canadian Natives, and Australian aboriginal children. African American children appear to have less disease than do American white children, but this observation has yet to be adequately explained.

Important epidemiologic factors include a higher incidence of OM in children attending daycare centers, a seasonal variation with more disease being present in the fall and winter versus the spring and summer, and a genetic predisposition to middle ear infection. Other epidemiologic factors include a lower incidence and duration of OM in breast-fed children and a higher incidence in children with altered host defenses.[2] Anatomic changes such as cleft palate and other craniofacial anomalies as well as congenital and acquired immune deficiencies are also important factors. One of the earliest signs of acquired immune deficiency syndrome (AIDS) in infants is recurring episodes of OM. This has been observed in more than 50% of neonates with AIDS.[3]

DEFINITIONS

Unfortunately, numerous terms have been used to describe the various inflammatory conditions of the middle ear space (Table 9–1). This has resulted in both confusion and a lack of uniformity in reporting. An attempt to standardize nomenclature was undertaken in the early 1980s and will be used throughout this chapter when discussing this disease process.[4]

Otitis media represents an inflammatory condition of the middle ear space, without reference to cause or pathogenesis.

Middle ear effusion is the liquid resulting from OM. An effusion may be either serous (thin, watery), mucoid (viscid, thick), or purulent (pus). The process may be acute (0 to 3 weeks in duration), subacute (3 to 12 weeks in duration), or chronic (greater than 12 weeks in duration).

Otitis media may occur with or without effusion. In those cases without effusion, inflammation of the middle ear mucous membrane and tympanic membrane may be the only physical finding. Occa-

TABLE 9–1. Synonyms Used in the Past for Otitis Media

Acute Otitis Media	*Otitis Media with Effusion*
Suppurative	Serous
Purulent	Secretory
Bacterial	Mucoid
	Glue ear
	Middle ear effusion

sionally, infection may involve only the tympanic membrane (myringitis), without involving the mucosa of the middle ear space.

This chapter deals primarily with two disease processes: acute otitis media (AOM) and chronic otitis media with effusion (COME). Acute otitis media represents the rapid onset of an inflammatory process of the middle ear space associated with one or more symptoms or local or systemic signs. These usually include otalgia, fever, irritability, anorexia, vomiting, diarrhea, or otorrhea. Physical examination usually reveals a thickened, erythematous or bulging tympanic membrane with limited or no mobility to pneumatic otoscopy. Erythema of the tympanic membrane may be an inconsistent finding and may be absent in certain systemic illnesses such as immune deficiency, when the patient cannot mount a sufficient inflammatory response to present this more classic finding. The acute onset of fever, otalgia, and, on occasion, a purulent discharge is usually evidence of AOM. Following such an episode, the patient may move into a subacute or even chronic phase in which fluid is present in the middle ear space, although active infection may be absent.

Chronic otitis media with effusion indicates the presence of asymptomatic middle ear fluid, usually resulting in conductive hearing loss. The tympanic membrane may present numerous physical findings including thickening, opacification, and impaired mobility. An air-fluid level and/or bubbles may be observed through a translucent tympanic membrane. This entity is distinguished from AOM in that the signs and symptoms of acute infection are lacking (eg, otalgia, fever, otorrhea).

ETIOLOGY

Eustachian tube (ET) dysfunction is considered the major etiologic factor in the development of middle ear disease. Politzer first proposed the ex vacuo theory of OM in 1867.[5] The theory postulates that chronic negative pressure, secondary to ET malfunction, results in the development of a transudate into the middle ear space. Numerous experiments have been carried out by many authors to substantiate this theory. It is traditionally maintained that the effusion is sterile; therefore, therapy should be aimed chiefly at relieving ET dysfunction. Most proponents of the theory have proposed that there are essentially two types of ET obstruction resulting in middle ear effusion: mechanical and functional. Mechanical obstruction may be either intrinsic or extrinsic. Intrinsic mechanical obstruction is usually caused by inflammation of the mucous membrane of the ET or an allergic diathesis causing edema of the tubal mucosa. Extrinsic mechanical obstruction is caused by obstructing masses such as adenoid tissue or nasopharyngeal neoplasms.

Some observers believe that infants and younger children may suffer from *functional ET obstruction* as a result of either decreased tubal stiffness or an inefficient active opening mechanism. Proponents believe that either form of obstruction results in inadequate ventilation of the middle ear with resulting negative middle ear pressure. This theory supports the development of a rational medical or surgical approach to alleviate the obstruction and thus overcome negative pressure.

The ET has three functions (1): ventilation of the middle ear associated with equalization of air pressure in the middle ear with atmospheric pressure, (2) protection of the middle ear from sound and secretions, and (3) drainage of middle ear secretions into the nasopharynx with the assistance of the mucociliary system of the ET and middle ear mucous membrane.

The second etiologic theory was first suggested by Brieger in 1914 and proposes an inflammatory origin to OM.[6] Since that time, several other authors have supported this theory. In 1958, Senturia reported a 41% incidence of positive bacterial cultures in 130 specimens taken from patients with a diagnosis of "serous otitis media."[6] Other series have supported this finding.[7] Sade's observation that the basic histopathologic mechanism in otitis media with effusion (OME) is an inflammatory hypertrophy of the middle ear mucous membrane and hyperplasia of its mucous glands also tends to support an inflammatory basis.[8]

Protein analysis of middle ear effusions indicates a higher concentration of total protein, lactate dehydrogenase, malate dehydrogenase, and acid phosphatase than in serum. This finding has led to the speculation that this material represents an exudate rather than a transudate, giving further evidence that this is an inflammatory process. Some proponents of the inflammatory theory feel, how-

ever, that inflammation occurs secondary to ET dysfunction; thus, the inflammatory response is not the primary etiologic factor.

Numerous other factors may well contribute to the development of middle ear disease. These include allergy, ciliary dysfunction, nasal and/or sinus disease, and immaturity of the immune system.

In the last 10 years, the role of allergy in OME has been extensively investigated.[9,10] Otitis media in the pediatric population is felt to be associated with allergy in 5 to 80% of cases. Inhalant allergies are felt to play a greater role than food allergies. Most studies have been unable to demonstrate, however, an increase in serum immunoglobulin E (IgE) in children with OME.[11]

Allergic rhinitis, itself, is not felt to be the cause of OME. A viral or bacterial infection may prime the environment, which, in response, produces inflammatory mediators. These mediators begin the physiologic cycle, creating ET dysfunction, pressure gradients, and transudation of fluid.[12]

Middle ear mucosa is rarely the target organ. The nasopharynx is felt to be the major site of action with subsequent spread via the ET to the middle ear. Allergy is felt to affect ET function in several ways. Nasal obstruction can occur secondary to mast cell degranulation with increased vascular permeability, increased mucosal blood flow, and increased mucus production. Retrograde extension of inflammatory mediators from the anterior nose to the nasopharynx as well as allergen contact with the nasopharynx can cause ET edema and obstruction with a secondary increase in the pressure gradient through nitrogen gas exchange and subsequent transudation of fluid.[12]

Animal models have demonstrated immune-mediated inflammation as a contributing factor in the pathogenesis of OME. The role of immune complexes in experimental models has also been established. Normal middle ear mucosa is not felt to be immunocompetent tissue. During inflammation, however, it is transformed into a secretory epithelium containing multiple goblet cells with infiltration of macrophages and lymphocytes. This immune response may result from circulating antibodies that enter the middle ear through increased vascular permeability.

Otitis media–prone children have been shown to have higher levels of IgG antibody and IgG-antigen immune complexes in their serum and middle ears versus a non–OM-prone cohort. Immunoglobulin may be the predominant immune mechanism in the middle ear. It is felt that bacteria may actually cause immunosuppression of cell-mediated immunity. Immunoglobulin A is thought to be a late defense mechanism and may actually prevent bacteriolysis by IgG and complement, acting as a blocking antibody. Others have proposed an IgE-mediated hypersensitivity reaction to viral antigens.

"Immune tolerance" is the concept that immunization through the oral or pulmonary route may modulate the middle ear immune response. Studies have shown that oral immunization after systemic sensitization actually increased immune-mediated OME. Vaccines are now becoming available, including the recently US Food and Drug Administration (FDA)-approved pneumococcal conjugate vaccine.[13]

MICROBIOLOGY

Numerous studies have documented the microbiology of OM.[14,15] Approximately 30% of middle ear effusions demonstrate known pathogens for OM.[16] Although the treatment of AOM is directed toward the elimination of the bacteria from the middle ear space, viruses may also play an important etiologic role in this disease process.[17] The most common bacterial pathogens responsible for acute infection include *Streptococcus pneumoniae* and nontypable *Haemophilis influenzae*. These two microorganisms account for approximately 60% of the cases associated with bacterial infection. Group A *Streptococcus, Branhamella catarrhalis, Staphylococcus aureus*, and gram-negative enteric bacteria are less frequent causes of OM (Figure 9–1).

Because of the difficulty in obtaining viral cultures, fewer specific data are available regarding their occurrence in patients with OM. However, *respiratory syncytial virus* accounts for a majority of the viral infections of the middle ear space.[17] Otitis media may accompany exanthematous viral infections such as infectious mononucleosis and measles.

Over the years, chronic effusions have been thought to be sterile. However, more recent studies have confirmed the presence of bacteria in middle ear fluid, and studies show that the bacterial spectrum closely resembles that found in AOM.[18] This

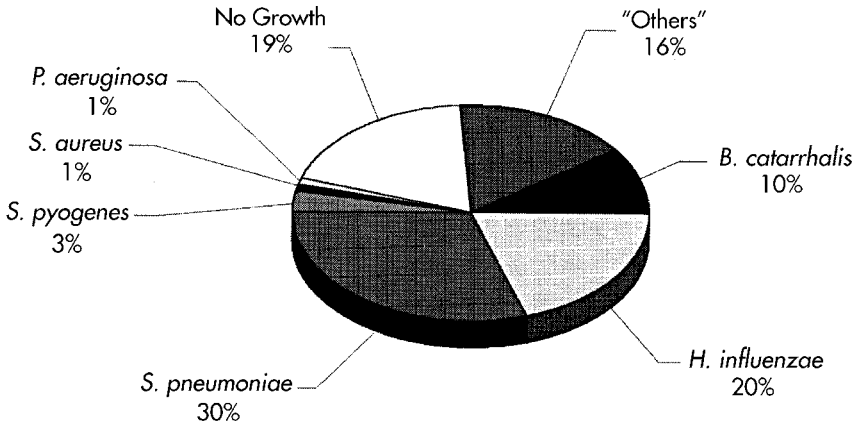


FIGURE 9–1. Bacterial incidence in acute otitis media with effusion.

information becomes increasingly important when consideration is given to the treatment of patients with both of these disease processes (Figure 9–2).

DIAGNOSIS

In most cases, a careful history and physical examination will lead to the accurate diagnosis of OM.

A careful history should elicit classic symptoms of OM. In the patient with the acute form of the disease, otalgia, fever, irritability, vomiting, and diarrhea may be present. Less frequently, otorrhea, vertigo, and facial paralysis may be associated with an acute infection of the middle ear space. In those patients in whom infection has spread into the mastoid air cell system and beyond, swelling of the postauricular area may be present.

In COME, hearing loss may be the only symptom. The most definitive part of the diagnosis is an appropriate physical examination to confirm the presence or absence of middle ear pathology. A complete examination of the head and neck should be undertaken first to identify the possibility of any predisposing condition such as craniofacial anomaly, nasal obstruction, palatal defect, or adenoid hypertrophy. In patients with unilateral OM, the nasopharynx should be visualized to rule out the possibility of neoplasm.

Otoscopy represents the most critical part of the examination to establish the diagnosis of OM (Figure 9–3). Use of the pneumatic otoscope is essential. The existence of chronic middle ear effusion is most easily confirmed when there is a definite air-fluid level or when bubbles are clearly visible within the middle ear space (Figure 9–4). However, findings commonly associated with OME include a severely retracted tympanic membrane with apparent foreshortening of the handle of the malleus and a reduction in tympanic membrane mobility (Figure 9–5). Occasionally, the tympanic membrane may be dull or thickened and have an amber hue (Figure 9–6). In severe cases, middle ear fluid may become purplish or blue, indicating hemorrhage within the tympanic cavity.

The color of the tympanic membrane is important but is not conclusive in making a diagnosis. An erythematous tympanic membrane alone

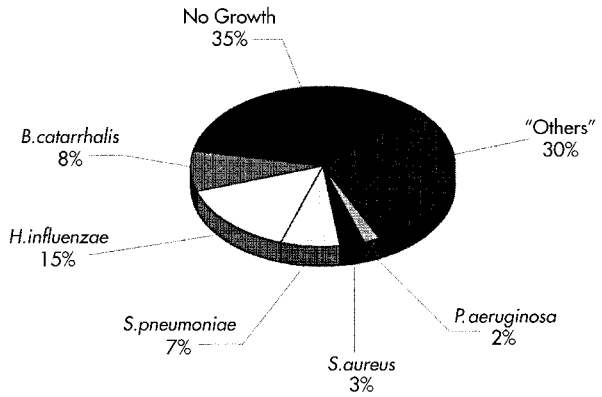


FIGURE 9–2. Bacterial incidence in chronic otitis media with effusion.

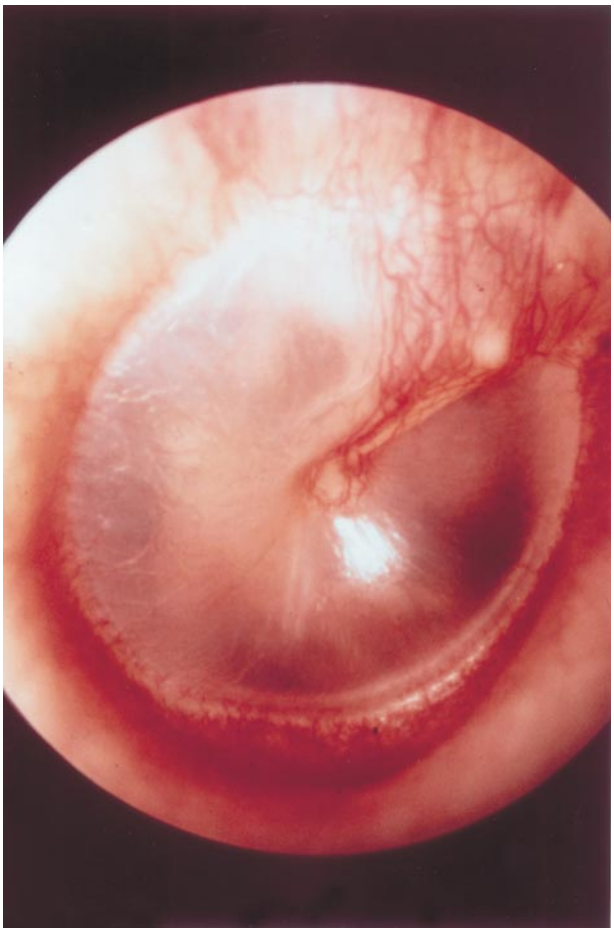


FIGURE 9–3. Normal right tympanic membrane at otoscopy.

may not be indicative of a pathologic condition because the vasculature of the tympanic membrane may be engorged as a result of the patient's crying or the presence of fever.

Acute otitis media usually presents with a hyperemic tympanic membrane that is frequently bulging and has poor mobility (Figure 9–7). Occasionally, perforation may be present, and purulent otorrhea is clearly visible.

The use of tympanometry has been popularized to confirm the findings of pneumatic otoscopy. This modality provides an objective assessment of the mobility of the tympanic membrane as well as the ossicular chain. By measuring tympanic membrane impedance, one can accurately predict conditions of the middle ear space (see Chapter 5).

The ultimate diagnostic test to confirm the presence of OM involves aspiration of middle ear contents. In acute situations, myringotomy or tympanocentesis may be undertaken to confirm the diagnosis, obtain material for culture, and relieve pus under pressure in an effort to avoid further complications. This may be necessary in patients who are unusually ill or toxic secondary to OM or in patients with severe suppurative complications. It may also be necessary in those patients who are having an unsatisfactory response to antibiotic therapy, in toxic newborns, or in patients who are significantly immune deficient.

With the rise in bacteria resistant to initial antibiotic therapy, tympanocentesis has been examined for its role in more specific antibiotic therapy.[19] Bluestone recommends tympanocentesis for several indications: (1) OM in patients with severe otalgia or toxic patients; (2) unsatisfactory response to antimicrobial therapy; (3) onset of OM in a patient already receiving antibiotics; (4) OM associated with a confirmed or potential suppurative complication; and (5) OM in a newborn, sick neonate, or immunosuppressed patient. This procedure does have significant risks including conductive and sensorineural hearing losses if not done properly. Otolaryngologists and pediatricians must be well trained in the technique to prevent these serious complications.

The hearing loss associated with OM should be documented whenever possible, especially in patients in whom chronic effusion is present. Although the presence of a conductive hearing loss does not confirm the diagnosis of COME, its presence does contribute to the confirmation of middle ear fluid. It is also important in documenting response to therapy.

MANAGEMENT

Acute Otitis Media

Acute otitis media represents one point in a continuum of the disease process known as "otitis media." The current standard of care strongly indicates that patients diagnosed as having an acute middle ear process should receive antimicrobial therapy for at least 10 to 14 days. In light of the fact that culture material is usually not readily available, therapy is begun on an empiric basis with treatment being aimed at the more common microorganisms found

FIGURE 9–4. Otitis media with effusion. Note bubbles confirming the presence of fluid.

in the acute process. Some have recommended withholding antimicrobial agents in certain cases.[20] However, in light of the fact that suppurative complications have markedly declined during the antibiotic era, antibiotic therapy is still strongly recommended in the acute process.[20]

The standard initial treatment for AOM is amoxicillin, 40 mg/kg every 24 hours in three divided doses, or ampicillin, 50 to 100 mg/kg every 24 hours in four divided doses for 10 days. In children allergic to penicillin, a combination of erythromycin, 40 mg/kg every 24 hours, and sulfisoxazole, 120 mg/kg every 24 hours in four divided doses, may be substituted. If β-lactamase-producing *H. influenzae* or *B. catarrhalis* is suspected, either amoxicillin-clavulanate, 40 mg/kg every 24 hours in three divided doses, or trimethoprim-sulfamethoxazole, 8 mg/kg of trimethoprim and 40 mg/kg of sul-

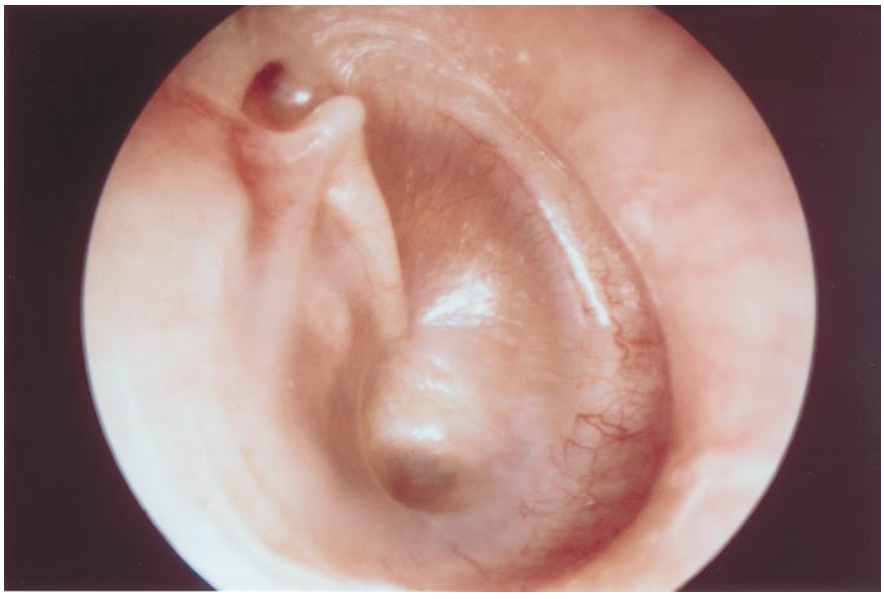


FIGURE 9–5. Advanced otitis media with effusion with markedly retracted right tympanic membrane, with an apparent foreshortened handle of the malleus and thick, mucoid effusion.

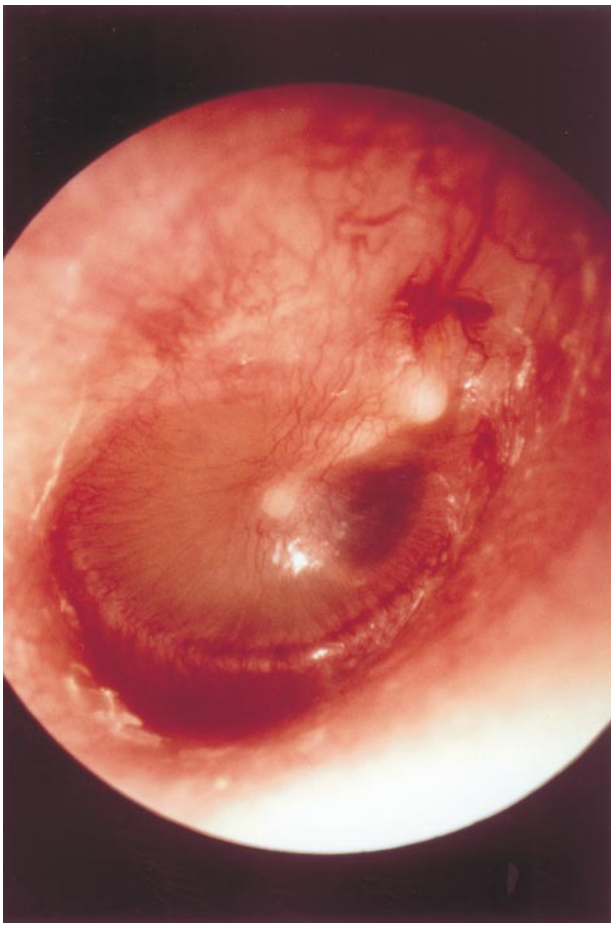


FIGURE 9–6. Chronic otitis media with effusion.

famethoxazole every 24 hours, may be used in two divided doses. Cefixime (Suprax, Lederle Laboratories, Wayne, NJ), 8 mg/kg in one dose, or cefprozil (Cefzil, Bristol-Myer-Squibb, Princeton, NJ), 15 mg/kg every 24 hours in two divided doses, may also be used effectively. A single intramuscular injection of ceftriaxone (Rocephin, Hoffman-Roche, Nutley, NJ), 50 to 100 mg/kg mixed with 1% lidocaine, not to exceed 1.5 mL, may be used in patients with vomiting (Table 9–2). Newer treatment for β-lactamase-producing microorganisms include high-dose amoxicillin at 80 mg/kg every 24 hours.[21] Dosages of appropriate antimicrobial agents for use in AOM are noted in Table 9–2.

Most patients who are receiving antibiotic therapy for AOM have significant improvement within 48 hours. If the child has not improved or the condition has worsened, tympanocentesis for culture and possibly myringotomy for drainage may be indicated. The patient may be re-examined some time during the course of therapy to ensure that the treatment has been effective.

Most children will have an effusion present at the completion of a 10- to 14-day course of antibiotic therapy. Such effusions may last up to 12 weeks before spontaneous clearance can be expected.[22]

Additional therapy such as analgesics, antipyretics, and oral decongestants (antihistamines and sympathomimetic amines) may be useful. Oral decongestants may relieve nasal congestion, providing some aeration of the ET. Their efficacy has not been proven, however.

The patient may remain in a subacute phase of the disease process for several weeks.

Repeated episodes of AOM plague many children, especially during the first 3 to 4 years of life. Several issues should be considered and evaluated in the management of such patients. A search should be made for a concomitant source of infection in the upper respiratory tract such as chronic adenoiditis or chronic sinusitis. Mild immune immaturity, especially in the IgG subclass group, may be responsible for this relentless process. Testing of immunoglobulins should be considered in relentless cases.

Such patients may be divided into two groups, the first being children who clear their effusion between episodes, whereas the second is made up of patients who have a persistent effusion between

TABLE 9–2. Daily Dosage for Common Antibiotic Agents in Acute Otitis Media

Agent	*Dosage/24 hr*
Amoxicillin	40 mg/kg in 3 doses
Ampicillin	50–100 mg/kg in 4 doses
Erythromycin-Sulfisoxazole	40 mg/kg (E) and 120 mg/kg (S) in 4 doses
Amoxicillin-clavulanate	40 mg/kg in 3 doses
Trimethoprim-sulfamethoxazole	8 mg (TMP) and 40 mg (SMZ)/kg in 2 doses
Cefixime	8 mg/kg in 1 dose

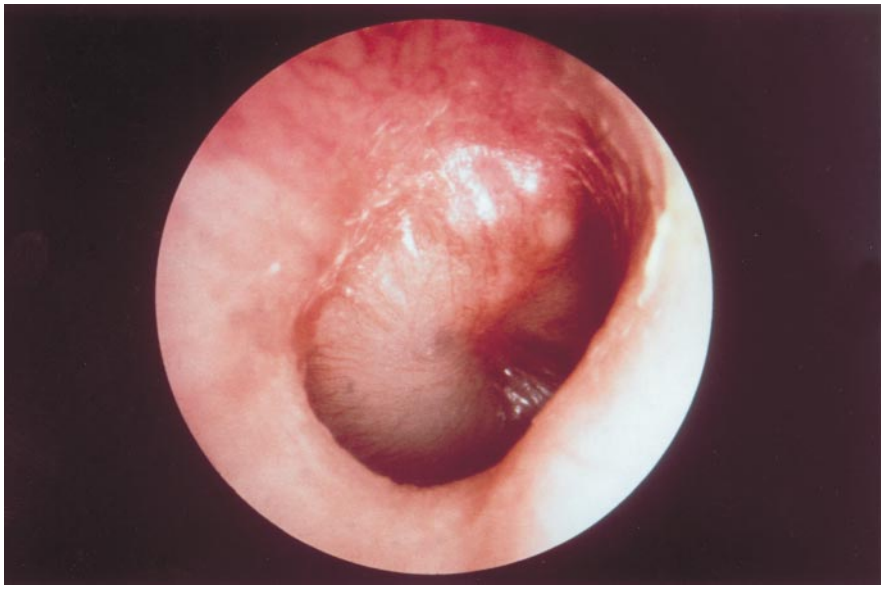


FIGURE 9–7. Acute otitis media showing an erythematous, bulging tympanic membrane.

episodes. In the latter group, chronic hearing loss becomes an additional major issue, especially as it impacts speech and language development.[1]

Several options are available for those patients who clear their effusion between episodes. The first option includes antibiotic therapy for each separate episode; a second option would be the use of antibiotic prophylaxis on a prolonged basis, whereas the third would include myringotomy and ventilation tube insertion. The administration of pneumococcal vaccine also may be useful.

Amoxicillin, or a suitable substitute in penicillin-allergic patients, may be given once daily at bedtime as prophylaxis. This form of therapy is usually administered during the months in which OM has its highest prevalence.

In children in whom the middle ear does not clear between acute episodes, ventilation tubes may be necessary to address the concomitant hearing loss associated with the process. Their use may also be necessary in those patients with antibiotic allergy or intolerance.

Chronic Otitis Media with Effusion

Medical Therapy Chronic otitis media with effusion may occur as a sequela to AOM or in patients who have had no documented recent episodes of acute suppurative disease. Numerous associated factors must be considered; thus, a careful history should be taken for the possibility of underlying allergy, sinus disease, or nasopharyngeal obstruction, which may be secondary to hypertrophic adenoids or even neoplasm.

Numerous methods of management have been advocated over the years for the persistent form of the disease. Antihistamine-sympathomimetic amine preparations were used frequently to clear the effusion. However, controlled clinical trials have demonstrated a lack of efficacy.[23] The use of corticosteroids, either applied topically in the nose or given systematically, has been reported to be advantageous in clearing middle ear fluid.[24] Unfortunately, there is a paucity of data to demonstrate efficacy; therefore, their use cannot be strongly recommended at this time.

The most effective medical therapy used to this point has been antibiotic administration. Numerous trials have concluded that some patients may respond to a 21- to 30-day course of full-dose antibiotic therapy.[25] The demonstration of viable bacteria in the middle ear effusions of chronically diseased ears has led to this recommendation. In light of the similarity of the bacterial spectrum, the same antibiotics recommended for AOM may be used in this disease. This form of therapy is strongly rec-

ommended in any child who has not received antibiotic treatment before consideration is given to myringotomy and ventilation tube insertion and/or adenoidectomy.

Historically, there was debate as to whether OME should be treated with antibiotics since there exists a certain spontaneous resolution rate. Physicians in other countries have traditionally not treated OM with antibiotics as frequently as is done in the United States. Rosenfeld and Post's meta-analysis of studies of antibiotic therapy for OME found a statistically significant efficacy for short-term resolution of OME.[16] Gates stated that approximately 40% of children with OME treated with antibiotics still have an effusion after 30 days, whereas 10% have a documented effusion 3 months after initial therapy.[26] Another factor that has not been studied well to date is what the impact of COME is on language development and learning in the child with conductive hearing loss secondary to COME and at what time point this deficit becomes an unacceptable risk. The success of prophylactic antibiotic therapy for OME had also not been studied in a systematic fashion.

In children with concomitant disease of the upper respiratory tract such as chronic sinusitis or adenoiditis, consideration must be given to the simultaneous control of these diseases. In addition, systemic problems such as allergy or immunodeficiency must also be addressed if long-term reversal of the middle ear abnormality is to be achieved.

With the emerging role of allergy in OME, the role of corticosteroid therapy has been re-examined. Investigators have demonstrated that a brief course of oral corticosteroids is efficacious for the short-term cure of OME. The proposed therapeutic mechanisms include stabilization of phospholipids to prevent arachidonic acid formation and subsequent inflammatory mediator formation, possible decrease in peritubal lymphoid tissue size, enhanced secretion of ET surfactant, and reduced middle ear fluid viscosity by action on mucoproteins.[27] The risks of oral corticosteroid therapy in children remain a concern, however. Studies have demonstrated an increase in the development of AOM if corticosteroids are not prescribed with antibiotics, transient depression in adrenal function, and the potential for disseminated varicella infection and its complications. Because of these risks, Rosenfeld recommends a trial of antibiotics and corticosteroids only if the next step would be tympanostomy tube placement. There have been no good randomized controlled trials to examine the risk-to-benefit ratio.[27] Relapses have been demonstrated within several weeks after treatment. The effects of nasal corticosteroids or immunotherapy on OME have not been studied.

Surgery The use of ventilation tubes with or without adenoidectomy has become the ultimate treatment of COME. This surgical intervention immediately corrects the conductive hearing loss associated with the middle ear process and diminishes the patient's tendency toward recurrent infection. It should be strongly considered in the following situations:

1. Recurrent AOM
   a. Unresponsive to antibiotic therapy
   b. Significant antibiotic allergy or intolerance
2. Negative middle ear pressure with impending cholesteatoma
3. Chronic effusion of the middle ear space with a duration of greater than 3 months
   a. Conductive hearing loss of greater than 15 dB
   b. Nasopharyngeal neoplasm for which treatment such as radiation therapy may be necessary

Although some controversy exists over the use of ventilation tubes, they do provide a safe method for normalizing middle ear pressure and, in most cases, restoring hearing to normal. They are usually associated with minimal morbidity, although tympanosclerosis or persistent perforation may be seen in a few instances. In most circumstances, it is difficult to determine whether these findings are a result of the underlying middle ear pathology with middle ear atelectasis or secondary to the ventilation tube itself. Numerous types of tubes are available for ventilation of the middle ear space, but, basically, they are all designed to provide equalization of pressure across the tympanic membrane (Figure 9–8). Some designs favor intubation of greater duration but also carry a slight risk of an increased incidence of persistent perforation upon extrusion.

Depending on the age of the patient, tube insertion may be carried out under local or general anesthesia.

It is usually advisable to allow spontaneous extrusion of the tubes to occur. Most tympanostomy tubes remain in the tympanic membrane for

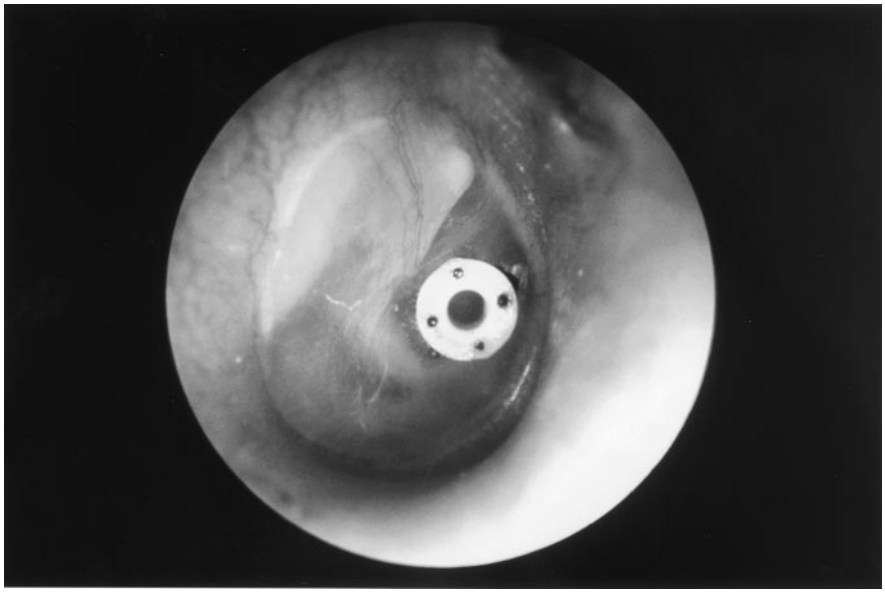


FIGURE 9–8. Ventilation tube in place in the right tympanic membrane.

approximately 6 to 12 months, with some extruding earlier and some later.

The most common complication of tube insertion is otorrhea. This may be secondary to either reflux of nasopharyngeal secretions, especially during an upper respiratory infection, or as a result of pathogens entering the middle ear through the lumen of the ventilation tube. It is commonly believed that contaminated water that is allowed to enter the middle ear through the tympanostomy tube may result in AOM media with otorrhea. In these instances, the external ear canal should be carefully cleaned and a culture obtained from the middle ear by aspirating through the tympanostomy tube. After this has been accomplished, oral antibiotics should be initiated as well as topical antibiotic therapy. The usual treatment includes the use of antibiotic agents commonly prescribed for other forms of AOM.

To avoid this possible complication, water protection of the ears is usually advised when contamination may be a possibility. This may be accomplished through the use of earplugs or cotton covered with petrolatum jelly, both of which can be inserted into the ear canal to provide adequate protection. In summary, tympanostomy tubes can provide a useful means of ventilating the middle ear space, controlling hearing loss secondary to middle ear effusion, and controlling recurrent infection.

Adenoidectomy may be useful as an adjunct to myringotomy in the treatment of middle ear effusion.[28] Removal of adenoid tissue improves the ventilatory function of the ET, thus allowing for appropriate equalization of pressure. The adenoid has been felt to play a role in OM in two ways: when hypertrophic, by causing mechanical obstruction of the ET, and when small, as a bacterial reservoir.[29] The adenoid is thought to be an important site for primary contact of inhaled microorganisms. Two studies are frequently cited regarding the use of adenoidectomy in OME. Gates demonstrated greater long-term efficacy in the treatment of OM in children 4 to 6 years of age when adenoidectomy was added to tympanostomy tube placement or myringotomies even if this was the first surgical intervention in a child.[26] Paradise et al, on the other hand, recommend adenoidectomy only if a child fails initial tympanostomy tube placement.[30] Studies have also shown that the recurrence rate of AOM may be reduced by adenoidectomy.[30] In addition, other confounding factors may warrant adenoidectomy such as evidence of chronic adenoiditis, nasal obstruction, or recurrent or chronic sinusitis secondary to nasal obstruction.

CONCLUSION

The future solution to OM, acute or chronic, does not lie in current therapy regimens. Manipulation of the immune system to provide enhanced protection will probably result in a significant decrease in the incidence of this disease process. Vaccination may play a significant role in this endeavor. Both nontypable *H. influenzae* and *S. pneumoniae* are leading etiologic factors in the development of bacterial OM and have polysaccharide capsules. Polysaccharide vaccines lack immunogenicity in infants and children under 2 years. Based on the same premise as the successful *H. influenzae* type b vaccine, in which the capsular polysaccharide is conjugated with a protein, a heptavalent pneumococcal conjugate vaccine, PCV7 (Prevanar, Wyeth-Lederle Vaccines), has become available. PCV7 and other pneumococcal vaccines may prove to be an important step in the prevention of AOM. This vaccine is recommended for universal use in children 23 months and younger. The American Academy of Pediatrics recommends that if immunization is initiated in infants younger than 7 months, four doses of PCV7 be given concurrently with other recommended childhood immunizations at 2, 4, 7, and 12 to 15 months. If immunization is initiated in infants who are between 7 and 23 months of age, who have not received their prior doses of PCV7, fewer doses are recommended. Children with congenital immunodeficiency, chronic cardiac disease, and infection with human immunodeficiency virus, chronic pulmonary disease, and other serious chronic conditions are especially vulnerable to pneumoccocal infection.[31] Attempts are under way to develop conjugated vaccines against nontypable *H. influenzae* and *B. catarralis*, which also has a polysaccharide capsule. It is too early to determine the impact of vaccination on the incidence of AOM. Other factors such as the role of gastroesophageal reflux or the role of such substances as surfactant have yet to be determined.

Currently, OM represents a costly medical problem worldwide. In some cultures, its morbidity with associated hearing loss, learning difficulties, and secondary central nervous system complications are almost unmeasurable. This significant medical problem deserves more research attention than it receives so that significant costs and morbidity associated with it can be reduced. Berman et al have developed a cost-effectiveness model for OME therapy.[32] Their recommendations include a trial of corticosteroids and antibiotic therapy at the 6-week follow-up visit if no resolution, another course of antibiotics at the 9-week visit, and tympanostomy tube placement if no resolution at 12 weeks. With OM continuing to be one of the most common reasons for physician visits in children today, finding the safest, most efficacious, and cost-effective treatment remains an important goal.

REFERENCES

1. Teele DW, Klein JO, Rosner B. Otitis media with effusion during the first three years of life and the development of speech and language. Pediatrics 1984;74:282–7.
2. Daly KA. Epidemiology of otitis media. Otolaryngol Clin North Am 1991;24:775–86.
3. Ammann AJ, Shannon K. Recognition of acquired immune deficiency syndrome (AIDS) in children. Pediatr Rev 1985;7:101–7.
4. Senturia BH, Bluestone CD, Klein JO, et al. Report of the Ad Hoc Committee on Definition and Classification of Otitis Media with Effusion. Ann Otol Rhinol Laryngol 1980;89:3–4.
5. Politzer A. Diagnose und Theraphaie der ansammIung seroeser Fluessigkeit in der Thrommelhoehle. Wien Med Wochenschr 1867;17:244–7.
6. Senturia BH. Studies concerned with tubo-tympanitis. Ann Otol Rhinol Laryngol 1958;67:440–67.
7. Healy GB, Smith HG. Current concepts in the management of otitis media with effusion. Am J Otol 1981;2:138–44.
8. Sade J. The biopathology of secretory otitis media. Ann Otol Rhinol Laryngol 1970;81:59–70.
9. Bernstein JM. Immunologic reactivity in the middle ear in otitis media with effusion. Otolaryngol Clin North Am 1991;24:845–58.
10. Bernstein JM. The role of IgE-mediated hypersensitivity in the development of otitis media with effusion: a review. Otolaryngol Head Neck Surg 1993; 109:611–20.
11. Bernstein JM, Doyle WJ. Role of IgE-mediated hypersensitivity in otitis media with effusion: pathophysiologic considerations. Ann Otol Rhinol Laryngol 1994;103:15–9.
12. Bernstein JM. Role of allergy in eustachian tube blockage and otitis media with effusion: a review. Otolaryngol Head Neck Surg 1996;114:562–8.

13. Lieu TA, Ray GT, Black SB, et al. Projected cost-effectiveness of pneumococcal conjugate vaccination of healthy infants and young children. JAMA 2000;283: 1460–8.
14. Bluestone CD, Klein JO. Microbiology in otitis media in infants and children. Philadelphia: WB Saunders; 1988.
15. Bluestone CD. Current therapy for otitis media and criteria for evaluation of new antimicrobial agents. Clin Infect Dis 1992;14 Suppl 2:S197–203.
16. Rosenfeld RM, Post JC. Meta-analysis of antibiotics for the treatment of otitis media with effusion. Otolaryngol Head Neck Surg 1992;106:378–86.
17. Mein JO, Teele DW. Isolation of viruses and mycoplasmas from middle ear effusions: A review. Ann Otol Rhinol Laryngol 1976;85 Suppl 25:140–4.
18. Healy GB, Teele DW. The microbiology of chronic middle ear effusions in young children. Laryngoscope 1977;87:1472–8.
19. Bluestone CD. Role of surgery for otitis media in the era of resistant bacteria. Pediatr Infect Dis J 1998;17: 1090–8.
20. VanBuchem FL, Peters MF, Van't Hof MA. Acute otitis media: a new treatment strategy. Par Med J (Clin) Res 1985;290:1033–7.
21. Jacobs MR. Drug-resistant *Streptococcus pneumoniae*: rational antibiotic choices. Am J Med 1999;106 Suppl 5A:19S–25S.
22. Teele DW, Mein JO, Rosner BA. Epidemiology of otitis media in children. Ann Otol Rhinol Laryngol 1980;89:5–6.
23. Cantekin EI, Mendel EM, Bluestone CD, et al. Lack of efficacy of a decongestant-antihistamine combination for otitis media with effusion ("secretory" otitis media) in children. N Engl J Med 1983; 308:297–301.
24. Schwartz RH, Puglese J, Schwartz DM. Use of a short course of prednisone for treating middle ear effusion: a double-blind cross-over study. Ann Otol Rhinol Laryngol 1980;89 Suppl 68:296–300.
25. Healy GB. Antimicrobial therapy of chronic otitis media with effusion. Int J Pediatr Otorhinolaryngol 1984;8:13–7.
26. Gates GA. Adenoidectomy for otitis media with effusion. Ann Otol Rhinol Laryngol 1994;103:54–8.
27. Rosenfeld RM. New concepts for steroid use in otitis media with effusion. Clin Pediatr 1992;31:615–21.
28. Gates GA, Avery CS, Phihooa TJ, et al. Effectiveness of adenoidectomy and tympanostomy tubes in the treatment of chronic otitis media with effusion. N Engl J Med 1987;317:1444–88.
29. Honjo I, Takahashi H, Sudo M, et al. Pathophysiological and therapeutic considerations of otitis media with effusion from viewpoint of middle ear ventilation. Int J Pediatr Otorhinolaryngol 1998;43:105–13.
30. Paradise JL, Bluestone CD, Rogers KD. Efficacy of adenoidectomy for recurrent otitis media: results from parallel random and non random trials. Pediatr Res 1987;21:286–91.
31. Black S, Shinefield H, Fireman B, et al. Efficacy, safety, and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group. Pediatr Infect Dis J 2000;19:187–95.
32. Berman S, Roark R, Luckey D. Theoretical cost effectiveness of management options for children with persisting middle ear effusions. Pediatrics 1994;93: 353–63.

CHAPTER 10

Chronic Otitis Media

Steven A. Telian, MD, Cecelia E. Schmalbach, MD

CLASSIFICATION

At first glance, chronic otitis media (COM) might appear to be an overwhelming and enigmatic topic. A great deal of the confusion stems from the lack of a universally accepted system of nomenclature for chronic middle ear disease and the related surgical procedures.

A unifying definition of the term "chronic otitis media" is any structural change in the middle ear system associated with a permanent defect in the tympanic membrane (TM). Usually, but not always, there is associated inflammatory mucosal disease in the middle ear, which may also involve the mastoid cells. If there is persistent or intermittent otorrhea through a nonintact TM, the amplified designation "chronic suppurative otitis media" is preferred. The condition is considered "chronic" if the TM defect is present for a period greater than 3 months. Thus, a draining middle ear that is associated with a perforation from acute otitis media would not qualify for this diagnosis if it responds to treatment within 3 months. Possible causes for the TM defect include infection, trauma, retraction pockets developing from chronic negative middle ear pressure, and therapeutic interventions such as ventilation tubes. Histologically, COM is defined as irreversible mucosal changes within the middle ear cleft.

Perforations of the Tympanic Membrane

Perforations of the TM are described according to their anatomic location and are separated into two categories.

Central perforations involve the pars tensa and are circumferentially surrounded by residual TM. The umbo of the malleus is used as a reference point to divide the pars tensa into four quadrants, allowing one to describe the location of the perforation. Subtotal perforations describe large defects in which there is only a narrow rim of residual pars tensa near the annulus. Central perforations are rarely associated with cholesteatoma and for this reason have generally been considered "safe" ears. However, serious complications in the setting of a central perforation, such as intracranial abscess, have been reported.[1] The term "tubotympanic disease" is sometimes used to describe COM with a central perforation. This term stems from the fact that the TM defect exposes both the middle ear mucosa and eustachian tube but generally does not produce inflammatory changes in the mastoid.

Marginal perforations have no remnant of TM adjacent to the bone of the posterior canal wall. As a result, the bony external canal wall, attic, antrum, and mastoid cells can be involved with inflammation. Hence, this condition has been referred to as atticoantral disease. Retraction pockets produce marginal perforations and also occur in the pars flaccida, where they are known as attic perforations. In addition, total perforations with complete loss of the pars tensa result from necrotizing otitis media.

Cholesteatomas

Cholesteatomas are retraction pockets or cysts lined with squamous cell epithelium and filled with keratin debris occurring within the pneumatized spaces of the temporal bone. They are frequently associated with marginal perforations. The name is a misnomer as cholesteatoma is not a true neoplasm and does not contain cholesterol, but this familiar term retains wide popularity in clinical parlance. Cholesteatomas have a propensity for growth, bone destruction, and chronic infection. Therefore, COM with cholesteatoma is considered an "unsafe" ear and generally requires surgical treatment.

Cholesteatomas are categorized as congenital or acquired. Congenital cholesteatomas appear as white pearly masses deep to normal, intact TMs. Prior history of otorrhea, perforation, or otologic procedures excludes this category. However, patients may have a prior history of uncomplicated acute otitis media.[2]

Acquired cholesteatomas are much more common and develop in the setting of a retracted or perforated TM. Primary acquired cholesteatomas arise from retracted but intact drumheads, most often within an attic. Secondary acquired cholesteatomas result from ingrowth of squamous cell epithelium into the middle ear, usually through marginal perforations.

Cholesteatomas may also be described according to their anatomic location. Pars tensa cholesteatomas usually involve the posterosuperior quadrant of the TM and commonly cause erosion of the long process of the incus with discontinuity of the incudostapedial joint and conductive hearing loss. The disease frequently extends into the posterior tympanic spaces, facial recess, and sinus tympani. Attic cholesteatomas arise from pars flaccida defects and tend to be associated with scutum defects owing to erosion of the outer wall of the epitympanum. The ossicles are frequently engulfed or eroded by the cholesteatoma, which may then extend into the mastoid antrum.

Atelectatic and Adhesive Otitis Media

Although not associated with an actual perforation of the TM, atelectatic otitis media is also considered a form of COM. This condition involves collapse and retraction of an atrophic TM into the middle ear cleft. As a result, the middle ear space becomes partially or completely obliterated, but the membrane is not adherent to the underlying mucosa. Adhesive otitis media refers to the condition in which the middle ear space is completely obliterated, and the TM is irreversibly adherent to the medial bony wall. In adhesive otitis media, the normal mucosal lining of the middle ear space is replaced with a scarred membrane comprising the mucosa and the remnant of the TM.

EPIDEMIOLOGY

Because of the variable expression and insidious nature of COM, epidemiologic studies are limited and rely heavily on retrospective research. It is recognized that specific racial groups such as Native Americans, Canadian and Alaskan Inuits, Australian Aborigines, and New Zealand Maoris are more prone to COM. The Canadian Inuits offer a unique population for the study of COM. Prior to the 1950s, ear disease was almost nonexistent in this population. However, after World War II, the Inuits began living in overcrowded, overheated, prefabricated homes with significant smoke exposure, substandard sanitation, and poor dietary habits. These conditions probably represent risk factors for COM.[3] This theory is further supported by the decrease in COM among Maori children between 1978 and 1987 when improvements were made in their socioeconomic status, living conditions, and access to medical care.[4] Studies of COM involving these distinct ethnic groups must be interpreted with caution since findings in these communities may not generalize to other populations. Chronic otitis media in these distinct ethnic groups is rarely associated with cholesteatoma, and the disease tends to present earlier, usually by the age of 2 years.

The natural history of COM has been investigated with a retrospective review of 200 Israeli patients with perforations of the TM.[5] Seventy-six percent of the patients were found to have perforations, whereas the remaining 24% had COM with cholesteatoma. Regardless of the presence of cholesteatoma, approximately one-third of the patients had middle ear disease of the contralateral ear. Surprisingly, an average duration of 10 to 14 years elapsed between the onset of symptoms and presentation to an otologist. During this period, patients were seen by a general practitioner an average of four times a year for the first 7 years of life and were subsequently treated for an additional 5 to 8 years by a general otolaryngologist. However, cultures, audiograms, and radiographic studies were rarely used. This delay in referral was thought to be related to the slow, sometimes relatively asymptomatic evolution of COM as well as some lack of awareness of therapeutic interventions that are available.

Studies specific to COM with cholesteatoma reveal an annual incidence ranging between 6 and 12 per 100,000.[6–8] The current literature suggests that 10% of cholesteatomas are bilateral.[7] In a 10-year retrospective study of 500 patients with cholesteatoma, approximately three-fourths of the patients suffered from suppurative otitis media on one or more occasions. In addition, 10% of the

patients underwent placement of one or more sets of pressure equalization tubes.

PATHOGENESIS

Chronic otitis media is an insidious process, and patients tend to present with long-standing disease. As a result, the etiology and natural course of this process remain obscure, although several credible theories have been advanced.

Continuum Theory

Traditionally, COM has been thought to follow a bout of acute otitis media (AOM) that resulted in TM perforation. However, this direct correlation has fallen out of favor for several reasons. First, AOM is one of the most common childhood diseases. Comparatively speaking, COM is quite rare. In addition, the majority of TM perforations secondary to AOM result in complete healing of the drumhead.[9] Second, whereas streptococcal otitis media, which causes necrotizing infections resulting in large perforations, is seldom seen today, the incidence of COM has remained constant.[10] Third, in a study of 200 patients with TM perforations, only 50% clearly recalled an acute, painful ear infection associated with the onset of otorrhea. Instead, 40% described the insidious onset of drainage or gradual hearing loss.[5]

Although COM may not result directly from a single episode of AOM, it has been suggested that all cases of otitis media represent different stages in a continuum of events.[11–13] For example, histologic studies have demonstrated that persistent effusion in chronic secretory otitis media leads to degradation of the fibrous layer of the TM. Loss of the fibrous layer results in a weakened, atrophic, two-layered drumhead that is vulnerable to atelectasis or perforation and hence chronic middle ear disease.[9,14]

Atelectatic and Adhesive Otitis Media

It is acknowledged that eustachian tube dysfunction plays an important role in the development of COM.[15] The eustachian tube serves to ventilate the middle ear so that pressure equalization occurs between this space and the surrounding environment. In persisting eustachian tube dysfunction, especially as seen in Down syndrome and cleft palate, the middle ear space is continually exposed to negative pressure. As a result, the TM is retracted medially. In atelectatic or adhesive OM, the middle ear space is partially or completely obliterated. In long-standing atelectasis, patients are at risk for secondary acquired cholesteatomas.

Poor mastoid pneumatization is also associated with chronic middle ear disease.[15] Although pneumatization is not completed until adulthood, the majority of the process takes place during the first 5 years of life. Infancy and early childhood infections occurring during this period are thought to prevent normal cellular development of the mastoid and thus lead to chronic middle ear disease. Temporal bone histopathologic studies also demonstrate that infection of a pneumatized cleft incites sclerosis, obliteration of air cells, and chronic middle ear disease in the setting of poor mastoid pneumatization.[9]

Ventilating tubes can be placed within an atelectatic TM in an attempt to equalize the pressure and allow the TM to return to its normal anatomic position. Fifty percent of patients with pressure equalization tubes experience at least one episode of otorrhea, and 3% will have symptoms persisting beyond 6 weeks.[16] After extrusion of the tube, the majority of iatrogenic TM perforations will heal. Residual perforation rates are 2 to 3% for button or grommet tubes but as high as 47% for T tubes.[17,18] Continued eustachian tube dysfunction, persistent otorrhea, and ingrowth of squamous cell epithelium through the defect can prevent spontaneous healing. In accordance with the continuum theory, patients are then at risk for secondary infections and COM.

Chronic Otitis Media without Cholesteatoma

Recurrent infections of the middle ear generally result in irreversible mucosal changes. Histologic studies have shown that as the inflammatory process enters the chronic phase, there is a shift in cellular population from infiltrating leukocytes toward mononuclear cells such as macrophages, lymphocytes, and plasma cells.[19] These mononuclear cells secrete inflammatory mediators and growth factors that increase capillary permeability and lead to edema and hyperemia of the middle ear mucosa. In chronic inflammation, the mucosa undergoes metaplasia from a single layer of ciliated cuboidal or columnar epithelium to mucosa resembling that of the respiratory tract with increased numbers of gob-

let and glandular cells. Consequently, there is an increase in the volume and viscosity of the mucus. These changes further overwhelm the already compromised mucociliary clearance capability of patients suffering from chronic middle ear disease.[19,20]

Granulation tissue consisting of vascular connective tissue with inflammatory infiltrates has been found to be the prominent pathologic feature of COM. Granulation tissue was identified in over 95% of the temporal bones studied from individuals with a history of COM.[21] Tympanosclerosis was present in 43%, cholesteatoma in 36%, and cholesterol granuloma in 21% of patients in one large histologic study of temporal bone pathology. A pathologic review of 800 temporal bones revealed that granulation tissue had both a higher prevalence and more generalized distribution when compared to cholesteatoma.[22] Both studies demonstrated identical pathologic changes within the middle ear cleft regardless of the presence of a TM perforation.

As granulation tissue matures, it becomes dense and fibrotic with decreased vascularity. This process leads to scarring and adhesions associated with the ossicular chain and TM.[23] Irreversible changes such as subepithelial edema and mucoperiosteal fibrosis occur deep to the epithelial lining.[24] As the inflammation persists, sclerosis, along with new bone formation, can cause a reduction in mastoid and antral pneumatization. Other late changes such as bone erosion, tympanosclerosis, and cholesterol granuloma are discussed in detail below.

Chronic Otitis Media with Cholesteatoma

Congenital cholesteatomas are thought to result from an error in embryogenesis that causes squamous epithelial cell arrest behind an intact TM. This theory has gained popularity since the identification of an "epidermoid formation," consisting of squamous cells in the anterior epitympanum of developing fetal temporal bones.[25] This location coincides with the most common location of congenital cholesteatomas. The epidermoid formation plays a role in organizing the formation of the tympanic ring, medial layer of the TM, and epithelial lining of the eustachian tube. It is not generally found beyond 33 weeks gestation, but failure to atrophy is thought to result in congenital cholesteatomas. Another theory is that squamous cell epithelium intended for the external canal is pinched off during the formation of the TM. Consequently, squamous cell epithelium trapped in the middle ear with an intact drumhead leads to congenital cholesteatoma formation.

The pathogenesis of acquired cholesteatoma has been debated for well over a century. Although no single theory has been universally accepted, the most popular explanation for attic cholesteatomas is the invagination theory.[26] Eustachian tube dysfunction is thought to cause retraction of the TM. The pars flaccida, lacking a fibrous layer, is more vulnerable to this effect. As the retraction pocket deepens, desquamated keratin accumulates and cannot be cleared from the pocket, and a cholesteatoma develops. On otoscopic examination, it may appear that the patient has a perforated TM, but closer examination reveals an intact drumhead with an attic retraction pocket. The term primary acquired cholesteatoma is applied to describe this mechanism of cholesteatoma formation.

It has also been theorized that cholesteatomas may develop as a result of the ingrowth of squamous cell epithelium from the lateral surface of the TM through a perforation. Convincing evidence comes from a study of cytokeratin 10, which is found in both the meatal epidermis and cholesteatoma matrix but not in the middle ear mucosa.[27]

The implantation theory proposes that squamous cell epithelium is displaced into the middle ear owing to surgery, trauma, or a foreign body. The prevalence of cholesteatomas following tympanostomy tubes is only 0.5%.[28] However, it is thought that ventilation tubes probably prevent far more cholesteatomas than they cause.

Another, less popular concept is the metaplasia theory in which healthy cuboidal cells transform into squamous cell epithelium in chronic inflammation. Although in situ metaplasia has been demonstrated, actual formation of cholesteatoma has not been proven.[29] In addition, this theory does not explain the common location of cholesteatomas in the posterosuperior aspect of the middle ear despite diffuse mucosal inflammation throughout the entire middle ear space. Lastly, it is proposed that cholesteatomas may develop from papillary ingrowth of basal cells through the basement membrane.[30]

Bone Erosion

Bone erosion is an important aspect of COM because it can lead to various complications ranging from hearing loss and labyrinthine fistula to facial

nerve paralysis and life-threatening intracranial complications. Originally, it was believed that pressure from the cholesteatoma itself led to necrosis of bone. However, experimental evidence is lacking. Cholesteatomas have been found to exert pressures between 1.31 and 11.88 mm Hg. These pressures fall well below the capillary perfusion pressure of 25 mm Hg. For this reason, it is unlikely that cholesteatoma, which is often associated with a perforation and open cavity, can produce enough pressure to occlude directly capillary flow and cause tissue anoxia. However, it is thought that the pressure exerted on adjacent bone by cholesteatoma causes bone erosion through the activation of osteoclasts.[31]

Histologic examination of the lytic zone between the submucosa and the necrotic bone reveals marked capillary proliferation along with infiltration of histiocytes that contain lysosomal enzymes such as protease, hyaluronidase, cathepsin, and acid phosphatase.[32] For this reason, it is argued that hyperemia in the setting of osteolytic enzymes, as opposed to anoxia, is responsible for bony erosion. Animal and human studies further support this theory.[33] In bone erosion, osteoclasts, which are known to be responsible for bone reabsorption, and collagenase have been identified. Prostaglandin E_2 and interleukin-1α are among various other mediators found to regulate bone reabsorption.[34]

Studies clearly demonstrate that bone erosion occurs in COM with or without the presence of cholesteatoma.[33,35] Although the true mechanism is not fully understood, it is generally accepted that inflammation is a major factor since granulation tissue is often associated with ossicular erosion.[36] The inflammatory process is thought to lead to infiltration and activation of osteoclasts and mononuclear cells containing various proteolytic enzymes. Microscopically, a subepithelial layer of granulation tissue has been identified adjacent to eroded bone. The cellular components of this layer are identical in cases with and without cholesteatoma.[37] The higher frequency of bone destruction in the setting of cholesteatoma may be related to the excellent environment cholesteatomas provide for persistent bacterial infection and chronic inflammation.

Tympanosclerosis

Tympanosclerosis is common in patients with a history of recurrent AOM, COM, or multiple ventilation tubes. Macroscopically, white crescent-shaped deposits are seen within the TM. In this setting, the term myringosclerosis is sometimes applied to indicate that the process is restricted to the TM. Microscopically, hyalinization of collagen and deposition of calcium are found within the lamina propria. Tympanosclerosis need not be limited to the TM. It may extend into the middle ear cleft to involve the ossicular heads and basement membrane of the middle ear mucosa. If the oval window is involved, stapes fixation may result and contribute to a significant conductive hearing loss.

Cholesterol Granuloma

Cholesterol crystals are frequently found within the submucosal lining of ears with COM. The process probably results from hemorrhage into the mucosa of the middle ear or within obstructed mastoid air cells, with the crystals representing breakdown products of the erythrocyte cell membrane. A foreign body response to these crystals leads to formation of a cholesterol granuloma. On macroscopic examination, a yellow-brown murky viscous material is found within the middle ear space or within the obstructed mastoid cells. Microscopically, the cholesterol crystals, foreign body giant cells, and associated inflammatory cells are identified.

DIAGNOSIS

History

Approximately one-third of individuals with COM have their diagnosis made as an incidental finding during routine physical examination. However, when symptomatic, the two hallmark presenting symptoms are otorrhea and hearing loss. Pain is unusual with COM and indicates either a reactive external otitis or the possibility of a developing intratemporal or intracranial complication.

The nature of the otorrhea is helpful in describing the specific type of COM. Profuse, intermittent, mucoid drainage is commonly noted in chronic suppurative otitis media without cholesteatoma. Malodorous otorrhea is rare in this setting. Conversely, patients with COM associated with cholesteatoma often describe scant but persistent, purulent, and foul-smelling otorrhea. Blood-stained drainage is often noted with granulation tissue or polyps. It is important to realize that cholesteatoma is frequently hidden beneath this abnormal inflammatory tissue.

Some patients with chronic middle ear disease will have normal hearing or only mild hearing loss because the granulation tissue or cholesteatoma successfully transmits sound energy to the inner ear. When present, hearing loss is usually conductive or mixed in nature.

The degree of hearing loss will depend on the size and location of the TM perforation and the status of the middle ear. Large perforations will generally cause greater hearing loss compared with smaller defects. In addition, perforations overlying the posterior part of the mesotympanum, and thus the round window niche, usually cause more severe degrees of conductive hearing loss because the TM is no longer protecting the round window membrane from direct sound energy transfer. As a result, there is reduction of the "baffle" effect, leading to a change in the cochlear mechanics.[38] Ossicular chain involvement will also cause conductive hearing loss. Chronic otitis media is often associated with bone erosion, leading to discontinuity of the ossicular chain. The long process of the incus is particularly vulnerable to postinflammatory necrosis, resulting in discontinuity or fibrous union at the incudostapedial joint. Tympanosclerosis may also arise in the setting of chronic inflammation. Usually, this is limited to fairly innocent changes in the TM but may involve the middle ear as well. As a result, the stapes footplate or other ossicles may become fixed. Finally, fibrosis or inflammatory granulation tissue within the middle ear can lead to limitation of TM and ossicular mobility.

Patients who have previously undergone a tympanoplasty with or without mastoidectomy for treatment of middle ear disease will often be left with a grafted TM that may be thickened, poorly mobile, and lacking the normal conical shape. Such changes, along with a shallow middle ear space and disruption of the ossicular chain, will also lead to conductive hearing loss.

Although an association between COM and conductive hearing loss is clearly acknowledged, the relationship to sensorineural hearing loss (SNHL) remains controversial. Several retrospective studies have compared the preoperative bone-conduction thresholds of diseased ears to the normal contralateral ears in patients with unilateral COM. These results have demonstrated an association between SNHL and COM.[39,40] However, the magnitude of hearing loss was small, ranging between 5 and 12 dB depending on the frequency tested. From a clinical standpoint, this hearing loss is less significant compared with the conductive losses caused by COM.

One proposed mechanism for SNHL is that inflammatory toxins cross the round window membrane, entering the scala tympani and causing irreversible cochlear damage. In this setting, one would expect high-frequency hearing thresholds to be most affected owing to the tonotopic arrangement of the cochlea, with high frequencies stimulating those cells closer to the round window in the basal turn. One retrospective review, limited to those patients diagnosed with unilateral COM with cholesteatoma, showed statistically significant interaural differences in both speech discrimination scores and pure-tone thresholds at 4,000 Hz independent of patient age.[41] Given that speech discrimination is believed to be significantly influenced by high-frequency cochlear function, these findings lend support to the theory of inflammatory toxins entering the inner ear. However, temporal bone studies of individuals with a history of COM failed to demonstrate histologic changes associated with inflammation near the round window membrane, even though an interaural difference in bone-conduction thresholds was appreciated.[42] Thus, although there is convincing evidence demonstrating a relationship between COM and SNHL, the clinical implications and actual mechanism of injury remain to be determined.

Physical Examination

Examination should include inspection of the pinna and postauricular region. Any scars from prior otologic surgery should be noted. Otoscopic examination with the use of a binocular microscope is imperative. Care should be taken to clear all cerumen and debris from the external auditory canal (EAC) so that the entire drumhead is visible. The EAC should be evaluated for signs of secondary otitis externa as well as scarring, which, again, would suggest prior surgery. The presence and specific location of aural polyps and granulation tissue should be noted. Attention is then turned to the TM. The size and location of any perforation should be detailed in the medical record, usually in the form of a diagram or photograph. Chronic otitis media in the absence of cholesteatoma is usually associated with a central perforation, whereas cholesteatomas are commonly observed in the setting of defects in

the pars flaccida or marginal perforations of the TM. Depending on the size of the perforation, as well as the presence of granulation tissue and polyps, the middle ear mucosa can often be examined. In the case of a dry perforation, the mucosa may appear completely normal. However, otorrhea is usually associated with edematous, inflamed mucosa. An attempt should be made to suction the middle ear space free of all secretions. If microbiologic specimens are indicated, fluid should be obtained from the middle ear, rather than the EAC, and it should be sent for both aerobic and anaerobic cultures. Perforations of the TM can also provide an opportunity to inspect the middle ear bony structures. Particular attention is focused on bone erosion and ossicular disruption. Depending on the patient's age and pain threshold, the microscopic examination and cleaning might have to be performed in the operating room with the use of sedation.

The physical examination should also include a fistula test if the patient reports vestibular symptoms, inspection of the nasopharynx, including the eustachian tube orifices, and gross assessment of hearing with a 512 Hz tuning fork.

Audiometric Testing

Every initial evaluation for COM should include audiometric testing with air and bone pure-tone thresholds. Appropriate clinical masking is imperative as patients often have bilateral involvement and mixed hearing loss. The degree of hearing loss is often helpful in determining the extent of the middle ear disease. Perforations of the TM can account for 15 to 20 dB of conductive hearing loss. When perforations are accompanied by ossicular chain damage, the hearing loss can increase to between 30 and 50 dB. Finally, ossicular chain disruption with an intact drumhead can account for 55 to 65 dB of conductive hearing loss.

Speech discrimination testing is also useful. Specifically, speech reception thresholds can help determine whether a patient is a candidate for middle ear reconstructive surgery. Usually, the speech discrimination and hearing loss correlate, but a small percentage of patients have an unexpectedly low speech reception threshold and thus may benefit less from surgical intervention. Furthermore, ears with very poor speech discrimination scores are not suitable for middle ear reconstruction unless there is reason to believe that the poor scores are owing to an inability to present a stimulus intense enough to obtain the maximal discrimination score.

Radiographic Studies

With the advent of high-resolution computed tomography (CT) in the 1980s, temporal bone plain films and polytomography have essentially become obsolete. The advantages of CT scanning over these traditional studies include superior definition of both bone and soft tissue detail and lower levels of exposure to radiation compared with polytomography (Figure 10–1).

Standard temporal bone CT technique involves 1.5 mm sequential cuts in both the axial and coronal projections. The axial images should extend from the arcuate eminence superiorly to the jugular fossa inferiorly. These views will be most helpful in evaluating the sinus tympani, facial recess, lateral semicircular canal, ossicles, and horizontal portion of the facial nerve. The coronal images span from the bony portion of the eustachian tube anteriorly to the posterior semicircular canal posteriorly and will be of value in assessing Prussak's space, the tegmen tympani, the ossicles, and the perigeniculate and vertical seg-

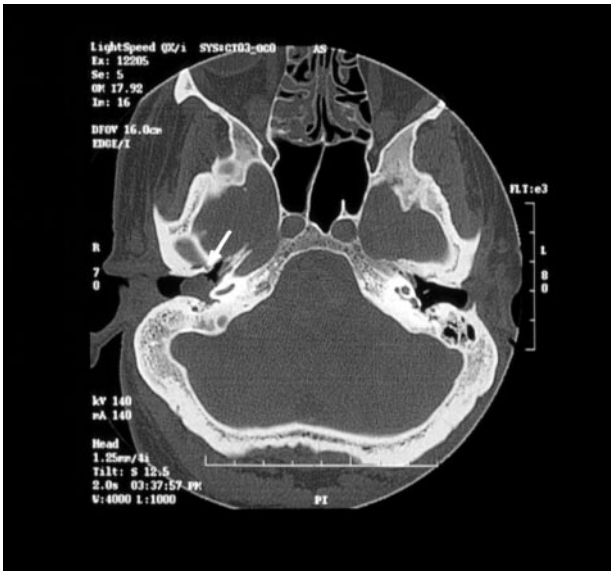


FIGURE 10–1. Axial computed tomographic scan showing inflammatory soft tissue (*arrow*) filling the right middle ear with a large fibrotic polyp extending into the external auditory canal and obscuring visualization of the tympanic membrane.

ments of the facial nerve (Figures 10–2 to 10–4). Cholesteatoma does not enhance with intravenous contrast, and it is not routinely administered unless there is a high suspicion for a vascular abnormality, tumor, or intracranial complication.[43]

Although superior for soft tissue resolution, magnetic resonance imaging (MRI) findings are usually nonspecific in COM. Bony details are not defined because both bone and air appear dark. Although an inflamed matrix may enhance with gadolinium administration, cholesteatoma is extremely difficult to differentiate from edematous mucosa, inspissated secretions, and granulation tissue. Studies have confirmed that CT scan is superior to MRI in the evaluation of cholesteatoma and uncomplicated COM.[44,45] However, MRI with gadolinium enhancement will serve to delineate intracranial complications better and is warranted when intracranial involvement is suspected (Figure 10–5).

Information provided by preoperative CT scans includes extent of disease and temporal bone pneumatization, both of which may influence the surgical approach and the management of the posterior canal wall. Anatomic variations such as a low-lying tegmen mastoideum, a large or anterior sigmoid sinus, or an anomalous facial nerve may also be appreciated and thus influence operative decisions. Computed tomography also allows identification of potential risks such as bony dehiscence

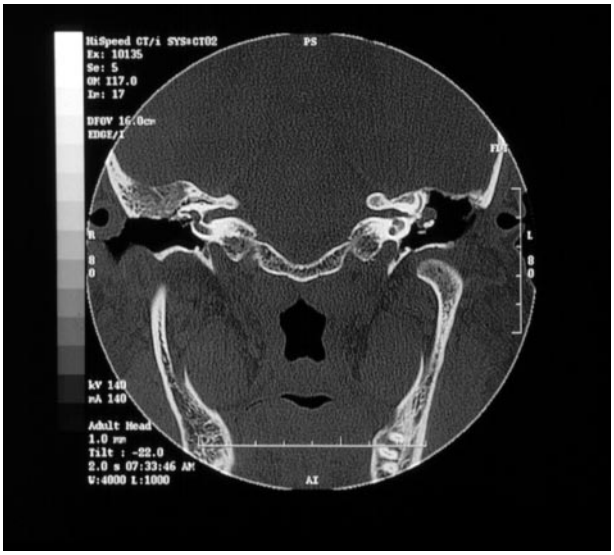


FIGURE 10–2. Coronal computed tomographic scan scan demonstrating attic cholesteatoma engulfing the ossicular chain in the left ear.

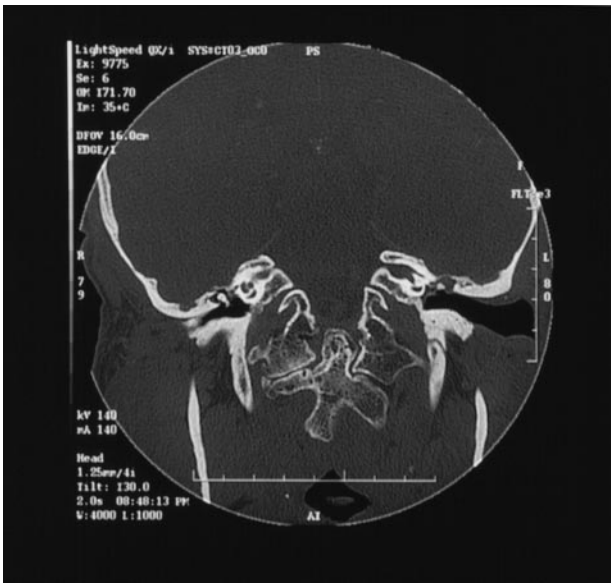


FIGURE 10–3. Coronal computed tomographic scan scan demonstrating cholesteatoma in Prussak's space with erosion of the scutum in the left ear.

of the tegmen tympani, facial nerve, lateral semicircular canal, jugular bulb, or sigmoid sinus. When using CT to evaluate disease in the temporal bone, one must remember that volume averaging of thin bony plates may lead to a false impression of bony dehiscence. Such misinterpretations can be minimized by examining the coronal sections of the temporal bone and comparing the diseased ear with the contralateral ear.

In an attempt to define the clinical utility of CT scans in COM, a study was conducted of 42 patients who underwent preoperative imaging prior to surgical intervention.[46] The most reliable finding was complete absence of soft tissue within the middle ear space. In all seven cases failing to demonstrate soft tissue on CT scan, no cholesteatoma was identified at the time of surgery. Soft tissue accompanied by bone erosion was found to correlate reliably with cholesteatoma. Only one-third of patients with soft tissue masses in the absence of bone erosion were ultimately diagnosed with cholesteatoma. Thus, CT scanning lacks specificity in differentiating cholesteatoma from granulation tissue, mucosal edema, and effusion unless bone erosion has occurred. Although CT is accepted as the "gold standard" radiographic study for COM, universal guidelines dictating its use have not been established.[47] Radiographic imaging does not alter surgical man-

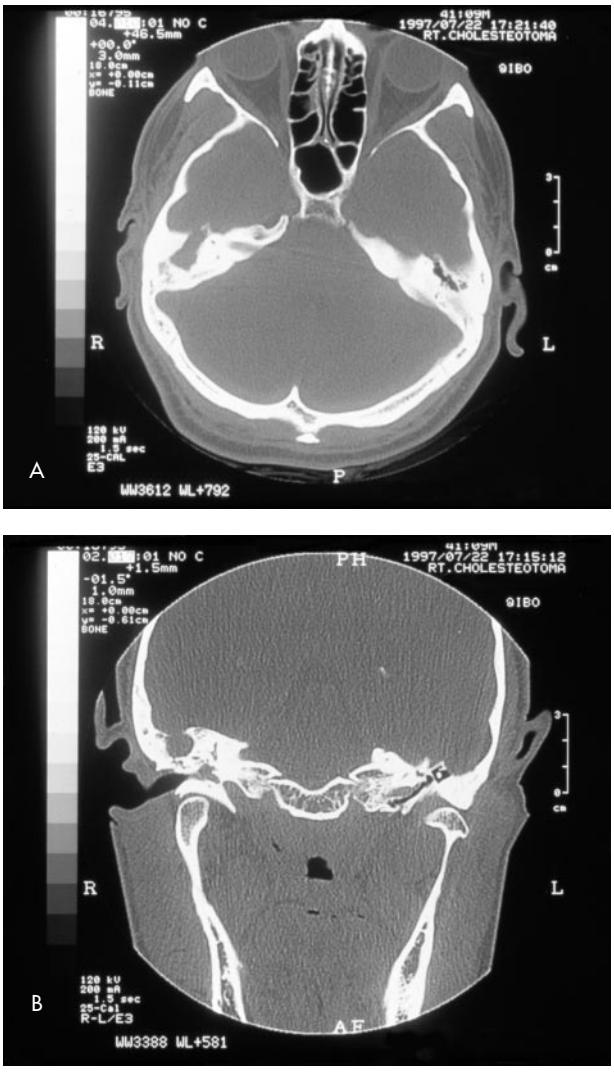


FIGURE 10–4. *A*, Axial scan showing poor pneumatization bilaterally, with aeration of the underdeveloped left mastoid antrum but cholesteatoma expanding the mastoid antrum on the right. *B*, Coronal scan of the same patient showing circular expansion of the antrum and dehiscence of the tegmen mastoideum by bone erosion.

agement of most COM cases because the extent of the middle ear and mastoid disease is ultimately exposed during surgical intervention. Given this limited clinical utility, along with cost considerations, radiation exposure, and need for sedation in children, most otologists recommend CT scanning only on a selective basis. Computed tomography is especially helpful in revision cases, when previous anatomic modifications are unknown to the surgeon, or when there is concern for recurrent cholesteatoma that may be hidden from otoscopic view. Other indications include congenital ear malformation, vertigo, a positive fistula test, and facial nerve paralysis. Finally, imaging is warranted in the setting of suspected central nervous system complications such as a brain abscess, cerebritis, or lateral sinus thrombosis.

Special Considerations

Histiocytosis X, an uncommon idiopathic disease associated with proliferation of histiocytes, can present in a fashion that mimics COM.[48] The most common otologic symptom is middle ear suppuration or otorrhea. Other manifestations include otitis externa, SNHL, vertigo, and mastoid destruction. Poor response to antibiotics, the presence of granulation tissue in the absence of suppuration, or associated signs of histiocytosis X such as hepatosplenomegaly, diabetes insipidus, bone lesions, and exophthalmos should alert one to this possibility. A mastoidectomy is often required to obtain tissue for diagnosis. A high index of suspicion is necessary because treatment of histiocytosis X entails low-dose radiation and occasionally chemotherapy. Eosinophilic granuloma, a related idiopathic condition that usually has a monostotic presentation, may present in a similar fashion. Again, the diagnosis is made based on histologic examination of tis-

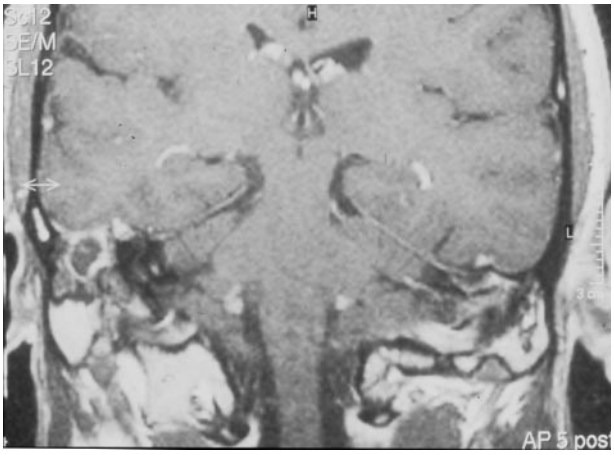


FIGURE 10–5. Magnetic resonance image of the patient in Figure 10–4 demonstrates hypointense signal in the center of the mass, consistent with cholesteatoma. Note also the enhancement of the surrounding matrix and the intact adjacent dura of the middle fossa. This finding rules out the possibility of an encephalocele, which may have been suspected from the computed tomographic findings.

sue obtained at surgery. These conditions should always be suspected when there is bone destruction in the absence of cholesteatoma.

Another rare cause of COM is tuberculosis. Patients present with thin, odorless otorrhea that is often insidious and painless. Multiple TM perforations are typically found on otoscopic examination. Labyrinthine involvement can lead to SNHL. Diagnosis is important because patients require treatment with multiple antituberculous agents. After successful medical eradication of active infection, tympanoplasty may be performed.

A final consideration in the evaluation of COM is the potential complication of Meniere's disease, which has been found to occur subsequent to, and in some cases simultaneously with, COM.[49] In such patients, endolymphatic hydrops develops that is secondary to the chronic inflammation, leading to vertigo and/or fluctuating SNHL. The differential diagnosis includes labyrinthine fistula and toxic (serous) labyrinthitis.

BACTERIOLOGY

Although a severe episode of AOM can precede COM, the microorganisms responsible for each diagnosis differ significantly. Whereas *Streptococcus pneumoniae, Haemophilus influenzae*, and *Moraxella catarrhalis* are the common microorganisms responsible for AOM, *Staphylococcus aureus* and *Pseudomonas aeruginosa* are the common aerobic isolates in COM. Both are indigenous microorganisms. *Staphylococcus aureus* is a gram-positive coccus that colonizes the nares. *Pseudomonas aeruginosa* is a gram-negative rod that is extremely common in moist environments and is generally found colonizing the EAC. Thus, it is not surprising that it is often an important pathogen in chronic middle ear disease.

Anaerobic microorganisms may also play an important role in COM. Individual isolates and percentages differ somewhat from study to study, most likely because of differences in culture technique.[50] However, when fastidious microbiologic techniques are employed, the most common anaerobic microorganisms cultured include *Fusobacterium* sp, pigmented *Prevotella, Bacteroides fragilis*, and *Porphyromonas* (previously known as *Bacteroides melaninogenicus*).[51]

Overall, COM should be viewed as a polymicrobial disease since multiple isolates are usually recovered from a single culture. In addition, nearly one half of all COM is caused by a combination of aerobic and anaerobic microorganisms.[50]

MEDICAL MANAGEMENT

The treatment of COM generally begins with local care of the ear and outpatient medical management. For medical management to be successful, aural toilet is imperative. This intervention requires repeat microscopic examination of the ear and diligent suctioning. The main goal is to remove debris from the EAC overlying the TM and middle ear cleft so that topical antimicrobial agents can successfully penetrate to the middle ear mucosa. For patients with otorrhea secondary to cholesteatoma, the hope is to minimize granulation tissue and perhaps achieve a dry ear prior to surgical intervention. However, this is not always feasible, and some refractory cases will require surgery to eliminate the otorrhea.

It may be possible to reduce the amount of granulation tissue contributing to the ongoing otorrhea, particularly in the infected mastoid cavity. Before proceeding, one must be reasonably convinced that the critical landmarks are properly identified and that there is not a dural defect with an encephalocele present. Judicious suctioning of immature granulation tissue or polyps may be undertaken provided that the otologist is certain of the landmarks in the cavity. Once the base of the granulation tissue is reached, topical cautery with silver nitrate or 30% trichloroacetic acid is helpful. Silver nitrate is more hemostatic. Sometimes topical 2% gentian violet is helpful as an astringent if the bed of granulation remains active despite initial or repeated treatment. If the granulation tissue is of significant thickness or if a polyp is present, cup forceps may be needed to reach the base of the granulation. A wick may be indicated if significant canal edema has resulted from secondary otitis externa.

Most often, it is appropriate to proceed to surgery if the ear does not respond to microscopic débridement and ototopical management. However, if more aggressive medical management is contemplated to avoid surgery or to obtain a dry ear prior to surgery, any soft tissue that is removed from the external canal or mastoid cavity should be sent for microbiologic culture and sensitivity testing. When otorrhea is cultured, specimens should ideally be obtained from direct middle ear aspiration rather than from the ear canal whenever possible. Care

must be taken to avoid contaminating the aspirate with debris from the EAC since microorganisms colonizing this region, *Pseudomonas* in particular, are recovered from the middle ear space in only 50% of cases.[52] Cultures with sensitivity are necessary to guide antibiotic therapy because most patients with COM have been treated with multiple prior antibiotic regimens. For this reason, it is not surprising that two-thirds of all COM patients are infected with β-lactamase-producing microorganisms.[53]

Topical medications may include antibiotics, antifungals, antiseptics, and corticosteroid preparations alone or in combination with other medications. If otorrhea is profuse, it may be helpful to have the patient irrigate the ear daily with a body temperature half-strength solution of acetic acid (50% white vinegar diluted with warm water) prior to application of otic drops.

Popular topical antimicrobial agents have included aminoglycosides such as gentamicin, tobramycin, and neomycin in combination with polymyxin B sulfate because of their antipseudomonal properties. Controversy has surrounded these agents since animal studies have demonstrated ototoxic effects.[54,55] However, human studies have failed to show evidence that otic drops cause SNHL in patients with COM.[56] This discrepancy may be primarily attributed to the fact that the animals used in these experimental models have a round window membrane that is much thinner than in humans, particularly when there is no reactive inflammation. Furthermore, the round window niche is extremely shallow and completely exposed in these experimental animals, but it tends to be deep and often protected by a mucosal pseudomembrane in humans.[55] More importantly, most otolaryngologists believe that the risk of SNHL from uncontrolled infection owing to COM itself is greater than the risk associated with potentially ototoxic drops.

The use of neomycin in ototopical preparations continues to be extremely widespread owing to long-standing prescribing habits and low cost, despite the fact that almost no strains of *Pseudomonas* remain sensitive to this medication. In addition, there is a fairly high incidence of localized and diffuse allergic reactions to the topical use of neomycin. For these reasons, preparations containing neomycin should eventually fade from the ototopical armamentarium.

More recently, fluoroquinolone antibiotics such as ciprofloxacin and ofloxacin have gained popularity because of their antipseudomonal properties, minimal bacterial resistance, lack of ototoxicity, and potential oral route of administration. Studies have found topical ciprofloxacin to be as efficacious as topical tobramycin and to be superior to topical gentamicin.[57,58] According to US Food and Drug Administration (FDA) regulations, fluoroquinolones are contraindicated in children, pregnant women, and nursing mothers.

Studies have revealed that the addition of topical corticosteroids in combination with antimicrobial agents improves response rates compared with placebo or corticosteroids alone.[59,60] The anti-inflammatory properties of corticosteroids are thought to allow increased antibiotic levels in the middle ear mucosa by decreasing tissue edema. Currently, the ototopical combination that provides the best profile of efficacy and safety consists of a fluoroquinolone and corticosteroid preparation in suspension form. This preparation is significantly more costly than traditionally prescribed eardrops.

The use of systemic antibiotics in COM is limited by several factors. Once again, antibiotic penetration into the middle ear may be hampered by mucosal edema. Systemic aminoglycosides carry a risk of ototoxicity and require parenteral administration with monitoring of serum levels. Oral ciprofloxacin has proven to be a safe and effective treatment for adults with COM; however, safety in patients under 18 years of age has not been established.[61] Thus, the pediatric population is limited to parenteral antibiotics such as broad-spectrum penicillins, cephalosporins, and aminoglycosides.

Patients with a TM perforation should be instructed about the importance of keeping their ear canal and middle ear dry. An ear plug or a cotton ball impregnated with petrolatum jelly should be used when showering, and ear plugs should be worn when swimming. Hearing aid molds can cause an increase in middle ear humidity, which may initiate or perpetuate mucosal infection. For this reason, patients with recalcitrant otorrhea who are resistant to medical treatment and are not candidates for tympanoplasty should be fitted with bone-conduction hearing aids or osseointegrated bone-anchored hearing aids.

Medical management of COM may be difficult for both the patient and the physician. Multiple

office visits are often required for adequate aural toilet. In addition, patients are asked to comply with a regimen that may include not only daily irrigation but also multiple administrations of otic drops throughout the day. Medical treatment usually requires 14 to 21 days. It is disappointing that up to 50% of patients will have actively draining ears even when diligently compliant with a combination of corticosteroid and antibiotic drops.[59]

SURGICAL MANAGEMENT

The primary goal of surgery for COM is to eradicate disease and obtain a dry, safe ear. Restoration of hearing is by necessity a secondary consideration because any attempt at middle ear reconstruction will fail in the setting of persistent inflammation and otorrhea.

Absolute indications for surgical intervention include impending or established intratemporal or intracranial complications as described in Chapter 11. Various pathologic conditions within the middle ear, such as cholesteatoma and chronic fibrotic granulation tissue, are irreversible and require elective surgical attention. In addition, patients with otorrhea failing to respond to medical treatment are surgical candidates, as well as those who respond but are left with a correctable conductive hearing loss or a TM defect.

There are numerous surgical techniques used to address COM. Ultimately, the procedure chosen will depend on the nature and extent of the disease, as well as the surgeon's training and experience. If possible, it is ideal to restore eustachian tube function prior to surgical intervention. However, attempts at performing eustachian tuboplasty have been found to be extremely cumbersome and ineffective. Studies have proven that transtympanic ventilation can lead to successful surgical results, even in the setting of eustachian tube dysfunction.[62]

Nomenclature in Surgery for Chronic Otitis Media

- *Myringoplasty*—an operation limited to superficial repair of TM defects, without exploration of the middle ear cleft. These are generally performed in an office or ambulatory operating room setting.
- *Tympanoplasty*—an operation involving exploration of the middle ear cleft through a transcanal approach or through a postauricular incision. This is performed to eradicate disease from the middle ear, repair TM defects, and reconstruct the ossicular chain. This procedure is frequently performed in conjunction with a mastoidectomy.
- *Canalplasty*—an operation to widen the bony portion of the EAC, generally performed to improve access and visualization for tympanoplasty and postoperative care. This is often required for prominent bony overhangs associated with the tympanomastoid and tympanosquamous sutures.
- *Ossiculoplasty*—a procedure to reconstruct the ossicular chain.

Open (Canal Wall Down) Mastoidectomy Techniques These procedures have as their unifying theme a surgical strategy involving removal of varying portions of the bony EAC to obtain improved access to the epitympanic and mesotympanic spaces for management of chronic ear disease. These approaches also leave some or all of the diseased spaces of the temporal bone permanently exteriorized to help avoid recurrent disease:

- *Radical mastoidectomy*—an operation to eradicate or exteriorize extensive middle ear disease by removing the posterior bony ear canal to open the middle ear, mastoid, and epitympanum into one common cavity. In doing so, remnants of the tympanic membrane, malleus, and incus are removed, leaving only the remaining portions of the stapes. The TM is not reconstructed, and the eustachian tube may be left open or permanently obstructed with tissue grafts.
- *Modified radical mastoidectomy*—differs from a radical mastoidectomy in that an attempt is made to preserve or reconstruct the middle ear. Sometimes healthy TM and ossicular remnants are preserved. In the classic Bondy modified radical procedure, atticoantral cholesteatoma is exteriorized without disturbing the intact pars tensa of the TM or the intact ossicular chain. This procedure is appropriate in the rare instance when the disease extends from the pars flaccida region lateral to the ossicles with extension into the antrum. More commonly, the ossicles are engulfed or eroded by disease, and primary or staged reconstruction is required. When associated middle ear reconstruction is performed,

many surgeons still apply the term modified radical mastoidectomy, but the procedure is more appropriately called an open or canal wall down mastoidectomy with tympanoplasty.

- *Atticotomy*—removal of ear canal bone including the lateral wall (scutum) of the epitympanum to expose and exteriorize limited attic disease, usually lateral to healthy ossicles.
- *Atticoantrostomy*—a strategy that ultimately creates a defect similar to the modified radical mastoidectomy by pursuing a posterior extension of the atticotomy approach. The surgeon's intent is to exteriorize rather than resect the matrix of the cholesteatoma. The operation is performed by entering the attic from the ear canal and then proceeding posteriorly, gradually removing posterior ear canal bone and exposing disease in the aditus and antrum until it is fully exteriorized. This approach is appealing in that it exposes only that portion of the mastoid affected by disease; however, it provides limited opportunity to identify the surgical landmarks needed to prevent injury to the facial nerve.
- *Meatoplasty*—an operation performed to widen the cartilaginous and bony external auditory meatus of patients undergoing a canal wall down procedure. The operation allows postoperative inspection of the mastoid cavity and routine cleaning of debris from the mastoid bowl.

Closed (Intact Canal Wall) Mastoidectomy Techniques

- *Cortical mastoidectomy*—removal of disease that is limited to the mastoid antrum and air cell system, preserving the posterior bony EAC wall. Generally, when this operation is performed without a tympanoplasty, it is for chronic mucosal disease rather than cholesteatoma. The middle ear contents are not disturbed, although a tympanostomy tube may be placed for improved ventilation.
- *Tympanoplasty with intact canal wall mastoidectomy*—an operation in which disease is removed from the mastoid and middle ear while preserving the posterior bony wall of the EAC. Often the mesotympanum is exposed by developing a posterior tympanotomy through the facial recess. Middle ear reconstruction may be performed primarily, but it is often staged in cholesteatoma cases.
- *Obliteration technique*—a procedure in which muscle or a musculoperiosteal flap is used to obliterate a portion of the cavity following a canal wall down mastoidectomy. If successful, the size of the defect is minimized, which may avoid the need for long-term cavity care.
- *Canal wall reconstructive technique*—reconstruction of the posterior canal wall may be undertaken following a canal wall down mastoidectomy, either primarily or at a later operation. In doing so, an attempt is made to convert the open cavity into a closed cavity.

Anesthesia and Patient Preparation

Otologic surgery can be performed using local or general anesthesia. The length of the procedure, patient's age and ability to cooperate, and the anticipated level of stimulation must all be factored into the decision. Children will obviously require general anesthesia. However, local anesthesia may be preferred in cases primarily involving ossicular reconstruction since it is possible to assess hearing improvement intraoperatively. When surgery involves the TM, some surgeons will request that the anesthesiologist avoid use of nitrous oxide, which may increase middle ear pressure and interfere with positioning of the graft or TM.

Facial nerve monitoring is recommended in congenital temporal bone anomalies because of the associated risk of injury to a malpositioned facial nerve. Monitoring is also recommended for revision cases if the surgeon is uncertain of prior anatomic modifications or if the canal of the facial nerve is known to be dehiscent. Some surgeons advocate the routine use of a facial nerve monitor in every chronic ear operation. Although this is a controversial issue, those on both sides of the argument strongly agree that monitoring is no substitute for detailed knowledge of temporal bone anatomy and adequate technical preparation on the part of the surgeon.[63]

The use of perioperative, prophylactic antibiotics has been debated for decades. Many surgeons advocate the preoperative use of a broad-coverage antibiotic such as a cephalosporin, although controlled studies have suggested that the use of antibiotics does not reduce the risk of postoperative infection.[64]

Local injection of 1 or 2% lidocaine with epinephrine (1:100,000) provides excellent vasoconstriction for transcanal procedures. More concentrated epinephrine solutions prepared by the surgeon or an assistant might provide improved vasoconstriction, but errors in preparation have led to fatal consequences owing to the induction of cardiac arrhythmias. A 27- or 30-gauge needle mounted on a 1 mL syringe is recommended for ear canal injections. The most important injection is along the thick vascular strip region of the canal skin, lying between the tympanomastoid and tympanosquamous sutures. A subperiosteal injection will dissect freely under the skin and provide excellent vasoconstriction and anesthesia. Another subperiosteal injection inferior to the tympanomastoid suture will control bleeding from the inferior canal wall skin. Slow infiltration will avoid the formation of a bleb beneath the canal skin that may obscure the surgical field. When performing surgery under local anesthesia, additional injections should be made at the bony-cartilaginous junction in all four quadrants. In addition, the local anesthetic is injected into the scalp overlying the mastoid for postauricular procedures. Such injections are made prior to preparation of the ear to allow adequate time for vasoconstriction.

Most cases will not require extensive shaving of hair. Instead, adhesive drapes can be used to keep hair out of the surgical field. The ear and postauricular region are prepared in the standard fashion, recognizing that it is not possible to sterilize the ear canal completely. If the patient has an open middle ear cleft owing to a TM perforation, a cotton ball should be placed in the canal to prevent ingress of the antiseptic solution.

Tympanoplasty

Hearing restoration is of secondary importance in the surgical management of COM since any attempt at tympanoplasty and ossicular chain reconstruction will fail in the setting of persistent otorrhea, inflammation, and eustachian tube dysfunction. Eradication of mucosal disease remains the priority, and if it is not reasonable to reconstruct the ossicular chain at the first operation, middle ear reconstruction can be performed at a later date.

Graft Material and Tympanoplasty Techniques The goal of tympanoplasty is to repair the TM with a connective tissue graft in the hope that squamous cell epithelium will migrate over the graft and seal the perforation. Various grafting materials are available. Autogenous temporalis fascia is used most often because it is readily available through a postauricular incision and is extremely effective. Other alternatives include tragal perichondrium, periosteum, and vein. Preserved homograft materials such as cadaveric TM, dura, and heart valve have limited application because of concern for viral infection and the transmission of Creutzfeldt-Jakob disease.

The two classic types of tympanoplasty are the lateral technique and the medial technique. The terms lateral and medial refer to the final relationship of the graft to the fibrous layer of the TM remnant and the annulus tympanicus. In both approaches, the graft is secured medial to the malleus handle, if present. The lateral technique is more technically demanding, but many surgeons believe that it provides more reliable results when repairing large anterior or pantympanic defects. It is also useful when the ear canal anatomy is unfavorable, requiring an extensive canalplasty. The medial technique tympanoplasty is commonly employed for smaller perforations in the posterior and central portions of the TM. However, with careful attention to surgical principles, this technique can be applied regardless of the size or location of the defect.

Incisions Although various incisions have been described in the past, the two most common approaches used today are the transcanal tympanomeatal flap and the postauricular approach with creation of a vascular strip. Each incision has advantages and disadvantages. Ultimately, the ear canal anatomy, the position of the perforation, and the surgeon's judgment will determine the approach selected.

The tympanomeatal flap is created by making longitudinal canal incisions superiorly, along the tympanosquamous suture line (12 o'clock) and inferiorly (6 o'clock). The incisions extend laterally from the TM approximately 6 to 8 mm and are connected by a transverse incision parallel to the annulus. Elevation of the skin and periosteum anteriorly creates a medially based U-shaped flap. The tympanomeatal flap allows for excellent exposure of the posterior part of the mesotympanum and is ideal for perforations located in this area. The flap is usually elevated using a transcanal approach and does not require a

postauricular incision. Closure is simple, postoperative pain is modest, and the ear heals quickly. However, it is important to realize that in inexperienced hands, the transcanal approach can be cumbersome owing to limited exposure imposed by the use of a speculum and the ear canal configuration.

The major disadvantage of the transcanal approach using a tympanomeatal flap is poor exposure of the anterior mesotympanum. As the flap is reflected forward, it remains within the surgical field and obscures visualization of this area. As a result, anterior perforations are difficult to address, especially in the setting of a prominent anterior canal wall bulge. It is also difficult to visualize and address irreversible mucosal disease in the protympanum and eustachian tube orifice, both of which may contribute to surgical failures by preventing satisfactory postoperative eustachian tube function. The swinging door technique may be used to improve anterior exposure. The tympanomeatal flap is bisected by an incision that extends medially through the annulus and TM remnant into the posterior margin of the perforation. The divided tympanomeatal flap can be reflected superiorly and inferiorly for increased exposure. Separating the drumhead remnant from the malleus handle may also be helpful in improving exposure to the anterior mesotympanum.

The postauricular, vascular strip approach is an alternative means to achieve excellent exposure of the entire mesotympanum. A superior ear canal incision is made longitudinally along the tympanosquamous suture line. An inferior incision is made along the tympanomastoid suture line. These incisions outline the so-called vascular strip, the thickest and most vascular portion of the ear canal skin. A medial incision is made parallel to the annulus, connecting the two longitudinal incisions. These incisions may be made through the ear canal at the beginning of the procedure. However, they are positioned more reliably when they are fashioned from behind after the posterior canal skin has been elevated through the postauricular incision. Some surgeons elect not to divide the vascular strip from the squamous cell epithelium of the TM, preferring instead to use this connection to simplify elevation of squamous cell epithelium from the fibrous remnant.

The postauricular incision is carried down to the level of the temporalis fascia superiorly and the mastoid cortex inferiorly. If a TM graft is planned, temporalis fascia or loose connective tissue lateral to the fascia can be harvested and set aside to dry (Figure 10–6). To ensure adequate exposure, the soft tissue elevation continues to the zygomatic root, exposing the superior portion of the bony canal. The ear canal skin is elevated and the vascular strip is incised, exposing the lumen of the bony EAC and the TM defect (Figure 10–7). Although the vascular strip approach provides excellent exposure of the middle ear, it does have certain drawbacks. The approach requires a postauricular incision and thus increases postoperative discomfort. Proper elevation and replacement of the vascular strip can be technically challenging, and there may be some residual areas of exposed bone after it is replaced. As a result, healing may be delayed and problematic.

Regardless of the approach selected, 1 to 2 mm of tissue at the margin of the TM perforation should be excised (Figure 10–8). This disrupts the mucocutaneous junction at the border of the perforation and starts the wound healing process essential to successful tympanoplasty. It is also essential to remove any squamous cell epithelium that might have extended under the edge of the perforation onto the medial surface of the TM. Retained squamous elements can later develop into a cholesteatoma, requiring repeat surgical intervention. In addition, any portions of the TM that are

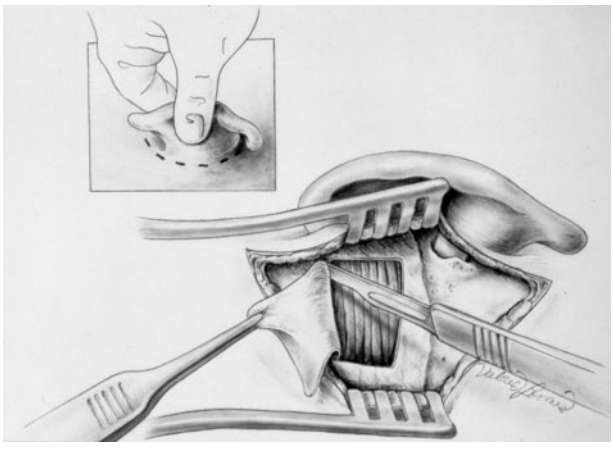


FIGURE 10–6. A right ear is pictured. The postauricular incision has been made, exposing the temporal bone and external meatus. A temporalis fascia graft is being harvested. Reproduced with permission from Telian SA, Kemink JL. Lateral technique tympanoplasty. Oper Tech Otolaryngol Head Neck Surg 1992;3:214–9.

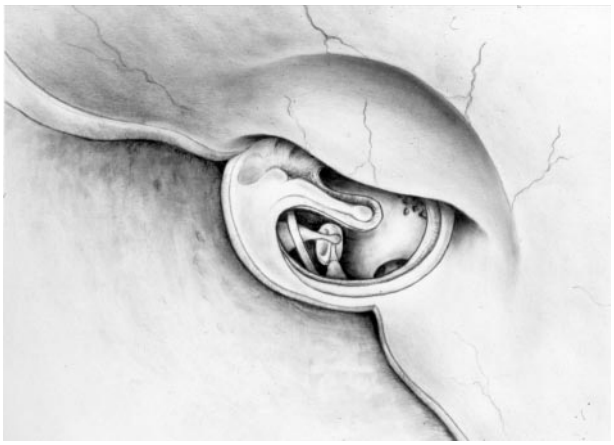


FIGURE 10–7. After the pinna and vascular strip are turned forward, the posterior wall of the bony external auditory canal and the perforation are visualized. A prominent bulge of the bony canal corresponding to the temporomandibular joint usually obscures the anterior portion of the tympanic membrane and middle ear. Reproduced with permission from Telian SA, Kemink JL. Lateral technique tympanoplasty. Oper Tech Otolaryngol Head Neck Surg 1992;3:214–9.

severely atrophic or calcified with tympanosclerosis should be excised.

Removal of overhanging areas of the tympanosquamous or tympanomastoid suture may be necessary to improve exposure for tympanoplasty. To visualize the anterior sulcus better and ensure proper placement of TM grafts in this area, removal of the anterior canal wall bulge may also be required. This step is essential when performing lateral technique tympanoplasty but is rarely required in medial technique tympanoplasty.

Lateral Technique Tympanoplasty The lateral technique tympanoplasty requires removal of the anterior canal skin after the vascular strip is retracted forward. An anterior canal incision is made just medial to the bony-cartilaginous junction, connecting the two longitudinal flap incisions. The remaining ear canal skin and periosteum are elevated medially to the level of the annulus tympanicus (Figure 10–9). Complete de-epithelialization of the TM is imperative since this technique entails placement of the graft lateral to the fibrous layer. This will prevent formation of cholesteatoma arising from any residual squamous cell epithelium left between the graft and the fibrous remnant. Elevation of the squamous cell layer begins medially at one of the canal incisions near the annulus. The entire bony ear canal and fibrous layer of the tympanic remnant are freed of squamous cell epithelium. The anterior canal skin is trimmed of useless tags and is stored in a moist sponge. The lateral edge

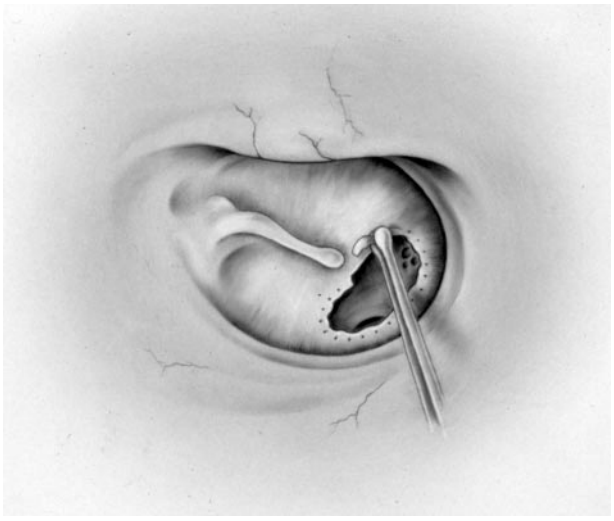


FIGURE 10–8. A fresh perforation is created by resecting the mucocutaneous junction at the margins of the perforation. This results in a slightly larger perforation and initiates wound healing along the graft.

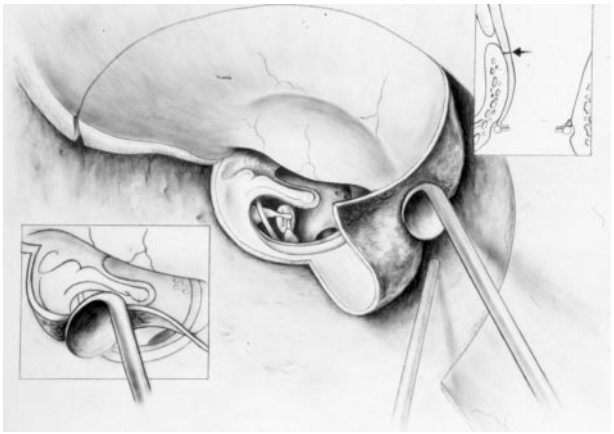


FIGURE 10–9. For a lateral technique tympanoplasty, the anterior canal skin is incised near the bony-cartilaginous junction (*upper inset*). The skin is elevated from the anterior and inferior canal wall and then from the tympanic membrane remnant (*lower inset*). Reproduced with permission from Telian SA, Kemink JL. Lateral technique tympanoplasty. Oper Tech Otolaryngol Head Neck Surg 1992;3:214–9.

of the squamous cell epithelium is marked to facilitate proper orientation when it is replaced.

The canalplasty is performed to remove the anterior canal wall bulge and allow visualization of the anterior sulcus. The goal is to change the acute angle formed by the TM and anterior bony canal to approximately a 90-degree angle, with the entire annulus easily visible through the canal. Prior to drilling down this bony ledge, the skin overlying the anterior canal wall is either removed entirely or reflected medially. Drilling begins in the superior and inferior corners of the anterior bony canal (Figure 10–10). When drilling in the medial portion of the superior corner, care should be taken to avoid contact with the short process of the malleus. Otherwise, a SNHL can result from inner ear trauma.

The bony bulge is gradually taken down between the two corners in a lateral to medial direction (Figure 10–11). It is also important to remember that this bony prominence represents the posterior wall of the temporomandibular joint (TMJ). Copious irrigation will allow underlying soft tissue to become evident before the joint is violated. If the soft tissue of the joint is violated, the posterior portion of the TMJ can be eroded, allowing prolapse of the mandibular condyle into the ear canal. This feared complication of lateral technique tympanoplasty is extremely difficult to correct and should be assiduously avoided. After achieving adequate exposure, the middle ear should be copiously irrigated to remove bone dust.

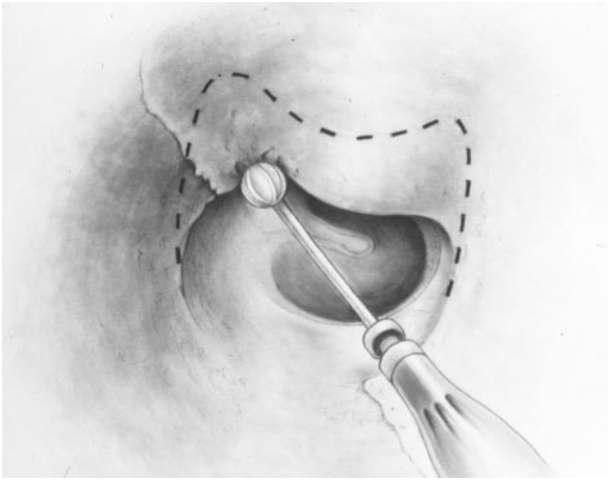


FIGURE 10–10. After all of the squamous cell epithelium is removed from the bony canal, the anterior canal wall is drilled to remove the bone as outlined. Reproduced with permission from Telian SA, Kemink JL. Lateral technique tympanoplasty. Oper Tech Otolaryngol Head Neck Surg 1992;3:214–9.

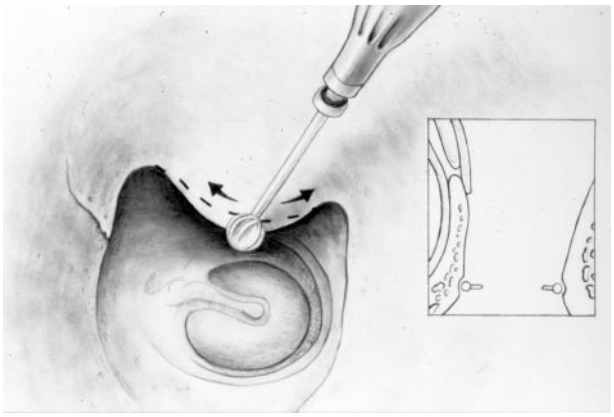


FIGURE 10–11. The superior and inferior corners are drilled initially, followed by careful polishing of the anterior bulge. A 90-degree angle at the anterior sulcus is desirable after the drilling is completed (*inset*). Reproduced with permission from Telian SA, Kemink JL. Lateral technique tympanoplasty. Oper Tech Otolaryngol Head Neck Surg 1992;3:214–9.

The annulus may be elevated from the sulcus tympanicus anteriorly in continuity with the mucosa of the lateral wall of the eustachian tube. This provides a nice pocket between the annulus and the bony wall of the eustachian tube to support the anterior portion of the graft. If there was a significant degree of mucosal resection performed in the middle ear, the surgeon may wish to place a sheet of gelatin film or thin Silastic sheeting over the promontory to prevent adhesions. Gelfoam can be placed in the middle ear to prevent retraction, but this is not mandatory with the lateral technique, particularly if there is a substantial remnant of the fibrous layer to support the graft peripherally. When a substantial remnant does remain, incisions should be made sharply alongside the malleus handle to permit insinuation of the connective tissue graft.

The dried tissue graft is trimmed so that it is long enough to drape onto the posterior canal wall and wide enough to cover the entire TM region and a portion of the superior canal wall (Figure 10–12). A superior slit is fashioned to allow the surgeon to wrap the graft around the neck of the malleus. The reconstruction is accomplished by placing the graft into the TM defect, medial to the malleus handle and

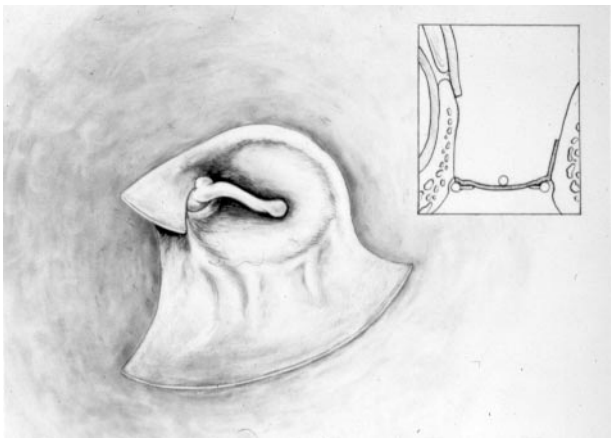


FIGURE 10–12. The fascia graft is placed medial to the malleus handle and lateral to the tympanic membrane remnant. A portion of the anterior fascia is draped over the region of the pars flaccida. *Inset:* cross-sectional view of graft position. Reproduced with permission from Telian SA, Kemink JL. Lateral technique tympanoplasty. Oper Tech Otolaryngol Head Neck Surg 1992;3:214–9.

lateral to the fibrous layer (see Figure 10–12, inset). It is best to position the graft medial to the malleus before it becomes floppy from rehydration. The tissue anterior to the slit is delivered onto the superior canal wall, and the base of the slit is pulled snugly onto the malleus neck. The region of the notch of Rivinus is covered by draping the tissue anterior to the slit onto the tissue that is posterior to the slit. The graft is then rotated around the axis of the malleus to bring the inferior and anterior edges into proper position. The anterior edge is positioned lateral to the fibrous annulus or tucked into the pocket created lateral to the eustachian tube orifice. Care must be taken to ensure that the graft does not extend onto the anterior bony canal wall because this may lead to anterior blunting or lateralization of the graft.

The anterior canal skin is retrieved from the moist sponge and returned to the ear canal, where it is shifted a couple of millimeters medially from its original anatomic position. It should just overlap the anterior portion of the previously placed connective tissue graft. The skin graft should be carefully positioned into the right angle formed by the bony canal and the TM graft. It is then supported by a rolled piece of pressed dry Gelfoam that is tucked firmly into the anterior sulcus and allowed to expand (Figure 10–13). Additional Gelfoam is placed over the tissue grafts, except for that portion of the fascia that extends onto the posterior ear canal wall. The pinna and vascular strip are then returned to the anatomic position. The vascular strip should overlie the portion of the fascia graft resting on the posterior canal wall. The remainder of the canal is packed with Gelfoam and a cotton ball is placed into the lateral third of the ear canal. The postauricular wound is closed in layers, taking care not to lift the vascular strip from the ear canal during suturing. A mastoid dressing is applied.

The lateral technique tympanoplasty offers the experienced practitioner the advantage of reliable (better than 95% in most published series) repair of large and anteriorly located TM perforations. However, the results of the technique in inexperienced hands are considerably poorer.[65] The primary disadvantages include increased technical demands of the surgery, longer operative time, and sometimes delayed healing of the ear canal. In addition to the risk of epithelial cyst formation or cholesteatoma from retained squamous cell epithelium mentioned earlier, there is also the possibility that the patient could develop a conductive hearing loss owing to postoperative blunting or lateralization of the TM graft, especially if certain key technical principles are neglected or misunderstood.[66]

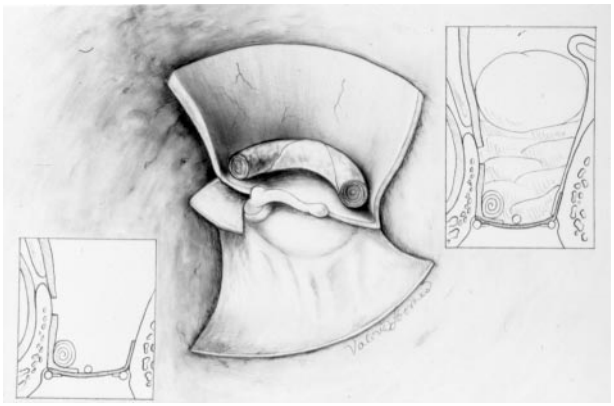


FIGURE 10–13. The canal skin graft is placed onto the fascia graft anteriorly and draped onto the canal wall. The anterior sulcus must be tightly packed, maintaining the 90-degree angle to prevent blunting (*lower inset*). After returning the vascular strip to the anatomic position, it should overlie the fascia posteriorly. The remainder of the canal is packed with Gelfoam and a cotton ball is placed in the lateral third of the canal (*upper inset*). Reproduced with permission from Telian SA, Kemink JL. Lateral technique tympanoplasty. Oper Tech Otolaryngol Head Neck Surg 1992;3:214–9.

Medial Technique Tympanoplasty As a result of the potential disadvantages discussed above, many otologic surgeons prefer the medial technique tympanoplasty. When the connective tissue graft is placed medial to the TM remnant, the risk of retained squamous cell epithelium beneath the graft and the subsequent development of cholesteatoma is essentially eliminated.

A medial technique tympanoplasty begins by elevating a tympanomeatal flap through a transcanal or postauricular approach, exposing the middle ear (see Figures 10–7 and 10–8). The drum may be completely elevated off the malleus handle, depending on the exposure required. The surgeon must take care to remove all squamous cell epithelial ingrowth on the medial surface of the TM remnant and malleus handle. Any irreversible mucosal disease is then removed from the middle ear. Unlike the lateral technique tympanoplasty, the fibrous layer of the TM has been elevated and will not provide support for the connective tissue graft. Therefore, the middle ear and eustachian tube orifice are filled with Gelfoam to prevent medial displacement of the graft. The connective tissue graft is seated medial to the handle of the malleus and TM remnant (Figure 10–14). The tympanomeatal flap is returned to its anatomic position, and the ear canal is filled with Gelfoam and a cotton ball.

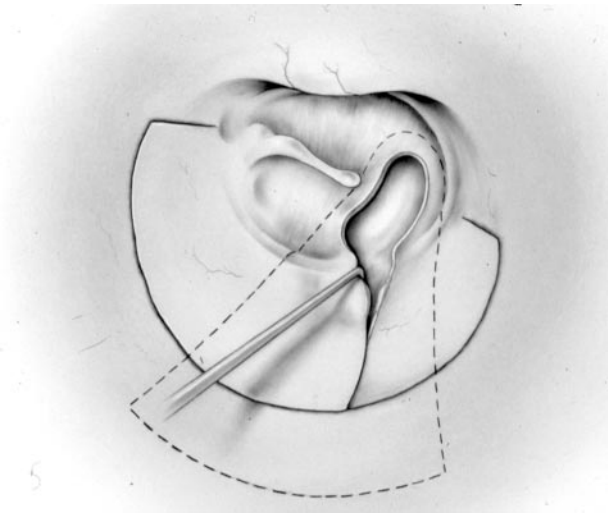


FIGURE 10–14. For a medial technique tympanoplasty, the fascia graft is placed medial to the tympanic membrane remnant and is supported by packing. After proper positioning of the graft, the divided limbs of the tympanomeatal flap are replaced into their original anatomic position to secure the graft.

Atelectatic Tympanic Membrane Nonperforated TMs as seen in atelectatic and adhesive otitis media may also require tympanoplasty if there is significant conductive hearing loss, retraction pockets with early cholesteatoma formation, or persistent infection. The atrophic, scarred TM is excised with preservation of the tympanic annulus. A tragal cartilage graft may be used to support the connective tissue graft in an attempt to prevent recurrence. Gelfilm (Upjohn, Kalamazoo, Michigan), or Silastic sheeting may be placed in an attempt to maintain ventilation of the middle ear cleft.

Postoperative Care Patients should refrain from strenuous exercise, air travel, and vigorous nose blowing during the immediate postoperative period to help prevent graft dislocation. The patient is examined approximately 1 to 2 weeks following the surgery, at which point the cotton ball is removed from the lateral canal and superficial Gelfoam is suctioned. Corticosteroid-antibiotic eardrops are prescribed for use two times a day. The patient should be examined every 2 weeks until healing is complete. All Gelfoam should be removed by the second or third postoperative visit. It is not unusual to require 8 to 12 weeks for complete healing, and crusting is common for the first few months. The ear should be observed periodically for the development of delayed healing problems, lateralization of the graft, or severe anterior blunting, which may require repeat surgery. If a pearl of cholesteatoma develops between the TM remnant and the graft, it will become evident on otoscopic examination. These cholesteatoma pearls are generally only a minor nuisance if detected early and can often be treated by simple marsupialization and extraction in the office without anesthesia.

Summary Techniques for tympanoplasty offer various options with respect to incision, graft material, and graft placement. Ultimately, the decision will depend on the location of the perforation, the bony anatomy of the EAC, and the surgeon's experience. Regardless of the method employed, experienced surgeons may expect that approximately 90% of tympanic perforations will be successfully repaired and 80% of the patients will have a con-

ductive hearing loss measuring 10 dB or less if the ossicular chain was intact at the time of surgery.[67]

Ossicular Chain Reconstruction

The numerous techniques and middle ear prostheses available to the otologic surgeon lend credence to the idea that ossicular chain reconstruction techniques remain to be perfected. However, in experienced hands, reconstruction should achieve closure of the air–bone gap to within 20 dB in two-thirds of the patients with an intact stapes arch and in one half of the patients missing the stapes superstructure.[68,69]

Some surgeons have been hesitant to perform ossiculoplasty in children for fear of failure associated with persistent eustachian tube dysfunction, higher infection rates, and increased risk of cholesteatoma. However, studies have demonstrated that hearing results in children can mirror those reported in adults.[70,71]

Autografts Autograft ossicles are removed from the patient and sculpted to serve as interposition grafts. The incus is used most often. Immediate availability, obvious biocompatibility, and a low extrusion rate have made autograft ossiculoplasty very popular. However, extensive bone erosion owing to middle ear disease may limit availability. Concerns have also been expressed regarding the use of ossicles from ears with cholesteatoma because of the risk of residual disease. Although cholesteatoma may involve the ossicular surface, histologic evidence of intraossicular cholesteatoma is lacking, even in the setting of gross bone erosion.[72] Once cholesteatoma remnants are removed from the ossicle with the use of a diamond bur, the bone is suitable for grafting. Caution must be taken when using autografts because bony fixation to adjacent bone of the cochlear promontory or the canal wall can lead to failure.

Homografts Homograft ossicles and en bloc TMs with attached ossicles can be harvested from human cadavers and are available for purchase through regional tissue banks. Graft rejection is extremely rare. Although donors are screened for known transmissible diseases and tissue preservation is performed using ethyl alcohol, formaldehyde, and glutaraldehyde, which should inactivate most viruses, the fear of potential transmission of HIV, hepatitis, and Creutzfeldt-Jakob disease limits the use of homografts.

Allografts More recently, bioengineers and surgeons have joined forces in an attempt to design biocompatible prostheses suitable for ossicular reconstruction. Allografts offer the advantage of sterility, availability, and, in some cases, tissue bonding to the ossicular chain or TM. However, these advantages come with a significant price tag compared with autograft ossicles.

Bioinert implants are so named because a mechanical bond, often consisting of fibrous tissue, forms between the implant and surrounding bone. Metals such as stainless steel and titanium fall within this category, as do solid polymers such as polytetrafluoroethylene (Teflon) and polydimethylsiloxane (Silastic). Owing to high extrusion rates, the use of metal and solid polymers has been largely abandoned.

In an effort to increase stability, porous polymers have been developed to allow ingrowth of human tissue into the prosthesis. Proplast 1 (Teflon-vitreous carbon composite) and Proplast 2 (Teflon-aluminum oxide composite) are two such examples (Vitek, Inc., Houston, Texas). Plastipore (Richards Medical Company, Inc., Memphis, Tennessee), a high-molecular-weight polyethylene sponge designed in a columellar fashion, has gained considerable popularity. Plastipore partial ossicular replacement prostheses (PORPs) are available to bridge the gap between the TM and stapes capitulum (Figure 10–15). In the absence of the stapes arch, a total ossicular replacement prosthesis (TORP) is seated between the drumhead and footplate (Figure 10–16). Intraoperative modifications can be made to the Plastipore prosthesis with the use of a knife. An alternative to Plastipore is Polycel (Treace Medical Inc., Memphis, Tennessee), consisting of thermally fused, high-density polyethylene sponge. It is available for both total and partial ossicular replacement.

Several large series have demonstrated that hearing levels achieved with Plastipore ossicular implants are equal or superior to those achieved with incus autografts.[73,74] The most common reason for failure is extrusion, especially in the setting of advanced mucosal disease.[75] Placement of cartilage, usually harvested from the tragus, between the TM and prosthesis head has decreased the extrusion rate to as low as 7%.[68]

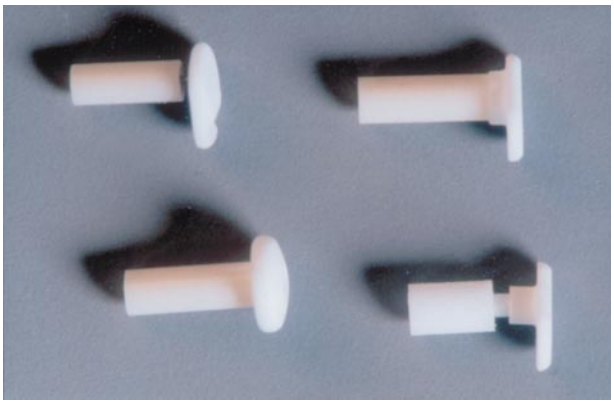


FIGURE 10–15. Photograph of several synthetic partial ossicular replacement prostheses (PORPs). Note the option of flat or rounded hydroxyapatite caps that may be placed directly under the tympanic membrane graft and alternative versions with a groove to allow secure positioning under the malleus handle.

Bioactive materials, which form chemical bonds allowing ion exchange with surrounding bone, have also been implemented in ossicular chain reconstruction. Included within this category are glass ceramics such as Cervital and Bioglass. The most popular bioactive material is hydroxyapatite (HA), a calcium phosphate ceramic resembling the mineral matrix of human bone. Although this substance is highly biocompatible, it is also quite brittle and therefore difficult to modify intraoperatively without risking a fracture of the implant. Thus, HA is generally used in the cap of the prosthesis, and the shafts are often made of a different substance that is easier to cut and perhaps even malleable. HAPEX, a homogeneous composite of HA and reinforced high-density polyethylene mixture, is also available (Richards Medical Company Inc.) and is much easier to cut than pure HA. Most prostheses are columellar in shape, with the exception of the Wehrs incus replacement and incus-stapes prosthesis.

FIGURE 10–16. Photograph of two synthetic total ossicular replacement prostheses (TORPs). The shafts may be made of rigid or malleable materials. The option exists to obtain smooth caps to place against the tympanic membrane graft or caps with a groove to engage the malleus handle.

A long-term study of HA prostheses demonstrated an extrusion rate of only 1.2%.[69] For this reason, it is argued that bioactive materials are better tolerated in the middle ear compared with their bioinert counterparts. Because of this low extrusion rate, many of the newer PORPs and TORPs in use today have a HA head, which is seated against the TM and therefore does not require the interposition of tragal cartilage. Nevertheless, some surgeons prefer to use cartilage in every case to minimize the long-term risk of extrusion.

Reconstructive Technique The technique chosen for ossicular reconstruction will ultimately depend on the remaining ossicles, with the two most important structures being the stapes superstructure and the malleus handle. The most commonly encountered ossicular defect is erosion of the long process of the incus (36%), followed by the stapes superstructure (18%) and the entire incus (17%).[10] Erosion of the incus long process should be expected with TM retraction onto the bone. In these cases, the incus should be inspected closely for erosion, and the integrity of the ossicle can be tested by gently palpating the long process with a pick. If there is a fibrous union at the incudostapedial joint instead of a healthy connection, palpation of the incus will not produce the usual motion of the stapes.

In COM, it is important to realize that conductive hearing loss may result from factors other than ossicular erosion by the disease process. For example, with cholesteatoma, removal of the incus and malleus head may be necessary to remove disease from the attic. Sometimes tympanosclerosis of the oval window or osteogenesis surrounding the stapes footplate may occur secondary to chronic inflammation. Such cases will require either a total stapedectomy with oval window grafting or removal of the stapes superstructure with stapedotomy to accomplish ossicular reconstruction. In addition,

tympanosclerosis of the malleus head and incus may cause fixation of the ossicles in the epitympanum, requiring mobilization or extraction of the ossicles.

If only the tip of the long process of the incus is absent, ossicular continuity can be restored using bone cement or a biocompatible incudostapedial joint prosthesis that fits onto the remaining long process. If complete erosion of the long process is discovered, the incus body can be extracted, sculpted, and used as an autograft to bridge the gap between the malleus handle and the stapes capitulum. A diamond bur is used to drill an acetabulum that fits over the stapes head and a groove to secure the autograft to the handle of the malleus. Figure 10–17 demonstrates two common ways in which the incus may be sculpted and placed into position. The method chosen depends on the relationship between the stapes and the malleus handle. Biocompatible prostheses are also available in similar configurations (Figure 10–18).

Loss of the incus long process along with the stapes superstructure is quite common in COM. Depending on the status of the residual incus, an interposition autograft from drumhead or malleus to stapes footplate can be used. The long process or short process of the incus is seated medially on the stapes footplate, depending on the height required to reach the TM. It is difficult to stabilize such a graft on the footplate, and the results are less reliable than those obtained when connecting to the stapes head. Alternatively, a TORP or various other HA incus stapes replacement prostheses are available.

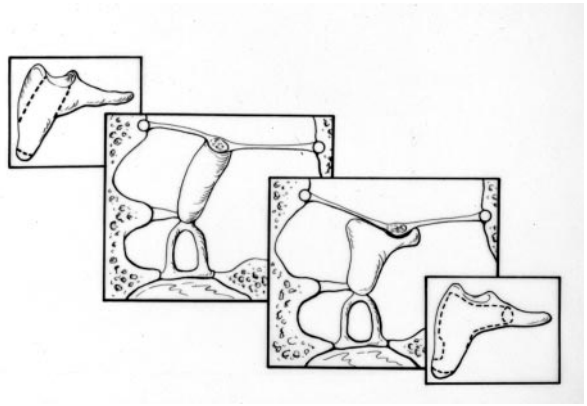


FIGURE 10–17. If the anatomic relationships are favorable, the incus may be sculpted to connect the malleus handle directly to the stapes capitulum. More commonly, the stapes is eccentrically placed posterior to the malleus handle, requiring an interposition graft that hooks onto the malleus and bridges the gap between the tympanic membrane and the stapes head.

FIGURE 10–18. Synthetic hydroxyapatite incus replacement strut.

If both the malleus handle and incus are absent when undertaking reconstruction of the ossicular chain, the use of a biocompatible PORP or TORP will be required, depending on the status of the stapes. With the probable exception of bioactive materials such as HA, any prosthesis seated directly against the TM will require a cartilage graft between the prosthesis cap and the drumhead to avoid extrusion.

Cortical Mastoidectomy

Tympanoplasty failures often occur with eustachian tube dysfunction and persistent inflammatory disease. For this reason, a cortical mastoidectomy is often recommended as an adjuvant to tympanoplasty.[76,77] The goal is to eliminate all irreversible mucosal disease, improve mastoid ventilation, and increase the buffering action of the mastoid cavity by enlarging its volume. In accordance with Boyle's law, significant changes in middle ear pressure may be better tolerated with minimal effect on the position of the TM.

A cortical mastoidectomy is usually performed through a postauricular incision. On children under

2 years of age, the mastoid tip has not fully developed, and care must be taken to avoid injury to the superficially positioned facial nerve. The incision is carried down to the mastoid cortex, and soft tissue is elevated to expose the surface landmarks of Macewen's triangle. The superior boundary is the temporal line extending posteriorly from the zygoma. This line corresponds to the probable level of the tegmen mastoideum, the bony plate separating the middle cranial fossa from the mastoid air system. The posterior boundary is a vertical line tangential to the posterior bony canal wall. Macewen's triangle is also bound by the bony projection of the posterosuperior external auditory meatus known as the spine of Henle. Contained within these boundaries is a collection of bony perforations know as the cribriform area. The mastoid antrum is directly deep to this area.

After identifying the surface landmarks, the cortex of the mastoid is opened using a large cutting bur. Dissection should always proceed from a lateral to medial direction, beveling all bony overhangs to allow adequate exposure of the cavity. The tegmen tympani and sigmoid sinus must be identified and clearly defined to avoid injury. In doing so, the sinodural angle is opened. The petrosquamous septum, also known as Körner's septum, is removed and all mastoid air cells anteromedial to the sigmoid sinus are opened (Figure 10–19). At this point, the mastoid antrum, the short process of the incus, and the horizontal semicircular canal are identified. Dissection of the mastoid tip air cells may be performed not only to increase the cavity volume but also to remove all mucosal disease that could potentially lead to future infection. On opening the mastoid tip, the digastric ridge is defined medially. On occasion, dissection of the zygomatic root is necessary to achieve adequate exposure of the epitympanum.

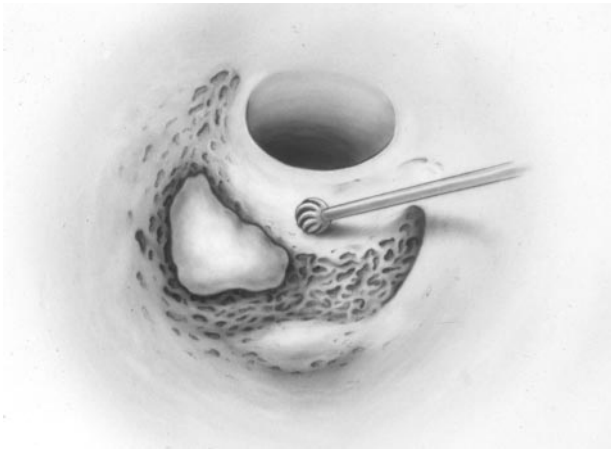


FIGURE 10–19. The cortical mastoidectomy is completed. Disease in the mastoid antrum has been exposed, with wide beveling of the cortical margins and skeletonization of the tegmen and sigmoid sinus plates.

It is important to realize that in the setting of chronic inflammation, the temporal bone tends to be sclerotic, limiting easy access to some of the important landmarks that often serve to guide the surgeon. It is vital to identify the tegmen and follow it forward to the antrum if injury is to be avoided. Liberal use of irrigation is essential, and final polishing of the tegmen plate with a diamond bur reduces bleeding.

Intact Canal Wall Tympanoplasty with Mastoidectomy

In an attempt to expose and eradicate middle ear disease better while preserving normal anatomic relationships for improved sound conduction, the posterior tympanotomy approach was introduced.[78] This technique allows access from a cortical mastoidectomy defect into the posterior mesotympanum by removal of the bony wall between the fossa incudis, the second genu of the facial nerve, and the chorda tympani. Since the access point into the middle ear is the facial recess of the posterior part of the tympanum, this operation is often referred to as a "facial recess approach." Alternative names for this approach include closed cavity tympanoplasty with mastoidectomy, intact canal wall tympanoplasty with mastoidectomy, and combined-approach tympanoplasty.

The major advantage of the intact canal wall tympanoplasty with mastoidectomy is the avoidance of a mastoid bowl that requires lifelong cleaning. In addition, healing is more rapid, and water exposure precautions are unnecessary. By avoiding a mastoid cavity and the associated meatoplasty, patients can usually wear hearing aids within the canal. However, these advantages come with a higher risk of residual and recurrent disease. The overall rate of recidivism is reported to range between 5 and 71% depending on the surgeon's experience, although most series report rates in the vicinity of 25%.[79,80] Preservation of the posterior canal wall limits visualization and access to the middle ear, especially to disease within the important

regions of the epitympanum, facial nerve, and oval window. As a result, residual middle ear disease may be left behind in any of these locations.

With respect to recurrent disease, an intact canal wall preserves a large cavity that must be ventilated by a marginally functional eustachian tube. If negative pressure persists, retraction pockets may recur and lead to the formation of recurrent cholesteatomas. Interventions such as the placement of Silastic sheeting within the middle ear may decrease the rate of postoperative adhesions and retraction pockets.[81,82] The middle ear mucosa is allowed to regenerate, and the plastic sheeting may be removed at a later date if ossicular reconstruction is undertaken at a second-stage procedure. In some cases, the presence of Silastic may be counterproductive by leading to intense middle ear fibrosis, further compromising ventilation. Reconstruction of canal wall defects that result from bone erosion by disease or the surgical approach itself has decreased the incidence of retraction pockets and associated recurrent disease.[83]

Because an intact canal wall tympanoplasty with mastoidectomy does not address the problem of negative pressure in the middle ear, ideal candidates are individuals with large pneumatized mastoids and well-aerated middle ear clefts. Patients with small sclerotic mastoids and poor eustachian tube function are often better served by removing the posterior canal wall as described in the next section.

As a result of the high rate of recidivism associated with an intact canal wall tympanoplasty with mastoidectomy procedures, most surgeons advocate a second-stage procedure when treating cholesteatoma. This is especially true for children since cholesteatomas in this population seem to behave more aggressively.[78,82,84] This second procedure usually takes place between 6 and 18 months postoperatively. During the operation, the middle ear and mastoid are explored for residual disease. If the middle ear is healthy, ossicular chain reconstruction is completed at that time. In the future, the use of endoscopes may replace the need for open second-look procedures in select cases.[85]

Intact canal wall tympanoplasty with mastoidectomy can be a technically challenging procedure. A postauricular approach is used, and the order of middle ear exploration and mastoidectomy remains the surgeon's decision. On completion of the cortical mastoidectomy, attention is turned to the facial recess, a variably pneumatized triangle bounded laterally by the TM, medially by the second genu and descending portion of the facial nerve, superiorly by the bone of the fossa incudis, and inferiorly by the chorda tympani. The bony EAC must be thinned to allow access for the facial recess dissection and eventual visualization of the most posterior mesotympanic structures, such as the round window and the stapedius tendon. Dissection should parallel the path of the facial nerve and chorda tympani nerve. Copious irrigation will reduce the risk of thermal injury to these nerves and will allow identification through intact thin bone of both the white neural sheath and associated capillary vessels. This process is known as "skeletonizing" the nerve and should be accomplished without exposing the nerve sheath. The facial nerve may be skeletonized and followed medially along the floor of the facial recess into the middle ear space.

Successful completion of the facial recess approach will allow visualization of the long process of the incus, the incudostapedial joint, the stapes superstructure, the distal tympanic segment of the facial nerve, the round window niche, and possibly the eustachian tube orifice and the cochleariform process. However, the oval window and stapes footplate are difficult to assess owing to the angle of visualization and the position of the facial nerve. The anterior epitympanum and sinus tympani are impossible to assess through the facial recess, and residual disease can be left in these areas unless visualized and removed through the ear canal. Often in the presence of cholesteatoma, the incus is eroded or requires removal. In this setting, the epitympanic dissection can be extended through the fossa incudus into the facial recess, improving visualization of the mesotympanic structures (Figure 10–20).

After removing all disease from the mastoid and middle ear, Gelfilm or Silastic sheeting may be placed in the middle ear space if extensive mucosal dissection was required. This is placed in an attempt to prevent adhesions between the TM graft and any exposed bone of the cochlear promontory, facial recess, and epitympanum. The TM graft is then placed. If there is a scutum defect owing to erosion of the superior ear canal bone, this may be reconstructed using a HA attic reconstruction plate or an autogenous cartilage graft. Tragal cartilage may be harvested with attached perichondrium and trimmed to the size of the scutum defect (Figure

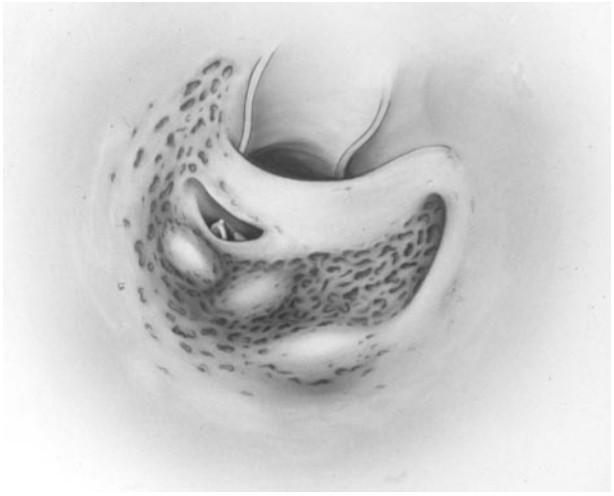


FIGURE 10–20. The facial recess has been dissected and now communicates with the epitympanum after removal of the incus. This allows visualization of the facial nerve, stapes, and other middle ear structures, facilitating dissection of disease using an intact canal wall approach.

10–21). The attached perichondrium may be used to repair a superior TM defect and/or may be draped over the remaining EAC bone to help secure the position of the cartilage in the defect. Ossicular remnants removed during exploration may be banked for possible use in ossicular chain reconstruction during a second-stage procedure. Closure and postoperative care are identical to that of a tympanoplasty, although special care should be taken to close the mastoid periosteal layer specifically to avoid a deep depression of the postauricular scar into the mastoid defect.

It is important to realize that although an intact canal wall tympanoplasty with mastoidectomy avoids the demands for routine cleaning of a mastoid cavity, close follow-up is nevertheless imperative. Recurrent cholesteatoma formation has been identified as late as 15 years after intact canal wall procedures.[86] For this reason, microscopic examination of the ear should be performed on a regular basis. With preservation of the posterior canal wall, the mastoid cavity and epitympanum will not be accessible for inspection. Thus, a potential disadvantage of the intact canal wall tympanoplasty with mastoidectomy is that early recurrent disease may go undiscovered. Radiographic imaging of the temporal bone region with a CT scan is helpful but not always diagnostic if there is concern for recurrent cholesteatoma (Figure 10–22).

Modified Radical Mastoidectomy

Gustave Bondy introduced the modified radical mastoidectomy, now often called the Bondy mastoidectomy, as a treatment for epitympanic cholesteatoma.[87] This procedure requires an open cavity mastoidectomy with removal of the posterior canal wall to create a common cavity between the mastoid and EAC. The middle ear space is not entered, and a tympanoplasty is not performed. The indications for a Bondy mastoidectomy are described above.

Today, the classic modified radical mastoidectomy is used infrequently since it addresses only the rare instance when one is treating isolated atticoantral cholesteatoma with disease lateral and posterior to the ossicles. Modifications have been made to the original approach to explore the middle ear and correct the conductive hearing loss, which often results in the setting of cholesteatoma. This combination of the open mastoidectomy and tympanoplasty with or without ossicular chain reconstruction is what most surgeons mean today when they use the term "modified radical

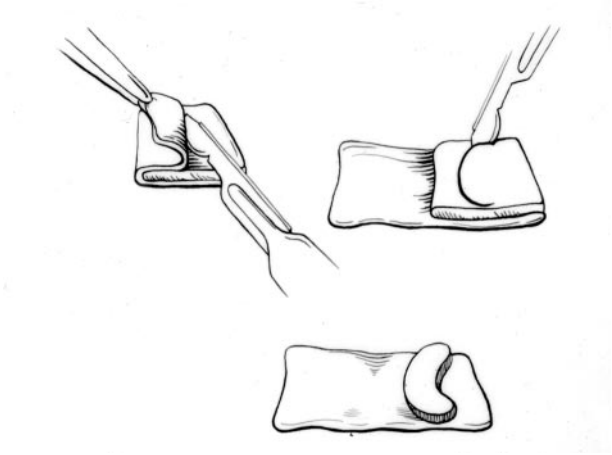


FIGURE 10–21. The tragal cartilage is harvested through a small incision behind the dome of the tragus, with its perichondrium attached on both sides. The perichondrium is elevated from one side of the cartilage, leaving the other side attached. The cartilage is trimmed to the size of the scutum defect, as pictured. After the cartilage is placed into the defect with the perichondrial side facing outward, the short extension of perichondrium is placed medial to the tympanic membrane remnant. The longer tail is draped onto the remaining ear canal bone and covered by the vascular strip.

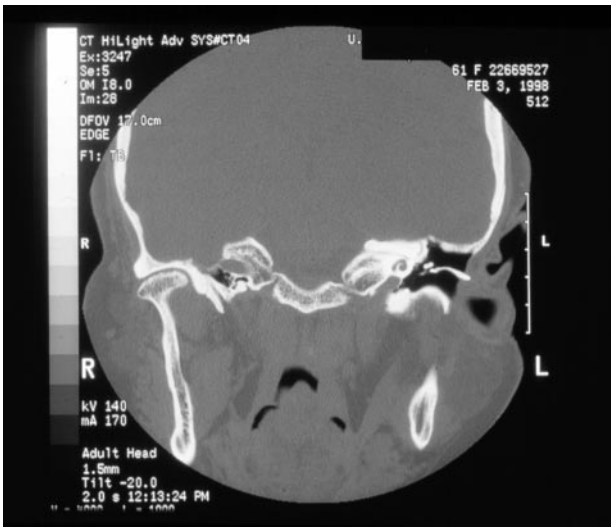


FIGURE 10–22. Coronal computed tomographic scan years after intact canal wall mastoidectomy shows the preserved bony canal wall and a well-aerated middle ear and mastoid cavity, free of any soft tissue density that might suggest recurrent disease. Such findings may eliminate the need for a second-stage procedure.

mastoidectomy." An alternate, less ambiguous term for this operation is "tympanoplasty with canal wall down mastoidectomy." Since the posterior canal wall is removed, it may also be called a canal wall down tympanoplasty with mastoidectomy or open-cavity technique.

The major advantage of a modified radical mastoidectomy is excellent exposure during dissection of cholesteatoma, which reduces the rate of residual or recurrent disease, usually precludes the need for a second-stage operation, and is therefore cost effective. Postoperative examination is ideal because only that disease hidden in the mesotympanum might go unnoticed. With removal of the canal wall, all other areas are exteriorized and can be inspected during follow-up.

The major disadvantage of this approach is the need for periodic mastoid bowl cleaning. To achieve adequate aural hygiene, a meatoplasty is performed at the time of the original surgery. This alteration in anatomy may lead to difficulty in the use of hearing aids within the meatus and precludes in-the-canal fittings. Water exposure restrictions are required in most patients to prevent infection as well as vertigo. Cold wind may also pass easily through the enlarged external auditory meatus and can produce dizziness. At one time, it was argued that hearing results following the open cavity technique were inferior to those of the closed cavity technique. However, with the advances in middle ear reconstruction, removal of the posterior canal wall appears to make little difference in the hearing outcome.[88,89]

One of the most passionate and long-standing debates within the field of otology surrounds the management of the posterior canal wall when performing a tympanoplasty with mastoidectomy. Ultimately, the decision will depend on the nature of the disease, the patient's anatomy, and the surgeon's training and experience. Often the final decision is made intraoperatively. A small, sclerotic mastoid, a low-lying middle cranial fossa dura, and an anteriorly positioned sigmoid sinus will limit surgical exposure and often necessitate removal of the canal wall. Sometimes this is best performed by simply enlarging the attic defect and extending it posteriorly until the entire cholesteatoma is exteriorized (Figure 10–23).

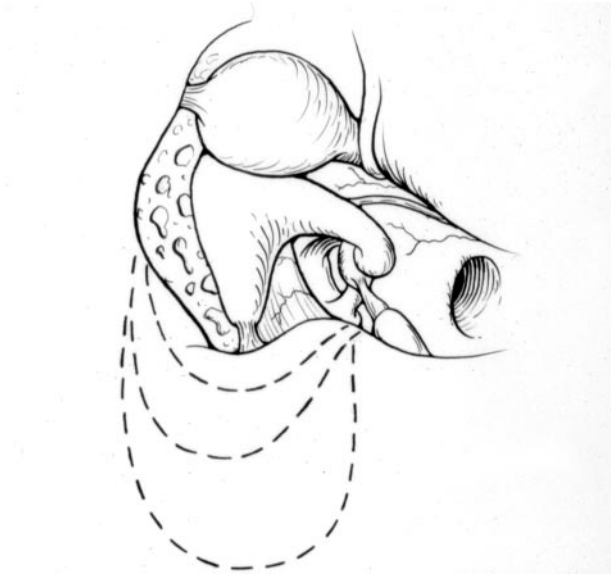


FIGURE 10–23. An atticotomy has been performed by removing the bone of the lateral epitympanum and exposing the ossicle heads, in preference to performing a formal mastoidectomy for attic cholesteatoma. The same defect may be progressively extended posteriorly (*dashed lines*) to expose disease in the antrum. In such cases, the operation is called an atticoantrotomy, or more colloquially an "inside-out mastoidectomy." This approach is most commonly used in very sclerotic mastoids but may also be used to avoid a large cavity when exteriorizing limited cholesteatoma in a well-pneumatized temporal bone.

This approach avoids the need for a more extended mastoidectomy but provides limited ability to identify surgical landmarks. Other indications to take down the canal wall include extensive cholesteatoma in the only hearing ear, presence of a large labyrinthine fistula, recurrent retraction cholesteatoma in the epitympanum, and significant destruction of the scutum or posterior canal wall. Patients who are unlikely to follow faithfully postoperative protocols, who are unwilling to undergo a second-stage procedure, or who are high risk for general anesthesia will also benefit from a canal wall down mastoidectomy.

The procedure is usually carried out through a postauricular incision. A tympanomeatal flap is elevated for middle ear exploration and eradication of disease as described in the tympanoplasty section. A cortical mastoidectomy is performed, with identification of the sigmoid sinus, antrum, short process of the incus, and horizontal semicircular canal. The digastric ridge may be followed forward to the stylomastoid foramen, identifying the height of the facial nerve as it descends from the second genu. After identification of the fallopian canal, the bony posterior canal wall is removed to the level of the facial nerve, leaving a thin layer of bone over the nerve called the facial ridge. To gain adequate exposure to the medial part of the epitympanum and the superior aspect of the mesotympanum, it is often necessary to remove the incus and the head of the malleus. The superior and posterior walls of the epitympanum are thinned, allowing for smooth communication between the tegmen mastoid and tegmen tympani. The anterior buttress of the ear canal is removed to produce a smooth transition between the anterior part of the epitympanum and the ear canal remnant (Figure 10–24).

The success of a canal wall down tympanoplasty with mastoidectomy depends on several key technical factors.[90] All air cells, especially within the sinodural angle, along the tegmen, and in the perilabyrinthine region, must be opened. The posterior canal wall must be lowered through the level of the facial recess to the descending portion of the facial nerve. In addition, the mastoid tip cells are exenterated and the bony tip is amputated, leaving the bony plate over the sigmoid sinus as the highest structure in the posterior part of the mastoid cavity. This allows for obliteration of the tip area and any cells posterior to the sigmoid sinus by postauricular scalp soft tissue. Otherwise, the patient is left with a dependent aspect of the mastoid cavity that will accumulate debris in the less accessible tip region. Ideally, the mastoid cavity should be widely saucerized with smooth bony walls achieved by wide beveling of the cavity during drilling.

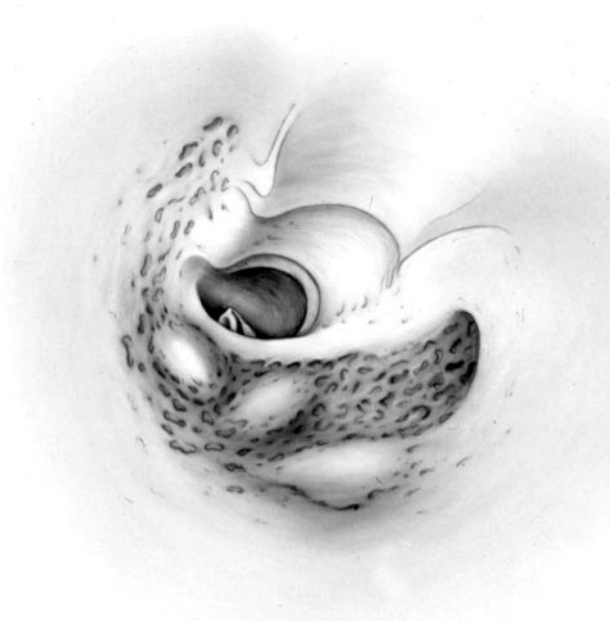


FIGURE 10–24. After removal of the posterior bony ear canal wall and smoothing of the facial ridge, an open cavity mastoidectomy has been completed. This allows maximum access to the middle ear structures for either removal or complete exteriorization of the cholesteatoma matrix.

Lastly, a successful canal wall down tympanoplasty with mastoidectomy requires an adequately sized meatoplasty. Enlargement of the meatus allows not only adequate visualization and cleaning of the mastoid bowl but also sufficient ventilation to maintain a dry cavity. The size of the opening should correlate with the size of the mastoid cavity. Ideally, the meatus should accommodate the surgeon's finger during construction. To create an adequate meatoplasty, it is essential to resect a significant portion of the conchal cartilage and underlying fibroadipose tissue while retaining the overlying skin to resurface a portion of the defect. A superior ear canal incision is made and carried to the facial skin by extending the incision through the incisura between the cartilage of the tragus and the helical root. An inferior canal incision is carried into the conchal bowl. The conchal skin and ear canal

skin are elevated, and a crescent-shaped portion of the conchal cartilage is excised. The underlying soft tissues are excised, and hemostasis is obtained (Figure 10–25).

After removal of the posterior canal wall, the grafted TM cannot be returned to its original anatomic location at the end of the procedure. Instead, it must be draped from the anterior annulus over the facial ridge and the horizontal semicircular canal. In doing so, a shallow middle ear space is created. If the stapes remains intact and the footplate is mobile, the TM graft can be seated directly on the stapes capitulum to allow adequate sound conduction. If the facial ridge is fairly high, it may be necessary to place a sculpted ossicle head or a piece of cartilage from the meatoplasty over the stapes to improve contact with the graft. In the absence of a stapes, a sculpted incus can be interposed, or a commercially available TORP may be selected. Some of these are specifically designed with short necks to accommodate the shallow middle ear space associated with an open cavity. If the entire cholesteatoma matrix has been removed, the fascia graft should be large enough to resurface the antrum and obliterate any perilabyrinthine air cells. This will result in a smaller cavity with a smoother lining and will help prevent mucosalization of the antrum during the healing phase.

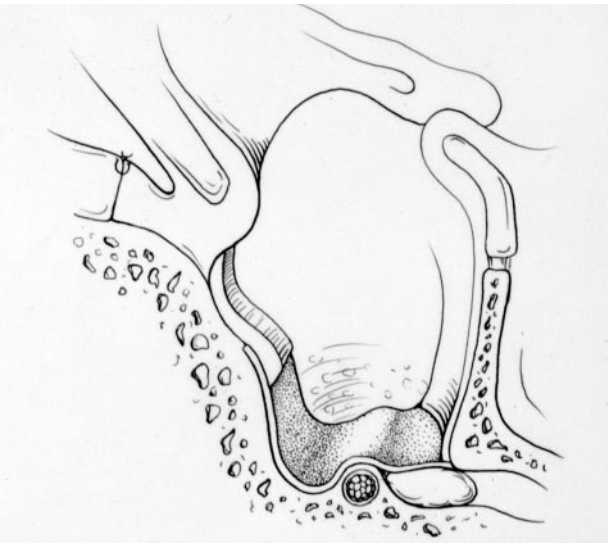


FIGURE 10–25. The successful canal wall down mastoidectomy requires a generous meatoplasty with removal of conchal cartilage and underlying soft tissues, with beveling of the cavity edges. This drawing demonstrates grafting of the middle ear cleft and draping of the graft into the posterior cavity, where it is overlaid with the vascular strip.

At the conclusion of the surgery, the meatus is tacked open by suturing the base of the skin flaps posteriorly to the mastoid periosteum. These sutures must be placed strategically to avoid distortion of the pinna's natural appearance. The periosteal layer is closed, and the conchal bowl and EAC skin flap are rotated medially into the defect. Ideally, the medial edge of the skin flap will reach onto the posterior aspect of the fascia graft used to graft the middle ear and line the antrum. The grafts are covered with a layer of Gelfoam, and the mastoid cavity is filled with antibiotic ointment. Merocel packs are placed laterally in the meatus and expanded with otic drops to prevent early collapse of the meatoplasty, which could lead to stenosis. A mastoid dressing is then applied and kept in place for 24 to 48 hours. Since the meatoplasty wound tends to have persistent oozing of blood, a light dressing may be needed subsequently for several more days. The packing and superficial Gelfoam are removed at the first postoperative visit, about 10 days after the operation. Open cavities take a considerable amount of time to heal, on the order of 6 to 10 weeks. During this period, patients are seen every 2 to 3 weeks for débridement of the cavity and management of granulation tissue. Early granulations are suctioned away, and the base is cauterized with silver nitrate or 30% trichloroacetic acid. Early neomembrane formation should be recognized and disrupted before fibroblast ingrowth causes mature scars that result in unfavorable cavity anatomy. Ototopical drops with an antibiotic and a corticosteroid should be continued until there is no granulation tissue and the cavity is lined with skin. As the healing progresses, immature squamous cell epithelium that tends to weep and form crusts may be painted with gentian violet. Patients can also participate in cavity hygiene during this later healing phase by performing daily irrigations with dilute acetic acid. Postoperative visits may become less frequent after this stage, although newly healed cavities tend to form crusts that may lead to inflammation and cavity infections. These crusts should be removed at each visit, and any underlying granulation tissue may be cauterized.

Long-term care entails visits every 6 months to 1 year, depending on the condition of the mastoid cavity. Microscopic examination and removal

of all crusts should be performed. Patients prone to accumulation of dry crust benefit from application of baby oil or mineral oil to the cavity once or twice a week. Some cavities require the ongoing use of occasional dilute vinegar irrigations, particularly if they are prone to infection from swimming, from hearing aid use, or during the humid months of summer.

Radical Mastoidectomy

Radical mastoidectomy entails exteriorization of the entire middle ear and mastoid by combining a canal wall down mastoidectomy with removal of the TM and the ossicles (with the exception of the stapes if present). In doing so, the mastoid, middle ear, and EAC become one common cavity. There is no attempt at middle ear reconstruction, and patients are left with a substantial conductive hearing loss. Given this significant functional deficit, radical mastoidectomy is considered a last resort, usually after previous surgical attempts have failed or when it is not possible to remove mesotympanic cholesteatoma. Radical mastoidectomy is also indicated if middle ear ventilation is impossible owing to complete inadequacy of eustachian tube function.

The incision, mastoidectomy, and removal of the posterior canal wall follow that described above. In addition, all remnants of the TM as well as the malleus and incus are excised. The inferior aspect of the tympanic ring is drilled down to remove disease in the hypotympanic air cells and allow communication between the cavity and the hypotympanum. The eustachian tube mucosa is inverted, and the orifice is sealed with an ossicular remnant, bone chips, muscle, or fascia in an attempt to prevent access of nasopharyngeal inflammation and secretions into the cavity. A meatoplasty is required for postsurgical care and surveillance. Postoperative care follows that of a canal wall down tympanoplasty with mastoidectomy.

Despite the extensive nature of this surgery, aural drainage may persist owing to mucosalization of the middle ear space from the eustachian tube or the hypotympanic cells. This is a particular problem if both ears are involved. In such cases, use of conventional hearing aids may be impossible, and a bone-conduction hearing aid or bone-anchored hearing aid is indicated. Other disadvantages of the radical mastoidectomy include lifelong need for mastoid cavity cleaning as well as water exposure precautions.

Mastoid Cavity Obliteration

The mastoid cavity created by a radical or modified radical mastoidectomy is at risk for chronic infection and persistent drainage. In the setting of a potentially large cavity, some surgeons elect to perform a mastoid obliteration procedure. The goal of the procedure is to provide a dry ear that is compatible with the use of a hearing aid and also to eliminate the need for annual cavity cleaning.

Various materials such as bone pate, cartilage, acrylic, and HA cement have been used to obliterate the mastoid cavity. Free abdominal fat grafts have been used; however, postoperative atrophy must be taken into consideration. Many surgeons prefer the use of regional soft tissue flaps because of the associated blood supply. Such alternatives include pedicled temporalis muscle or postauricular musculoperiosteal flaps as described by Rambo and Palva, respectively.[91,92]

The major disadvantage associated with obliteration of the mastoid is that disease can be buried within the cavity. A CT scan may be required to evaluate patients adequately for recurrent disease. Thus, contraindications to mastoid obliteration include widespread cholesteatoma and infected bone at the time of surgery. Generally, if the surgeon observes the surgical principles and postoperative care guidelines outlined above, the need for a mastoid obliteration procedure would be unusual.

Rarely, the surgeon may elect to perform a complete closure of the external canal to isolate the cavity internally rather than a meatoplasty to exteriorize it. This is most suitable when treating chronic inflammation without cholesteatoma in the setting of a profound hearing loss in the involved ear. In such a case, the need for extended healing and long-term cavity care are obviated. The modified Rambo closure is recommended to provide a cosmetically acceptable and problem-free external closure (Figure 10–26).[93] Care must be taken to remove all squamous elements from the ear deep to the point of closure. The primary and potentially serious risk of such a procedure is the development of an expanding undetectable cholesteatoma in the temporal bone. This may grow silently for many years, causing no symptoms until a serious problem such as facial

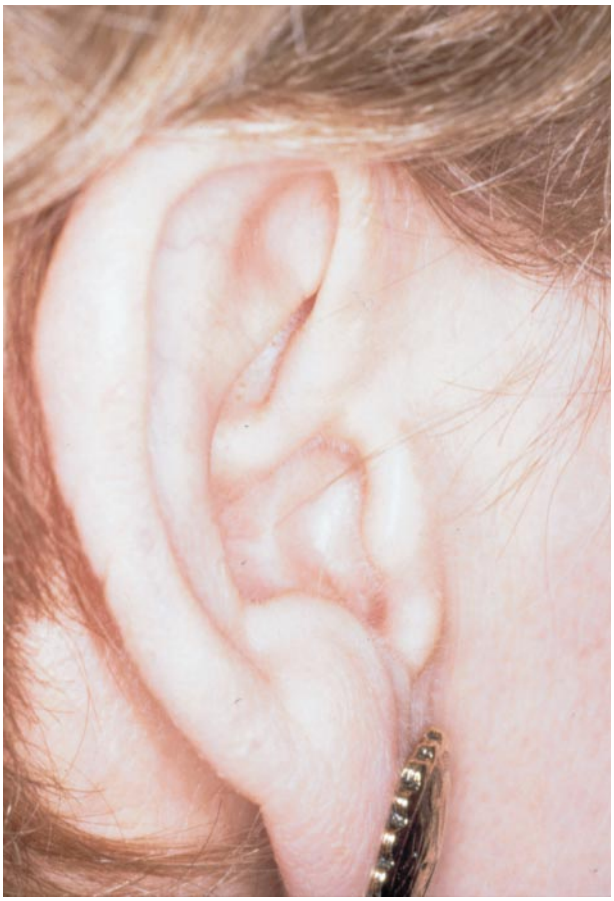


FIGURE 10–26. Photograph of an overclosed external auditory meatus after a modified Rambo technique. The smooth contours created using this method of closure eliminate any need for periodic cleaning of the meatus. In natural lighting conditions, the shadow that is cast by the tragus provides a natural appearance, creating the illusion of an external meatus.

nerve paralysis or an intracranial complication develops. Since the cavity cannot be examined otoscopically, the patient must undergo a postoperative CT scan about 2 years postoperatively. If the scan shows that the internal temporal bone defect is ventilated via the eustachian tube and is free of disease, no further imaging is required (Figure 10–27). If the defect contains soft tissue, additional interval CT scans will be required to rule out an expanding cholesteatoma.

COMPLICATIONS OF CHRONIC OTITIS MEDIA

Complications of COM range from mild hearing loss to life-threatening intracranial infections.

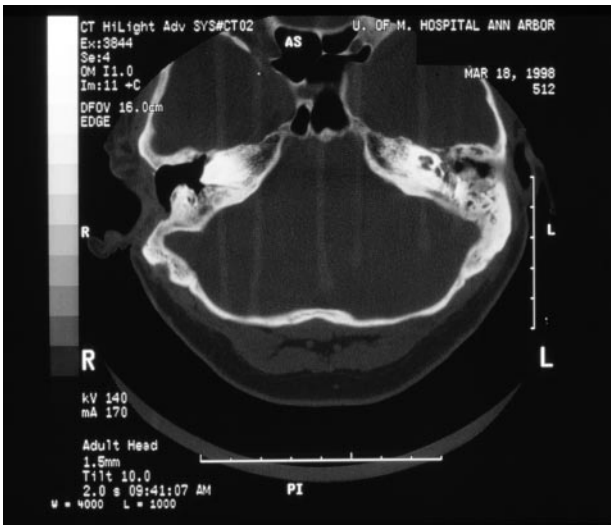


FIGURE 10–27. Axial computed tomographic scan of the patient pictured in Figure 10–26. Note the osteoneogenesis of the cochlea that led to deafness as a complication of chronic otitis media. Since hearing rehabilitation was impossible in this ear, a meatal closure was performed to eliminate the need for a mastoid cavity. The scan shows excellent aeration of the resulting cavity through the eustachian tube and no evidence of cholesteatoma or inflammatory disease.

Intratemporal complications include facial nerve paralysis, labyrinthitis, labyrinthine fistula, coalescent mastoiditis, subperiosteal abscess, postauricular fistula, and petrositis. If infection spreads beyond the confines of the temporal bone, intracranial complications such as epidural abscess, subdural abscess, lateral sinus thrombophlebitis, meningitis, and brain abscess may result. For a detailed description of the complications associated with COM and their management, see Chapter 11.

REFERENCES

1. Browning GG. The unsafeness of the 'safe' ears. J Laryngol Otol 1984;98:23–6.
2. Levenson MJ, Michaels L, Parisier SC. Congenital cholesteatomas of the middle ear in children: origin and management. Otolaryngol Clin North Am 1989;22:941–53.
3. Baxter JD. Chronic otitis media and hearing loss in the Eskimo population of Canada. Larynogoscope 1977;87:1528–42.
4. Giles M, Asher I. Prevalence and natural history of otitis media with perforation in Maori school children. J Laryngol Otol 1991;105:257–60.

5. Sade J, Halevy A. The natural history of chronic otitis media. J Laryngol Otol 1976;90:743–51.
6. Harker LA, Koontz FP. The bacteriology of cholesteatoma. In: McCabe BF, Sade J, Abramson M, editors. Cholesteatoma: first international conference. Birmingham (AL): Aesculapius; 1977. p. 264–7.
7. Kemppainen HO, Puhakka HJ, Laippala PJ, et al. Epidemiology and aetiology of middle ear cholesteatoma. Acta Otolaryngol (Stockh) 1999;119:568–72.
8. Tos M. Incidence, etiology and pathogenesis of cholesteatoma in children. Adv Otorhinolaryngol 1988;40:110–7.
9. Youngs R. Chronic suppurative otitis media—mucosal disease. In: Ludman H, Wright T, editors. Diseases of the ear. New York: Oxford University Press; 1998. p. 374–85.
10. Austin DF. Chronic otitis media. In: Ballenger JJ, Snow JB, editors. Otorhinolaryngology head and neck surgery. Baltimore: Williams and Wilkins; 1996. p. 1010–36.
11. Wright T, Ludman H. Introduction to middle ear and mastoid disease. In: Wright T, Ludman H, editors. Diseases of the ear. New York: Oxford University Press; 1998. p. 331–4.
12. Giebink GS. Otitis media update: pathogenesis and treatment. Ann Otol Rhinol Laryngol 1992;101: 21–3.
13. Yoon TH, Schachern PA, Paparella MM, Aeppli DM. Pathology and pathogenesis of tympanic membrane retraction. Am J Otol 1990;11:10–7.
14. Sade J, Berco E. Atelectasis and secretory otitis media. Ann Otol Rhinol Laryngol 1976;85:66–72.
15. Andreasson L. Correlation of tubal function and volume of mastoid and middle ear space as related to otitis media. Ann Otol Rhinol Laryngol 1976;82: 198–203.
16. McClelland CA. Incidence of complications from use of tympanostomy tubes. Arch Otolaryngol Head Neck Surg 1980;106:97–9.
17. Larsen PL, Tos M, Stangerup SE. Progression of drum pathology following secretory otitis media. In: Lim DJ, et al, editors. Recent advances in otitis media. Philadelphia: BC Decker; 1988. p. 34–8.
18. Von Schoenberg M, Wengraf CL, Bleeson M. Results of middle ear ventilation with Goode's tubes. Clin Otolaryngol 1989;14:503–8.
19. Wright CG, Meyerhoff, WL. Pathology of otitis media. Ann Otol Rhinol Laryngol 1994;103:24–7.
20. Yeger H, Minaker E, Charles D, et al. Abnormalities in cilia in the middle ear in chronic otitis media. Ann Otol Rhinol Laryngol 1988;97:186–91.
21. da Costa SS, Paparella MM, Schachern PA, et al. Temporal bone histopathology in chronically infected ears with intact and perforated tympanic membranes. Laryngoscope 1992;102:1229–36.
22. Meyerhoff WL, Kim CS, Paparella MM. Pathology of chronic otitis media. Ann Otol Rhinol Laryngol 1978;87:749–60.
23. Schuknecht HF. Pathology of the ear. Philadelphia: Lea and Febiger; 1993.
24. Jahn AF, Farkashidy J. Newer perspectives on the pathology of chronic otitis media. J Otolaryngol 1980;9:131–42.
25. Michaels L. An epidermoid formation in the developing middle ear: possible source of cholesteatoma. J Otolaryngol 1986;15:169–74.
26. Sie KC. Cholesteatoma in children. Pediatr Clin North Am 1996;43:1245–52.
27. Lee RJ, Mackenzie IC, Hall BK, Gantz BJ. The nature of the epithelium in acquired cholesteatoma. Clin Otolaryngol 1991;16:168–73.
28. Van Cauwenberge P, Vermeersch H, Otte L, Klutyskens M. Relationship between secretory otitis and cholesteatoma. Epidemiological aspect. In: Tos M, Thomsen J, Peitersen E, editors. Cholesteatoma and mastoid surgery. Amsterdam: Kugler & Ghedini; 1989. p. 319–24.
29. Sade J, Babiacki A, Pinkus G. The metaplastic and congenital origin of cholesteatoma. Acta Otolaryngol (Stockh) 1983;96:119–29.
30. Ruedi L. Pathogenesis and surgical treatment of the middle ear cholesteatoma. Acta Otolaryngol (Stockh) 1979; 361:1–45.
31. Orisek BS, Chole RA. Pressures exerted by experimental cholesteatomas. Arch Otolaryngol Head Neck Surg 1987;113:386–91.
32. Thomsen J, Bretlau P, Kristensen HK. Bone resorption in chronic otitis media. A light-microscopical and histological investigation of acid phosphatase activity. Acta Otolaryngol (Stockh) 1975;79:400–8.
33. Gantz BJ, Maynard J, Bumsted RM, et al. Bone resorption in chronic otitis media. Ann Otol Rhinol Laryngol 1979;88:693–700.
34. Kurihara A, Toshima M, Yuasa R, Takasaka T. Bone destruction mechanisms in chronic otitis media with cholesteatoma: specific production by cholesteatoma tissue in culture of bone-resorbing activity attributable to interleukin-1 alpha. Ann Otol Rhinol Laryngol 1991;100:989–98.
35. Chole RA. Osteoclasts in chronic otitis media, cholesteatoma, and otosclerosis. Ann Otol Rhinol Laryngol 1988;97:661–6.

36. Sade J, Halevy A. The aetiology of bone destruction in chronic otitis media. J Laryngol Otol 1974;88:139–43.
37. Thomsen J, Bretlau P, Balslev Jorgensen M. Bone resorption in chronic otitis media. The role of cholesteatoma, a must or an adjunct? Clin Otolaryngol 1981;6:179–86.
38. Mills P. Management of chronic suppurative otitis media. In: Kerr A, Booth JB, editors. Scott-Brown's otolaryngology. Oxford (England): Butterworth-Heinemann; 1997. 3/10/1-11.
39. Blakley BW, Kim S. Does chronic otitis media cause sensorineural hearing loss? J Otolaryngol 1998;27: 17–20.
40. Noordzij JP, Dodson EE, Ruth RA, et al. Chronic otitis media and sensorineural hearing loss: is there a clinically significant relation? Am J Otol 1995;16: 420–3.
41. Eisenman DJ, Parisier SC. Is chronic otitis media with cholesteatoma associated with neurosensory hearing loss? Am J Otol 1998;19:20–5.
42. Walby AP, Barrera A, Schuknecht HF. Cochlear pathology in chronic suppurative otitis media. Ann Otol Rhinol Laryngol 1983;103:3–19.
43. Blevins NH, Carter BL. Routine preoperative imaging in chronic ear surgery. Am J Otol 1998;19:527–35.
44. Mafee MF, Levin BC, Applebaum EL, et al. Cholesteatoma of the middle ear and mastoid. Otolaryngol Clin North Am 1988;21:265–93.
45. Koltai PJ, Earnes FA, Parnes SM, et al. Comparison of computed tomography and magnetic resonance imaging in chronic otitis media with cholesteatoma. Arch Otolaryngol Head Neck Surg 1989;115:1231–3.
46. Jackler RK, Dillon WP, Schindler RA. Computed tomography in suppurative ear disease: a correlation of surgical and radiographic findings. Laryngoscope 1984;94:746–52.
47. Campbell JP, Pillsbury HC. The use of computerized tomographic imaging in revision mastoid surgery for chronic otitis media. Am J Otol 1990;11:387–94.
48. McCaffrey TV, McDonald TJ. Histiocytosis X of the ear and temporal bone: review of 22 cases. Laryngoscope 1979;89:1735–42.
49. Paparella MM, deSousa LC, Mancini F. Meniere's syndrome and otitis media. Laryngoscope 1983;93: 1408–15.
50. Brook I. Role of anaerobic bacteria in chronic otitis media and cholesteatoma. Int J Pediatr Otorhinolaryngol 1995;31:153–7.
51. Brook I. Otitis media: microbiology and management. J Otolaryngol 1994;23:269–75.
52. Brook I. Chronic otitis media in children: microbiological studies. Am J Dis Child 1980;134:564–6.
53. Brook I. Prevalence of beta lactamase producing bacteria in chronic otitis media. Am J Dis Child 1985;139:280–3.
54. Kohonen A, Tarkanen J. Cochlea damage by ototoxic antibiotics by intratympanic application. Acta Otolaryngol (Stockh) 1969;68:90–7.
55. Brummet RE, Harris RF, Lindgren JA. Detection of ototoxicity from drugs applied topically to the middle ear space. Laryngoscope 1976;86:1177–87.
56. Fairbanks DNF. Antimicrobial therapy for chronic otitis media. Ann Otol Rhinol Laryngol 1981;90:58–62.
57. Fradis M, Brodsky A, Ben-David J, et al. Chronic otitis media treated topically with ciprofloxacin or tobramycin. Arch Otolarygol Head Neck Surg 1997;123:1057–60.
58. Tutkun A, Ozagar A, Koc A, et al. Treatment of chronic ear disease. Topical ciprofloxacin vs topical gentamicin. Arch Otolaryngol Head Neck Surg 1995;121:1414–6.
59. Browning GG, Gatehouse S, Calder IT. Medical management of active chronic otitis media: a controlled study. J Laryngol Otol 1988;102:491–5.
60. Crowler JA, Simpson D. Medical treatment of chronic otitis media: steroid or antibiotic with steroid ear-drops? Clin Otolaryngol 1991;16:142–4.
61. Gehanno P. Multicenter study of the efficacy and safety of oral ciprofloxacin in the treatment of chronic suppurative otitis media in adults. The French Study Group. Otolaryngol Head Neck Surg 1997;117:83–90.
62. Sheehy JL. Testing eustachian tube function. Ann Otol Rhinol Laryngol 1981;90:562–5.
63. Roland PS, Meyerhoff WL. Intraoperative facial nerve monitoring: what is its appropriate role? Am J Otol 1993;14:1.
64. Jackson CG. Antimicrobial prophylaxis in ear surgery. Laryngoscope 1988;98:1116–23.
65. Doyle PJ, Schluening AJ, Echevarria J. Tympanoplasty: should grafts be placed medial or lateral to the tympanic membrane? Laryngoscope 1972;82:1425–30.
66. Plester D, Pulsakar A. The anterior tympanomeatal angle: the aetiology, surgery, and avoidance of blunting and annular cholesteatoma. Clin Otolaryngol 1981;6:326–8.
67. Sheehy JL, Anderson RG. Myringoplasty: a review of 472 cases. Ann Otol Rhinol Laryngol 1980;89:331–4.
68. Brackmann DE, Sheehy JL, Luxford WM. TORPs and PORPs in tympanoplasty: a review of 1042 operations. Otolaryngol Head Neck Surg 1984;92:32–7.

69. Grote JJ. Reconstruction of the middle ear with hydroxylapatite implants: long-term results. Ann Otol Rhinol Laryngol Suppl 1990;144:12–6.
70. Kessler A, Potsic WP, Marsh RR. Total and partial ossicular replacement prostheses in children. Otolaryngol Head Neck Surg 1994;110:302–3.
71. Daniels RL, Rizer FM, Schuring AG, Lippy WL. Partial ossicular reconstruction in children: a review of 62 operations. Laryngoscope 1998;108:1674–81.
72. Rupa V, Krishnaswami H, Job A. Autograft ossicle selection in cholesteatomatous ear disease: histopathological considerations. J Laryngol Otol 1997;111:807–9.
73. Brackmann DE, Sheehy JL. Tympanoplasty: TORPs and PORPs. Laryngoscope 1979;89:108–14.
74. Silverstein H, McDaniel AB, Lichtenstein R. A comparison of PORP, TORP, and incus homograft for ossicular reconstruction in chronic ear surgery. Laryngoscope 1986;96:159–65.
75. Emmett JR, Shea JJ Jr, Moretz WH Jr. Long-term experience with biocompatible ossicular implants. Otolaryngol Head Neck Surg 1986;94:611–6.
76. Jackler RK, Schindler RA. Myringoplasty with simple mastoidectomy: results in eighty-two consecutive patients. Otolaryngol Head Neck Surg 1983;91:14–7.
77. Ruhl CM, Pensak ML. Role of aerating mastoidectomy in noncholesteatomatous chronic otitis media. Laryngoscope 1999;109:1924–7.
78. Naclerio R, Neely JG, Alford BR. A retrospective analysis of the intact canal wall tympanoplasty with mastoidectomy. Am J Otol 1981;2:315–7.
79. Sheehy JL. Mastoidectomy: the intact canal wall procedure. In: Brackmann DE, Shelton C, Arriaga MA, editors. Otologic surgery. Philadelphia: WB Saunders; 1994. p. 211–24.
80. Parisier SC, Hanson MB, Han JC, et al. Pediatric cholesteatoma: an individualized, single stage approach. Otolaryngol Head Neck Surg 1996;115: 107–14.
81. Sheehy JL. Tympanoplasty with mastoidectomy — a reevaluation. Laryngoscope 1970;80:1212–30.
82. Brackmann DE. Tympanoplasty with mastoidectomy: canal wall up procedures. Am J Otol 1993;14: 380–2.
83. Weber PC, Gantz BJ. Cartilage reconstruction of the scutum defects in canal wall up mastoidectomies. Am J Otolaryngol 1998;19:178–82.
84. Sismanis A, Hutchinson L, Abedi E. Chronic otitis media: surgical failures and management. Am J Otol 1989;10:460–5.
85. Rosenberg SI, Silverstein H, Hoffer M, Nichols M. Use of endoscopes for chronic ear surgery in children. Arch Otolaryngol Head Neck Surg 1995; 121:870–2.
86. Cody DTR, McDonald TB. Mastoidectomy for acquired cholesteatoma. Follow up to 20 years. Laryngoscope 1984;94:1027–8.
87. Bondy G. Totalaufmeisselung mit Erhaltung von Trommelfell und Gehork nochelchen. Monatsschr Ohrenheilkd 1910;44:15–23.
88. Toner JG, Smyth GDL. Surgical treatment of cholesteatoma: a comparison of three techniques. Am J Otol 1990;11:247–9.
89. Dodson EE, Hashisaki GT, Hobgood TC, Lambert PR. Intact canal wall mastoidectomy with tympanoplasty for cholesteatoma in children. Laryngoscope 1998;108:977–83.
90. Nadol JB. Causes of failure of mastoidectomy for chronic otitis media. Laryngoscope 1985;95:410–3.
91. Rambo JHT. Primary closure of the radical mastoidectomy wound; a technique to eliminate postoperative care. Laryngoscope 1958;68:1216–27.
92. Palva T. Mastoid obliteration. Acta Otolaryngol Suppl (Stockh) 1979;360:152–4.
93. Meyerhoff WL, Stringer SP, Roland PS. Rambo procedure: modification and application. Laryngoscope 1988;98:795–6.

CHAPTER 11

Cranial and Intracranial Complications of Acute and Chronic Otitis Media

Lee A. Harker, MD

Complications of acute and chronic otitis media can cause grave morbidity and even mortality. Even though the incidence and prevalence of these complications have dramatically declined, their potential gravity requires physicians to have a thorough understanding of the diagnosis and management of each one. Because of the decreased prevalence of these complications and changes in the health care delivery system, many otolaryngologists-in-training today have no opportunity to see patients with these problems. This chapter first provides an overview of the complications of acute and chronic otitis media, including their etiology, pathology, pathophysiology, diagnosis, and treatment. The latter part of the chapter presents a review of each complication.

Cranial and intracranial complications of otitis media can occur in individuals of any age, but they occur much more commonly in children in the first two decades of life. Table 11–1 illustrates the age distribution of extracranial (cranial), intracranial, and combined complications in a large series of patients from the rural Natal province in South Africa. Nearly 80% of cranial complications and 70% of intracranial complications occurred in children in their first two decades of life. Cranial complications, led by postauricular abscess, most commonly occurred in children under age 6 years.[1] In a series of 93 intracranial and cranial complications of otitis media that occurred in Turkey during the 1990s, 58% were present in patients under age 20 years.[2] For unexplained reasons, males are affected nearly twice as often as females.

People who are poor, live in overcrowded surroundings, and have poor personal hygiene, poor health, decreased resistance to infection, inadequate health education, and limited access to medical care have the highest incidence of these complications.

TABLE 11–1. Age Distribution of 268 Patients with Complications of Otitis Media between January 1985 and December 1990.

Age (yr)	*Cranial (n = 87) %*	*Intracranial (n = 150) %*	*Cranial and Intracranial (n = 31) %*
0–5	33.3	8.0	6.5
6–10	23.0	21.3	19.4
11–20	21.8	39.3	38.7
21–30	9.2	8.6	12.9
31–40	3.4	12.6	9.7
41–50	5.7	6.0	6.5
51–60	3.4	0.6	3.2
> 60	0.0	3.3	3.2

Adapted from Singh B and Maharaj TJ.[1]

For example, most of the current reports of otogenic brain abscesses come from South Africa and developing countries. Only an occasional reported series is from North America. Although the number of patients with acquired or iatrogenic immunodeficiency has increased, large series of patients with complications of otitis media have not been reported in patients with human immunodeficiency virus (HIV) infection or acquired immune deficiency syndrome (AIDS) or in those receiving immunosuppressive therapy following organ transplantation, even though such individuals do have suppurative ear disease.

Table 11–2 shows the classification of cranial and intracranial complications, and Table 11–3 illustrates the relative frequencies of those complications reported in two recent series.[1,2] The dominant cranial complication is postauricular abscess, and the dominant intracranial complication is meningitis. Unlike many areas of medicine, the complications tend to occur multiply, especially the intracranial complications, as shown in Tables 11–4 and 11–5. Whereas all of the complications originate from infection in the pneumatized spaces of the middle ear and mastoid, the mechanisms by which complications occur in acute otitis media differ from those leading to complications in chronic otitis media. Therefore, these two entities are briefly discussed separately.

Acute Otitis Media

An estimated 85% of all children experience at least one episode of acute otitis media, making it the most common bacterial infection of childhood.[3] Predisposing factors include being of a young age or the male sex, receiving bottle feedings, and being exposed to a daycare environment, crowded living conditions, or smoking within the home. Medical conditions such as cleft palate, Down syndrome, and mucous membrane abnormalities such as cystic fibrosis, ciliary dyskinesia, and immunodeficiency states also predispose individuals to otitis media.

Acute otitis media is a bacterial infection of the middle ear space characterized by vascular dilatation and proliferation (manifested externally by erythema), mucosal edema, exudation, bacterial proliferation, white blood cell infiltration, and pus formation. Acute otitis media refers only to an acute infection that arises de novo, from a previously normal middle ear, rather than an acute clinical infection arising in a chronically abnormal middle ear, such as one with long-standing otitis media with effusion. This distinction is important because the pathophysiology of the otitis and the development of complications are different in acute versus chronic otitis media. Acute otitis media may resolve completely and spontaneously with or without treatment, or it may cause a complication while remain-

TABLE 11–2. Classification of Complications of Acute and Chronic Otitis Media

Cranial Complications	*Intracranial Complications*
Coalescent mastoiditis	Meningitis
Chronic mastoiditis	Brain abscess
Masked mastoiditis	Subdural empyema
Postauricular abscess	Epidural abscess
Bezold's abscess	Lateral sinus thrombosis
Temporal abscess	Otitic hydrocephalus
Petrous apicitis	
Labyrinthine fistula	
Facial nerve paralysis	
Acute suppurative labyrinthitis	
Encephalocele and cerebrospinal fluid leakage	

TABLE 11–3. Distribution of Intracranial and Cranial Complications

	Osma et al[2]		*Singh and Maharaj*[1]	
Complication	*n*	*%*	*n*	*%*
Intracranial				
Meningitis	41	71.9	22	12
Brain abscess	10	17.5	93	51
Epidural abscess	4	7.0	19	10
Lateral sinus thrombosis	1	1.8	36	20
Cerebritis	1	1.8		
Subdural empyema			36	20
Cranial				
Mastoid abscess	25	64.1	65	75
Labyrinthitis	5	12.8	—	—
Facial nerve paralysis	5	12.8	15	14.9
Bezold's abscess	4	10.3	5	5.7
Petrous apicitis	—	—	2	2.2

Adapted from Osma U et al[2] and Singh B and Maharaj TJ.[1]

ing in its acute stage. Conversely, it may persist into a subacute or chronic state. Factors that favor the acute infection progressing to a subacute or chronic otitis media include microorganism virulence, decreased host resistance, and inadequate or inappropriate antibiotic therapy.

After the first few weeks of life, acute suppurative otitis media is caused primarily by three microorganisms: *Streptococcus pneumoniae, Haemophilus influenzae,* and *Branhamella catarrhalis,* compromising roughly 30, 20, and 10% of isolates, respectively.[3] Optimal treatment for acute suppurative otitis media with complications includes appropriate antibiotic treatment of the underlying infection for 10 days, in addition to tympanocentesis, or myringotomy and placement of a ventilating tube. Tympanocentesis is used primarily to obtain material for culture to identify the offending microorganism and its antibiotic sensitivities, but it can also somewhat reduce the bacterial population. Doing a myringotomy and tube placement also provides the physician with material to identify the involved microorganism while simultaneously allowing the physician to remove most of the bacteria from the middle ear space and to introduce topical antibiotic therapy. After treatment, the physician should document that the acute otitis media has completely resolved by otoscopy and a normal tympanogram. If the complication was intracranial, a computed tomographic (CT) scan or a magnetic resonance image (MRI) should be obtained.

Chronic Otitis Media

Whereas acute otitis media is primarily a middle ear infection that extends into the contiguous mastoid, chronic otitis media often displays a dominant mastoid infection with concurrent otitis media. In the mastoid, infection persists in the numerous small air cells easily or more easily than it does in the relatively large single cell of the middle ear. Chronic otitis media is present when an infectious process persists for a period longer than the 1 to 3 weeks usually necessary for acute otitis media to resolve in a previously normal ear. Chronic otitis media can occur with or without a cholesteatoma when there is a tympanic membrane perforation. A third type of chronic otitis media is emerging, especially in young children who develop persistent otorrhea with a patent middle ear ventilating tube.

TABLE 11–4. Intracranial Complications in 181 Patients with Otitis After Singh and Maharaj[1]

Complication	*n*	*No. with Associated Complications*	*AC%*
Meningitis	22	4	18.2
Brain abscess	93	15	16.1
Subdural empyema	36	12	33.3
Epidural abscess	19	19	100.0
Lateral sinus thrombosis	36	10	27.8

AC% = percentage of patients with each complication who also had additional cranial or intracranial complication(s). Adapted from Singh B and Maharaj TJ.[1]

When the infection in the middle ear and mastoid will not resolve, mucosal edema and exudation increase, and the mucous glands and secretory elements proliferate. Mucosal edema in the small spaces between the mesotympanum and the epitympanum and in the aditus between the epitympanum and mastoid antrum block the normal pathways for aeration and decrease both oxygenation and vascularity. At the same time, the blockage prevents topical antibiotic and anti-inflammatory agents from reaching the attic and mastoid. Radiographically, the mastoid air cell system is partly or completely opaque, reflecting the loss of aeration.

These changes are accompanied by a dramatic change in the bacterial flora from acute to chronic otitis media and mastoiditis. Harker and Koontz cultured 30 cholesteatomas at surgery and isolated at least one anaerobic microorganism in 67% of the cases, at least one aerobic microorganism in 70%, and both aerobes and anaerobes in 50%.[4] In 57% of the cholesteatomas, multiple microorganisms were cultured, and in 30%, five or more bacteria were identified. Even when there was no clinical infection, anaerobic microorganisms, such as *Propionibacterium acnes,* were frequently isolated. Significantly, the study showed that an ear with a cholesteatoma is highly likely to harbor multiple bacteria of both anaerobic and aerobic types. Multiple microorganisms and anaerobic bacteria also have been identified frequently in chronic otitis media without cholesteatoma.

Chronic otitis media that develops in patients with indwelling middle ear ventilating tubes has a different bacterial flora. In most instances, these cases begin during an upper respiratory infection or when the external auditory canal is contaminated with water. A series of events, including treatment with antibiotic drops, treatment with oral antibiotics, repeated contamination, repeated cultures and sensitivities, increasing patient and physician frustration, decreasing patient compliance, and fungal overgrowth, frequently results in resistant microorganisms developing. The resistant bacteria most often found are *Pseudomonas aeruginosa, Achromobacter xylosoxidans,* and even methicillin-resistant *Staphylococcus aureus*. It is also common to have a clinically significant fungal infection of both the external auditory canal and the middle ear and mastoid. Although the author has not seen cranial and intracranial complications from this type of infection, granulation tissue he has noted at surgery suggests that possibility.

Patients with intracranial or multiple complications often appear more systemically ill than they should be from otitis alone. They can present with toxicity or with obtundation, manifesting depressed

TABLE 11–5. Interrelationship of Intracranial Complications

Complication	*n*	*Men*	*BA*	*EA*	*SE*	*LST*
Meningitis	22		—	2	—	1
Brain abscess	93	—		3	9	3
Subdural empyema	36	—	9	2		1
Epidural abscess	19	2	3		2	5
Lateral sinus thrombosis	36	1	3	5	1	

Adapted from Singh B and Maharaj TJ.[1]
Men = meningitis; BA = brain abscess; EA = epidural abscess; SE = subdural empyema; LST = lateral sinus thrombosis.

levels of consciousness that can vary from lethargy to total unresponsiveness. Focal neurologic signs may be absent, subtle, or florid. The physical examination of the ear itself usually does little to pinpoint a specific complication unless there is an obvious postauricular, cervical, or temporal abscess.

Pathophysiology

Sometimes a complication occurs during the first few days of an episode of acute otitis media. This can result from bacteremia accompanying acute otitis media initiating a hematogenous complication such as meningitis, or the bacteria proliferating in the middle ear can gain access to adjacent or distant structures and continue the infection at the new location. A patient may develop partial or complete facial paralysis if he or she has a congenital dehiscence of the bony fallopian canal above the stapes. If there are preformed pathways leading to the meninges or the labyrinth, patients with acute otitis media are at risk of developing meningitis, subdural effusion, or suppurative labyrinthitis. Developmental labyrinthine abnormalities, such as enlarged vestibular aqueducts or Mondini's deformity, often have such pathways. In older children and adults, bacteria can propagate along preformed pathways left from previous surgery or temporal bone fractures. But with acute otitis media, rarely is there granulation tissue formation or bone destruction; complications develop by hematogenous dissemination or by direct extension of infection along preformed pathways. Accordingly, the usual medical treatment of the acute otitis media will resolve the otitis portion, and mastoidectomy is not necessary. Therefore, it is critical to know that the middle ear was normal before the current bout of otitis media began.

In chronic otitis media and mastoiditis, complications are always associated with some combination of bone destruction, granulation tissue formation, or the presence of cholesteatoma. Bacteria gain access to the involved structures most commonly by direct extension from mastoid infection and by infecting and propagating along veins leading from the mastoid to adjacent structures. Direct extension can come about as a result of bone resorption from cholesteatoma or osteitis, or it can occur without bone erosion if the patient has preformed pathways from previous mastoid surgery, temporal bone fracture, congenital dehiscences, or other conditions that removed the intervening bone. The mechanisms by which bone is actively resorbed, although incompletely understood, probably include enzymatic degradation, suppuration, and decreased blood supply.

Diagnosis

History Because a thorough understanding of the chronology of the events is critical to determining the pathophysiology, the clinician must obtain a complete history of the present illness. The clinician must establish when the patient's ear was last free of disease and perfectly normal to differentiate acute otitis media from chronic otitis media because the bacteriology, medical treatment, and most probable complications are different for each. Answers to questions such as (1) when did a physician last examine the involved ear? (2) what is the past history and treatment of the patient's otitis media? (3) did the patient use an antibiotic for otitis media in the involved ear in the past month? and (4) what were the order of appearance and the magnitude of the different symptoms that comprise the present problem? are essential to an accurate diagnosis. Finally, the physician must obtain any objective evidence that the ear was normal recently (tympanogram, radiographic study that included the ears). The physician must gather from the patient or a knowledgeable accompanying person or through telephonic contact with physicians' offices and hospitals information about the involved ear including previous diagnoses, all operations, and recent medical and surgical treatment.

Symptoms of acute otitis media include pain, fever, and fussiness, but otorrhea is relatively uncommon. The cardinal symptom of chronic otitis media is painless, purulent drainage. Changes in the drainage may suggest the possibility of a complication. The cessation of drainage with the onset of pain suggests the presence of aditus or epitympanic block with increased pressure within the mastoid. The onset of imbalance or vertigo at the same time suggests the possibility of imminent labyrinthitis.

Intracranial complications, most notably meningitis, intraparenchymal brain abscess, and subdural empyema, can alter the patient's level of consciousness. Fifteen percent of Singh and

Maharaj's patients were drowsy on admission, 18% were stuporous, and 2% presented in coma.[1] Establishing the chronology of this alteration of sensorium will help the physician differentiate among diagnoses of brain abscess, meningitis, and subdural empyema. A brain abscess takes weeks to develop, whereas it takes only a few hours to several days for meningitis and subdural empyema to be fulminant and progress to coma.

In eliciting the patient's past history, social history, family history, and review of systems, the clinician should gather not only information relevant to the present illness but also all of the information the anesthesiologist needs to evaluate a patient for a several-hour surgical procedure.

Physical Examination The vital signs, especially the temperature, provide a pretreatment baseline and one parameter for following the course of the disease and the treatment; however, if the patient has previously received oral or parenteral antibiotics, he or she may present without a fever. Some patients remain afebrile during the entire course of illness, despite having significant cranial or even intracranial complications, but the temperature curve can sometimes provide useful information (see "Lateral Sinus Thrombosis").

A complete neurologic examination is essential. Make a thorough mental assessment of the patient, and note any alteration in mental state. Evaluate the station and gait and perform Romberg's and sharpened Romberg's tests. Evaluate the motor and sensory function of the extremities and perform a complete cranial nerve evaluation, including an assessment of vision, extraocular muscle function, facial nerve function, and facial sensation. Note whether there is nystagmus, either in the straight-ahead cardinal position or with the eyes deviated 30 degrees to the left and to the right. Evaluate cerebellar function by checking the alternate motion rate of the extremities, determining if past pointing is present or absent, and performing the finger-to-nose test. Assess ocular saccades and smooth pursuit. It is critical to determine if nuchal rigidity is present and, if so, to perform Kernig's and Brudzinski's tests. Observe the optic disks with an ophthalmoscope to determine if papilledema is present. Write down all of your positive and negative findings so that there is a written record of the patient's status at a specific time to compare to any subsequent changes. Changes in the patient's symptoms and the physical signs provide valuable information for understanding the evolving pathophysiology and determining the appropriate treatment.

Begin the otologic examination with an assessment of the color, size, shape, and position of the pinna compared with the opposite side. Make note of any erythema, tenderness, or drainage and any evidence of trauma, excoriation, or protrusion outward or downward. Next observe the regions adjacent to the auricle and note any swelling, erythema, tenderness, purulent drainage, or fluctuance.

Examine the external auditory canal and tympanic membrane using a microscope and fine suction. If purulent secretions are present, culture them. Document the presence of any edema and whether it primarily affects the posterior superior bony canal wall or if it affects the entire canal circumference. Make a drawing of the tympanic membrane illustrating any perforation, granulation tissue, or epithelial debris and any erosion of the scutum. Perform pneumatic otoscopy looking for conjugate deviation of the eyes that strongly suggests the presence of a labyrinthine fistula. Make a clinical impression whether the patient has acute otitis media, chronic otitis media with perforation, or a cholesteatoma before obtaining any imaging studies.

The tympanic membrane may appear normal or near normal even when an otitic complication is present because it reflects only the status of the middle ear medial to it. Whereas mastoid infection always begins with a middle ear infection, suppuration in these two locations may proceed differently in that the middle ear may revert to normal or near normal under treatment, whereas the mastoid may not. Especially in the presence of an aditus block, the middle ear may appear nearly or completely normal after several courses of antibiotics, whereas symptoms from the mastoid persist (see "masked mastoiditis"). When evaluating a patient with any infectious condition that could be caused by acute or chronic otitis media, the clinician should obtain a CT scan to rule out chronic otitis as the cause even if the tympanic membrane appears normal and there is no history of ear disease.

Laboratory Examination The complete blood cell count will document the patient's relative degree of leukocytosis and ensure that his or her hemoglobin and hematocrit are adequate for safe general anes-

thesia. The individual health needs of the patient and the necessity for a general anesthetic will determine the need for any other hematologic studies.

Imaging Techniques Computed tomographic scanning is essential for all patients suspected of having complications of otitis media. It is a fast and reliable method for assessing the status of the middle ear and mastoid air cell system and diagnosing intracranial complications of otitis media.[5,6] It is unparalleled for showing the bony details of the middle ear, epitympanic, and mastoid structures and documenting the degree to which the normally aerated pneumatized spaces may be opacified by the inflammatory process. It clearly shows demineralization and loss of the bony septae of air cells in coalescent mastoiditis and reveals erosion of the bony plates covering the sigmoid sinus and cerebellum or tegmen of the middle ear and mastoid, even erosions of the bony labyrinth itself. Computed tomographic scans must include views taken in an algorithm to assess soft tissue changes and wide ("bone") algorithm views to evaluate bone erosion.

Computed tomographic scans can help to establish the specific primary otologic diagnosis (acute otitis media, chronic otitis media, cholesteatoma) as well as the diagnosis of several of the specific cranial and intracranial complications of otitis media. In addition to their diagnostic value, CT scans are useful to assess the results of therapy and to provide a baseline post-treatment study of the mastoid for comparison in case of further complications.

When the patient is somnolent or unstable and when intracranial complications are suspected, CT scanning may be the initial study of choice because it is fast and gives the health care team better access to the patient during the study than MRI does. Even without enhancement, CT scanning may be an adequate diagnostic tool for a febrile, stuporous patient with meningeal irritation when ruling out the presence of an intraparenchymal brain abscess or communicating hydrocephalus prior to performing lumbar puncture to establish the diagnosis of meningitis. However, when using CT to diagnose cerebritis, cerebral abscess, subdural empyema, and ventriculitis, intravenous injection of an iodinated contrast agent is essential.

Magnetic resonance imaging is a far more sensitive imaging technique than CT scanning for diagnosing intracranial complications because paramagnetic contrast agents such as gadolinium-DTPA (diethylenetriamine pentaacetic acid) cross the blood-brain barrier in areas of cerebritis or abscess. Meningeal enhancement is much more easily seen with MRI than with CT scanning in which the adjacent bony skull often obscures the meninges. The T_2-weighted images of an MRI can demonstrate intraparenchymal edema from subtle brain infection much earlier than can a CT scan. When otitic complications are suspected, the clinician usually obtains both CT and MRI scans if the patient's condition allows because much of the information from the two studies is complementary.

Lumbar Puncture To determine the presence of meningitis, the physician must perform a lumbar puncture, measure the patient's cerebrospinal fluid (CSF) pressure at the beginning and end of the procedure, examine the CSF sample for the presence of bacteria on direct smear, and determine the concentrations of glucose, chloride, and protein for comparison with their concentrations in serum. Lumbar puncture should be undertaken only after ophthalmoscopic examination and CT scan have ruled out significant increased intracranial pressure that could result in herniation of the cerebellar tonsils during or after the procedure. Lumbar puncture is contraindicated in the presence of brain abscess and subdural empyema.

Treatment

Although each complication has its own separate treatment, there are certain general management principles for treating the underlying otitis. For each patient, the treatment of the underlying acute or chronic otitis media and the treatment of any complications can be medical or surgical and can be administered sequentially or concurrently. In nearly all instances of complications resulting from acute otitis media, appropriate antibiotic therapy will be sufficient to resolve the otitis without any need for surgery. This will usually include aspiration by tympanocentesis or myringotomy, as previously discussed.

When the complications have resulted from chronic otitis media and mastoiditis, initial antibiotic therapy should employ broad-spectrum antibiotics effective against anaerobic as well as aerobic

microorganisms. In all of these cases, some form of mastoidectomy will be required. When surgery is necessary for intracranial complications, the neurosurgeon operates first, immediately followed by the otologist performing a mastoidectomy at the same sitting if the patient's condition permits. In most instances (except for brain abscess and subdural empyema), both the chronic otitis media and its complications are treated entirely through the mastoid. When both neurosurgery and otologic surgery are necessary, the surgeons must carefully and methodically plan and coordinate the order of the procedures, patient preparation, draping, and incisions to decrease the duration of general anesthesia and to optimize the chances of a safe and effective surgical result.

Performing a mastoidectomy under these circumstances is difficult because vascularity increases and landmarks are more obscured. The following general principles are helpful guidelines. After removal of the mastoid cortex, complete all bony dissection using diamond instead of cutting burs and use copious suction-irrigation. This will minimize the chance of inadvertent trauma to important underlying structures. If preoperative scanning identified the sites of the complications, complete the surgery in the portions of the mastoid away from the complication sites first. This will improve hemostasis and visibility. When no cholesteatoma is associated with the mastoiditis, leave the external auditory canal wall intact unless visibility is inadequate.[1] Remove the canal wall and perform an open cavity procedure in the presence of cholesteatoma.

The rate of mortality in these cases correlates most with the patient's level of consciousness on admission. In the South African series, mortality was 1% when patients were fully conscious, 7% when drowsy, 30% when stuporous, and 67% when comatose.[1] The complications most likely to be lethal are brain abscess, meningitis, and subdural empyema.

Postoperative follow-up is an essential part of the general treatment of patients who experienced life-threatening complications of acute or chronic otitis. Even when the patient responded perfectly, the surgeon should seriously consider follow-up CT scan to confirm objectively the status of the mastoid at the termination of treatment because there is a risk of recurrence or emergence of new intracranial complications. In patients who have experienced lateral sinus thrombosis, epidural abscess, subdural empyema, or brain abscess, follow-up evaluation with enhanced MRI 2 to 4 weeks after treatment is recommended.

CRANIAL (INTRATEMPORAL) COMPLICATIONS

Coalescent Mastoiditis

Etiology Sometimes when a patient has acute otitis media and mastoiditis that persists unabated for 2 to 4 weeks, coalescent mastoiditis sets in. This is an acute progressive clinical infection with corresponding changes in the bone and mucoperiosteum of the mastoid air cell system. Coalescent mastoiditis is a disease of the young, especially males. Most patients are 4 years old or younger. Bacterial virulence and decreased host resistance are important in its etiology, but mastoid development also plays a role. The condition rarely develops in children who have had chronic ear disease or in those with poorly pneumatized mastoids containing few air cells. Rather, it tends to occur in patients with well-developed air cell systems that contain numerous small pneumatic spaces and in those who have had little or no previous otologic disease.

Pathology Initially, hyperemia and edema of the mucoperiosteal lining of the mastoid air cells block the narrow aditus and disrupt aeration. The mucous membrane thickens, and impaired ciliary function prevents normal middle ear drainage through the eustachian tube. The serous exudate becomes purulent as inflammatory cells accumulate. Continued inflammation, hyperemia, and accumulation of purulent debris cause venous stasis, localized acidosis, and decalcification of the bony septae. The osteoclastic activity in the inflamed periosteum softens and decalcifies the bony partitions, causing the small air cells to coalesce into a larger cavity.[7]

Pathophysiology As the infection grows, pressure within the mastoid cavity increases and conditions become progressively more favorable for it to extend beyond the confines of the mastoid. In the presence of such an intense inflammation and infection, phlebitis and periphlebitis are common and spread the infection to the adjacent meninges, sigmoid sinus, cerebellum, and temporal lobe.[8] The infection sometimes extends directly to the meninges, sigmoid sinus, labyrinth, or facial nerve because of bone ero-

sion. The most common pathway for infection to extend beyond the mastoid is through the lateral cortex behind the ear. Less commonly, it can extend to the soft tissues in the upper portion of the neck (see "Bezold's Abscess") and rarely to the soft tissue anterior and superior to the auricle either by direct extension through eroded bone or by phlebitis and periphlebitis. Go et al reviewed the records of 118 children with acute mastoiditis at Texas Children's Hospital between 1986 and 1998 and found only 8 patients in whom the mastoiditis had caused an intracranial complication.[9]

Diagnosis The symptoms and signs of coalescent mastoiditis—purulent otorrhea, fever, toxicity, and ear pain—are the same as those seen in patients with uncomplicated acute otitis media. The strongest historical suggestion of coalescent mastoiditis is the chronology of the infection in which purulent drainage or significant otalgia persists for 2 or more weeks, recurs after 10 to 14 days, or significantly worsens after that time interval. As a group, children with coalescent mastoiditis look sicker and have more toxicity with higher and more persistent fevers than those with acute otitis media. Older children may be able to localize the pain to the postauricular area rather than the ear canal. Physical findings that are most helpful include mastoid tenderness to percussion, mastoid erythema, and sagging of the posterior superior external auditory canal wall.

The clinician should order a complete blood count and hematologic studies followed by CT scanning, which can establish the diagnosis by documenting the breakdown of the bony air cell walls and opacification of the pneumatized spaces. If any suggestion of an intracranial complication exists, the clinician should obtain an enhanced MRI scan.

Treatment The treatment for coalescent mastoiditis can be either medical or surgical. Without question, complete mastoidectomy with ventilating tube placement in conjunction with appropriate antibiotic therapy provides prompt, precise eradication of all infected tissue in an expeditious, cost-effective manner. However, because the increased vascularity and granulation tissue greatly increase the difficulty of the operation, it should not be undertaken lightly. Another consideration is that pneumatization has not progressed to incorporate the mastoid tip in children under the age of 2 years, so there is risk of surgical injury to the facial nerve. However, the author's experience with cochlear implant surgery in children 12 to 24 months of age suggests that the facial nerve does not exit through the lateral surface of the mastoid in this age group as is commonly taught, and the risk of facial nerve injury in this age group is not substantial.

It is equally true that appropriate intravenous antibiotics for a minimum of 3 to 6 weeks also will eradicate the disease process in the majority of infected infants who have no additional complications. The therapeutic choice for a specific patient will depend on the clinical factors present and regional preferences. At the end of medical therapy, it is essential to document that the disease process has been completely eradicated. The patient should continue receiving antibiotic therapy until a CT scan produces evidence that the mastoid air cell system is no longer opacified and the middle ear is normally aerated.

Chronic Mastoiditis

Etiology Chronic mastoiditis can occur in association with a long-standing tympanic membrane perforation, with cholesteatoma, or as a complication from an infection following placement of a middle ear ventilating tube. As noted above, ventilating tube mastoiditis tends to occur in young children who have experienced water contamination and have undergone cultures and treatment with multiple antibiotic drops and oral antibiotics. Mastoiditis with tympanic membrane perforation occurs when an episode of acute otitis media with perforation pursues a course of chronic infection rather than resolving or developing into coalescent mastoiditis. Chronic mastoiditis of this type may also begin when an uninfected ear with a long established central perforation becomes infected, the infection extends to the mastoid, and the ear continues to drain chronically. Even though cholesteatoma frequently remains uninfected for long periods of time, it tends to suppurate, form granulation tissue, and erode bone. Once any type of mastoiditis causes continuous purulent drainage for 8 or more weeks, the likelihood of a complete resolution with antibiotics significantly decreases. Chronic mastoiditis requires surgical intervention to heal, and an infected cholesteatoma requires surgical ablation,

regardless of duration. Complications in patients with chronic mastoiditis with tympanic membrane perforation may develop at any time, but they often occur only after weeks or months of otorrhea. In contrast, cholesteatoma typically requires months or years to produce complications.

Masked Mastoiditis

Chronic otitis media with granulation tissue formation and bone erosion can occur without otorrhea. It can persist in spite of a normal or near-normal tympanic membrane. This has been referred to as "masked mastoiditis," and it usually occurs in patients who have received numerous courses of antibiotics. In this complication, the epitympanum and aditus become blocked so that the middle ear responds to the antibiotics, but the mastoid does not. Recently, the author has seen three patients with this problem. Each patient had aural and postauricular pain, slight but definite tenderness to percussion of the mastoid cortex, and CT scan evidence of very localized areas of opacification in an otherwise normal mastoid. Surgical excision relieved the patients of all of their symptoms. There are two keys to the successful management of masked mastoiditis: (1) mastoid disease is not always reflected by the appearance of the tympanic membrane, and (2) chronic mastoiditis is a surgical condition, regardless of the appearance of the tympanic membrane. In this regard, the diagnosis and management of masked mastoiditis are no different from those of mastoiditis with otorrhea emanating from other causes.

Samuel and Fernandez described 21 cases of mastoiditis with retroauricular swelling and an intact tympanic membrane in a South African population (19 patients were under age 13 years) in whom the duration of symptoms was shorter than that usually seen with coalescent mastoiditis.[10] The predominant finding at mastoidectomy was granulation tissue filling the mastoid cavity and the antrum. The tympanic membrane was described as hyperemic, dull, bluish, bulging, or retracted. In addition to their mastoiditis, 10 patients had postauricular abscesses, 3 had Bezold's abscesses, and 2 had facial nerve palsies. Four patients had cellulitis or abscess in the posterior fossa. Although these cases do not represent masked mastoiditis, they demonstrate the important fact that medically significant mastoiditis can exist without purulent otorrhea.

Postauricular Abscess

Postauricular abscess is the most common complication of mastoiditis. It is most often seen accompanying coalescent mastoiditis in young children. The infection extends from the mastoid to the subperiosteal space. Usually, this occurs by direct extension subsequent to bone destruction or by phlebitis and periphlebitis of mastoid veins. The tiny pits in adult temporal bones that make up the cribriform area of the mastoid near the spine of Henle exist in newborns as a series of open vascular channels between the interior of the mastoid and the cortex. In very young children, infection can pass directly from the mastoid to the subperiosteal space until these channels have closed. Regardless of how it starts, soft tissue infection leads to tissue necrosis and abscess formation. The surrounding soft tissue will exhibit thickening, inflammation, erythema, tenderness, and fluctuance.

The diagnosis is usually obvious. The tissue edema and the abscess drive the auricle downward and laterally because only the upper part of the mastoid is pneumatized (Figure 11–1). In earlier stages, if fluctuance is not obvious, the clinician should use imaging studies or ultrasonography to document the presence of air within the soft tissue or an abscess

FIGURE 11–1. Postauricular abscess associated with coalescent mastoiditis. The auricle is displaced laterally and inferiorly.

cavity with its capsule. When mastoiditis has produced an abscess, the case for excision and drainage in conjunction with mastoidectomy is much stronger. Only in the most unusual of circumstances would treatment by prolonged antibiotics and drainage of the abscess without mastoidectomy be appropriate.

Bezold's Abscess

In his 1913 textbook *Diseases of the Ear*, Kerrison gave the following description of Bezold's abscess:

> This condition is caused by a perforation in the bony plate forming the inner surface of the tip of the mastoid. It occurs presumably in cases in which the tip cells are especially large and in which the bony plate forming the inner or medial wall of the tip is very thin, and the outer cortex thick. Pus escaping through such a perforation burrows downward in the neck beneath the sternomastoid, or may be confined between layers of the deep cervical fascia.[11]

The cervical infection develops into an abscess in the upper neck deep to the sternocleidomastoid muscle. Bezold's abscess can also develop without any erosion or penetration of the inner and outer cortex of the mastoid if phlebitis and periphlebitis propagate the infection to the same area. Because infants have limited mastoid pneumatization, Bezold's abscess occurs more commonly in older children in whom pneumatization has extended into the mastoid tip and in adults who have either chronic mastoiditis or cholesteatoma.

The diagnosis of Bezold's abscess is often not considered immediately in young patients with deep, tender, upper cervical masses because inflamed lymph nodes from many causes are so common. If the history and physical examination do not reveal a specific cause, the clinician should obtain a CT scan to identify or rule out a mastoid source.

The recommended treatment is complete surgical excision of the mastoid pathology, drainage of the abscess, and removal of any associated granulation tissue. The surgeon should use bipolar electrocautery, copious suction irrigation, and coarse diamond burs to allow him or her to visualize the pathology adequately and exenterate thoroughly all of the diseased cells. The surgeon should drain both the mastoid and the abscess cavity.

Temporal Root Abscess

The soft tissues above and even those anterior to the auricle can become infected and form an abscess from suppuration that involves the cells of the zygomatic process of the temporal bone. As is the case with postauricular and Bezold's abscesses, temporal root abscesses can form by direct extension through eroded zygomatic process cells or by phlebitis and periphlebitis. The clinical picture can be very confusing because abscesses in this location are rare, but there are so few other etiologic possibilities. Computed tomographic scanning is always recommended to rule out mastoiditis as the source, and surgery is the appropriate treatment.

Petrous Apicitis

A petrous apex can be undeveloped (sclerotic), can contain marrow, or can exhibit some degree of pneumatization. Pneumatization of the petrous apex develops in only 30% of temporal bones.[8] Much has been written about the different cell tracts that extend into and pneumatize the petrous apex. These cells can extend into the apex from above (supralabyrinthine; Figure 11–2, A and B), behind (retrolabyrinthine; Figure 11–3, A and B), beneath (infralabyrinthine; Figure 11–4, A and B), and/or in front of the labyrinth (anterior labyrinthine). Petrous apicitis is essentially mastoiditis that occurs in the petrous apex. It is rare because infection in sclerotic or marrow-containing petrous apices is uncommon and the prevalence of pneumatization is low. Petrositis develops by direct extension of a mastoid infection, but the mastoid may respond to medical or surgical treatment without resolution of the apical infection. Just as there can be disjunction between the state of infection in the middle ear and the mastoid, the same disjunction can exist between the mastoid and the petrous apex.

The pathology of the infection can mirror that seen in coalescent mastoiditis with dissolution of thin cellular septi and coalescence, or it can form granulation tissue and erode bone like chronic mastoiditis. Only rarely does cholesteatoma extend to the apex. Imaging studies will usually include both CT and MRI. A CT scan will show the bony details of the septae of the air cells and the size and contour of the apex as a whole. Magnetic resonance imaging will differentiate marrow from mucus or CSF. Both CT and MRI

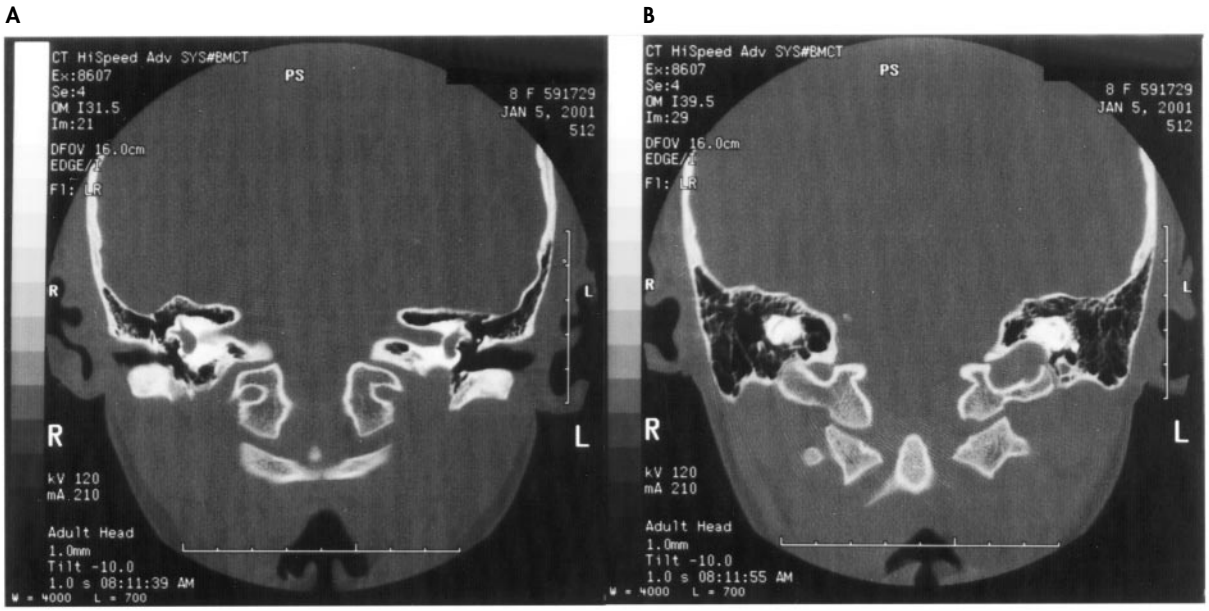


FIGURE 11–2. Petrous apex pneumatization via the supralabyrinthine cell tract at the level of the vestibule (*A*) and posterior semicircular canal (*B*).

studies are essential to establish opacification of the air cells in the suspected petrous apex. When one petrous apex is well pneumatized, a small sclerotic or marrow-containing apex on the opposite side can be misinterpreted as a pneumatized apex opacified by fluid or infection unless both studies are obtained.

The symptomatology of infection in the petrous apex reflects the innervation of the air cells and the structures adjacent to the apex itself. Whereas increased pressure within the mastoid air cell system causes pain in the region of the mastoid and the ear, increased pressure in the petrous apex is usually

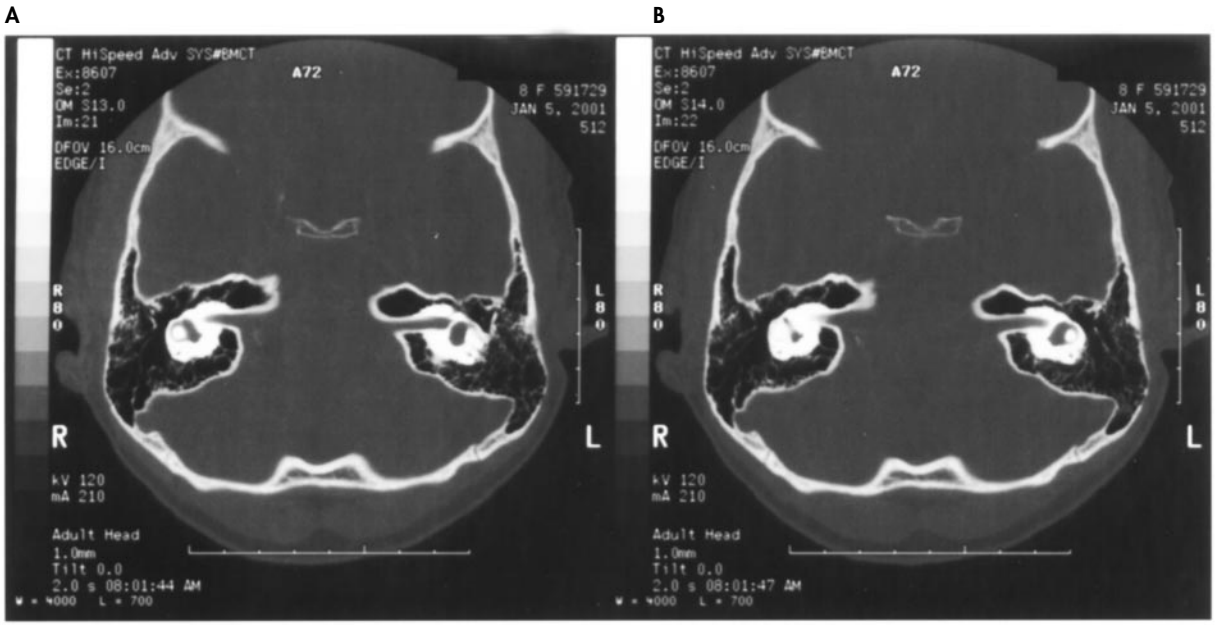


FIGURE 11–3. Petrous apex pneumatization via the retrolabyrinthine cell tract at two levels (*A* and *B*) of the internal auditory canal and vestibule.

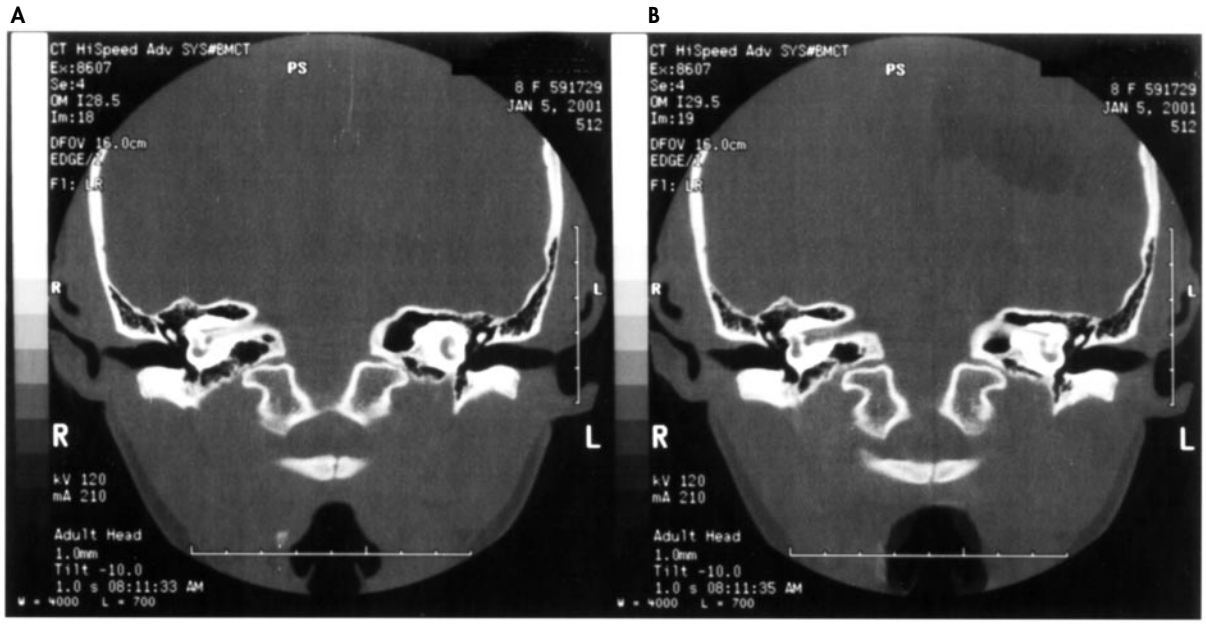


FIGURE 11–4. Petrous apex pneumatization via the infralabyrinthine cell tract in the right temporal bone at the level of the internal auditory canal.

referred to the retro-orbital area or deep within the skull. Patients with petrous apicitis have symptoms of infection in the middle ear and mastoid, as well as symptoms from the apex. The most common symptoms from infection in the petrous apex are retro-orbital pain from irritation of the contiguous trigeminal ganglion in Meckel's cave, paralysis of the sixth cranial nerve as it passes through Dorello's canal abutting the apex, or dysfunction of the facial, acoustic, or vestibular nerves as they course through the temporal bone. In 1904, Gradenigo described the triad of retro-orbital pain, sixth nerve paralysis, and otorrhea, which has since been known as Gradenigo's syndrome.[8] Only a minority of patients with petrous apicitis exhibit the full triad today.

How petrous apicitis is best treated depends on its duration, its severity, and whether additional complications are present. Because the labyrinth, carotid artery, and internal auditory canal are housed in the petrous pyramid, it is difficult or impossible to excise all of the air cells, which is done so routinely in the mastoid. Therefore, prolonged antibiotics are necessary, even when surgery is performed. When necrotic bone is obvious, surgical drainage is a necessary adjunct to intravenous antibiotic therapy. Even partial apicectomy can reduce the bacterial load sufficiently to allow antibiotics and host defenses to control the disease. The best surgical approach to the apex depends on the portion of the apex that is involved and which air cell tracts connect the apex to the mastoid. In some temporal bones, the apex can be reached by a route posterior to the posterior semicircular canal (see Figure 11–3, A and B); in others, the best route is infralabyrinthine (see Figure 11–4, A and B). In a small percentage of temporal bones, limited drainage can be achieved by removing cells extending over the top of the superior semicircular canal and through its center (the "hole in the doughnut"; see Figure 11–2, A and B). The original anterior petrous apex drainage procedure described by Julius Lempert incorporated a radical mastoidectomy operation and is rarely done today.[12] The middle cranial fossa approach can provide excellent access to obtain material for culture and to remove infected tissue when the air cell tracts from the mastoid to the apex are inadequate. Postoperative CT scans are useful to verify that the disease has been excised and as a reference for future studies.

Labyrinthine Fistula

A labyrinthine fistula results from the erosion of the endochondral bone of the bony labyrinth. The loss of this bone allows the underlying endosteum, perilymph, and structures of the endolymphatic compartment to move when the pressure in the external

auditory canal is changed. Motion of the fluids in the endolymphatic compartment brings on the symptoms of labyrinthine fistulae. Almost all labyrinthine fistulae affect the lateral semicircular canal; the superior and posterior semicircular canals, vestibule, and cochlea are rarely involved. Cholesteatoma is the cause in nearly all cases, and fistulae were seen to occur in 7% of the cholesteatomas in Gersdorff and Nouwen's large series.[13] Higher percentages have been reported, but the actual incidence is unknown. Most reports of labyrinthine fistulae are from tertiary care referral centers that include many patients with large or previously operated cholesteatomas, which artificially elevates the incidence. On the other hand, studies reviewing suppurative complications of otitis media often do not include labyrinthine fistulae if the fistula is not infected or does not cause significant symptoms or if the surgeon first detects it at operation.

The mechanisms by which a cholesteatoma causes bone erosion are not fully understood, but the dense endochondral bone first becomes demineralized and then is absorbed so that a progressively smaller amount of bone exists between the endosteal membrane overlying the perilymph and the cholesteatoma matrix above it. When the bone is completely resorbed, the surgeon can see what appears to be a "blue line" parallel to the underlying semicircular canal lumen because the illuminating light at the blue line is no longer reflected off the dense bone but is absorbed into the underlying fluid.

Manolidis recently reviewed the records of 111 inner-city Texas patients with labyrinthine fistulae and looked for any coexisting complications.[14] Two associations were prominent. The facial nerve was involved with cholesteatoma or was damaged by the cholesteatoma in 60% of the patients, and dehiscences of the tegmen occurred in 39%. The majority of these patients had two or more operations, with an average of 2.6 operations per patient. Surgeons should suspect a triad of complications: lateral semicircular canal fistula, facial nerve exposure, and tegmen erosion in all patients undergoing revision surgery for cholesteatoma or in patients with primary acquired cholesteatomas that developed in well-pneumatized mastoids.

Labyrinthine fistulae cause mostly vestibular symptoms. Patients describe short periods of imbalance, dysequilibrium, or vertigo but have normal equilibrium most of the time. Some recall that they feel suddenly off balance when they hear a sudden loud sound, which is called Tullio's phenomenon, or when they push on their external ear canal, for example, when cleaning their ear with a washcloth.

After obtaining a history that suggests the presence of a fistula, the physician should perform the fistula test, a specific maneuver that can establish the diagnosis. The examiner occludes the external auditory canal with the pneumatic otoscope and alternately increases and decreases the pressure slightly so that pressure changes are transmitted from the canal to the middle ear and mastoid air cell system through either an intact or a perforated tympanic membrane. During the test, the patient is told to look directly ahead at a specific object, and the physician observes the patient's eyes for any horizontal deviation. In a normal ear, changes in external canal pressure cause no motion of the eyes and no symptoms. When a lateral canal fistula is present, positive pressure transmitted from the external auditory canal compresses the lateral semicircular canal endosteum and causes utriculopetal endolymph flow. This produces a positive fistula sign, conjugate deviation of the eyes away from the side of the compression. Negative external auditory canal pressure produces a conjugate deviation toward the ear being tested. The patient may become slightly nauseated or perceive a to-and-fro motion of the environment during the test. Only 55 to 70% of patients with lateral canal erosions have a positive fistula test, but in those patients, it is a highly reliable indication that a fistula is present and allows the surgeon to plan the operation so that he or she can avoid entering the fistula during surgery. Inadvertently opening a labyrinthine fistula usually causes total loss of hearing in that ear.[15] Computed tomographic scanning also can suggest the diagnosis. Images in the bone algorithm will show erosion of the lateral semicircular canal and any other signs of cholesteatoma.

Surgery is the recommended treatment for labyrinthine fistulae. Usually, the surgeon will remove the posterior external auditory canal wall and perform an open-cavity mastoidectomy. If the patient has no other complications, the temporal bone is well pneumatized, and the exposure of the lateral semicircular canal is good, the surgeon can consider performing a canal wall up technique. Whether the canal wall is left up or removed, the surgeon completes the operation and removes all of the cholesteatoma except for a small area around

the fistula site. After carefully removing the overlying cholesteatoma debris without disturbing the matrix, the surgeon looks for the "blue line" that identifies the site of the fistula and the thin layers of lateral semicircular canal bone on either side. The surgeon develops a dissection plane between the cholesteatoma matrix and the endosteum using high microscopic magnification and a flat dissector that is 2 to 3 mm wide. The matrix is elevated with a fine microsuction to improve visualization of the dissection plane. A small piece of tissue or a thin cap of bone is placed over the site and secured with fibrin glue or packing after the cholesteatoma is successfully removed. If the endosteum is torn, it is best to replace the matrix and terminate the procedure.

Spontaneous vestibular symptoms usually improve after the fistula is repaired, but symptoms from ear canal pressure changes may persist for some time, and a positive fistula sign will persist until there is a regrowth of bone over the site of the fistula. The principal risk in removing the cholesteatoma matrix from the fistula is total or partial loss of hearing, which occurs in less than 20% of carefully managed fistula cases.[16]

Facial Nerve Paralysis

Facial nerve paralysis can result from acute otitis media, chronic otitis media without cholesteatoma, or cholesteatoma. Bacteria reach the nerve because of a congenital dehiscence of the bony fallopian canal or because the canal is eroded by granulation tissue or cholesteatoma. Facial nerve function is lost because of edema secondary to pressure and inflammation or because suppurative neuritis destroys the neural elements. If the edema persists, avascular necrosis of the axons can lead to axonal degeneration.

Facial paralysis in young children that is caused by acute otitis media is frequently incomplete and probably occurs only in infants with congenital dehiscence of the fallopian canal in the middle ear adjacent to the stapes. Facial weakness in these cases rarely lasts longer than 3 weeks even when complete paralysis is present.

Facial nerve paralysis secondary to chronic otitis media without cholesteatoma also usually affects the horizontal portion of the facial nerve near the stapes.[17] In these cases, the clinical course of the paralysis is more likely to be prolonged, with a gradual progression from slight weakness to full paralysis. Sometimes the progression to complete paralysis is rapid. Facial nerve paralysis secondary to cholesteatoma is also most likely to involve the nerve in the middle ear because it is there where most cholesteatomas arise. Extensive erosion of the horizontal segment of the fallopian canal is especially common in large, uninfected primary acquired cholesteatomas. When there is no infection, the nerve usually retains normal function. An erosive cholesteatoma can expose the facial nerve anywhere in the temporal bone and cause paralysis. In these cases, the onset of the paralysis is usually gradual, and sometimes the progression is so slow that patients do not seek medical attention for months. When the onset of facial paralysis is this slow, the paralysis is more likely to persist after surgical treatment.

The clinical presentation of facial paralysis is discussed more fully in Chapter 24. Facial nerve dysfunction secondary to temporal bone infection is not usually accompanied by significant pain or a facial tic but is simply expressed as muscle weakness or paralysis.

When facial paralysis is caused by acute otitis media, appropriate antibiotic therapy for the acute otitis may be adequate, although myringotomy with evacuation of the purulent material and reduction of the numbers of bacteria is recommended. When facial nerve paralysis follows chronic suppurative otitis media (with or without cholesteatoma), the surgeon should remove the infection surrounding the nerve as part of the mastoidectomy. The surgeon should address the diseased area of the facial nerve after the rest of the operation is completed in the same way that he or she would manage a labyrinthine fistula. The surgeon gradually approaches the granulation tissue or cholesteatoma overlying the nerve from the proximal and distal portions of the nerve that are uninvolved by the chronic infection using diamond burs to carefully remove the bone of the fallopian canal on both sides of the diseased portion. The surgeon uses a flat blunt instrument to dissect the chronic inflammatory tissue from the nerve while elevating the diseased tissue with a small suction tip. It may be necessary to use sharp dissection to separate the inflammatory tissue from the epineurium. The outcome for surgical decompression of the facial nerve secondary to chronic suppurative otitis media primarily depends on whether the nerve has undergone complete degeneration before surgery.

Acute Suppurative Labyrinthitis

Bacterial invasion of the labyrinth is always promptly followed by total loss of both auditory and vestibular function. Usually, acute otitis media extends into the labyrinth through a weakened or dehiscent oval window membrane, as seen in congenital labyrinthine deformities, such as Mondini's deformity and enlarged vestibular aqueduct syndrome, and in individuals who have undergone stapes surgery. Although it is not proven, suppurative labyrinthitis may be a common mechanism for unilateral anacusis in children with Mondini's deformity. These children are at additional risk. The foramina of the internal auditory canal opening into the medial aspects of the labyrinth may also be weak or dehiscent, and those foramina and the cochlear aqueduct can permit bacterial infection to progress from the labyrinth to the meninges or vice versa. It is not known how frequently suppurative labyrinthitis causes meningitis or how often meningitis subsequently causes bacterial labyrinthitis, but both probably occur, especially in the special population of children with congenital labyrinthine abnormalities.

Direct bacterial invasion of the labyrinth through a cholesteatomatous lateral semicircular canal fistula is another cause of acute suppurative labyrinthitis. In this situation, infected granulation tissue beneath the cholesteatoma matrix lies directly on the endosteal membrane and the underlying perilymph. The bacteria causing the labyrinthitis are those either of the underlying acute otitis media or of the cholesteatoma. The diagnosis is clinical. Within 30 to 60 minutes from the beginning of the infection, tinnitus and dizziness are followed by whirling vertigo, pallor, diaphoresis, nausea, and vomiting. The vestibular symptoms continue at their maximum for at least 8 to 12 hours even if the patient is totally immobile and receiving intravenous antiemetics. Brisk labyrinthine nystagmus directed toward the opposite ear accompanies the vertigo. After continuing for several hours, the spontaneous vertigo and nystagmus gradually abate. Symptomatic improvement continues during the next few days, but any motion of the head will evoke vertigo and nausea. Over the next 2 or 3 weeks, central nervous system compensation occurs, and normal or near-normal balance is restored. The tinnitus also abates, but all hearing is lost.

When the physician recognizes the typical clinical picture in a patient with a predisposing condition and acute or chronic otitis media, no specific diagnostic studies are necessary. Although there is no possibility of reversing the clinical course, appropriate antibiotic treatment for 10 days is recommended to eradicate the labyrinthine infection and prevent propagation to the meninges. Other therapeutic measures are dictated by the underlying otitis, but labyrinthectomy is not necessary in labyrinthitis secondary to acute otitis media.

Encephalocele and Cerebrospinal Fluid Leakage

Encephalocele (brain hernia, brain fungus, meningoencephalocele) and CSF leakage may be associated with other cranial or intracranial complications in acute or chronic otitis media. Three different clinical patterns can occur.

Spontaneous CSF leakage from defects in the tegmen tympani (less commonly the tegmen mastoideum) have been reported in more than 50 patients who have not undergone previous otologic surgery.[18,19] Seventy to 80% of the patients are over 45 years of age. The defects range in diameter from 2 mm to 2 cm and are sometimes multiple. Usually, an encephalocele protrudes through the defect. Although there is no agreement as to why the dura breaks down and allows cerebral herniation and CSF leakage in these older patients, Jackson and associates have suggested that aging, increased intracranial pressure, low-grade inflammation, arachnoid granulation, and irradiation can play a part.[20] Patients frequently present because of hearing loss owing to reduced ossicular motion from the middle ear fluid or the encephalocele. Myringotomy and placement of a ventilating tube release a profuse watery otorrhea that tests positive for CSF. Other patients present with signs and symptoms of meningitis, and many of these patients have already experienced one or more episodes of meningitis.

Encephalocele and CSF leakage secondary to chronic otitis media occur when a cholesteatoma and granulation tissue erode through the bony plates that separate the mastoid from the temporal lobe and the cerebellum and then erode through the dura. The mastoid and epitympanic tegmen are involved more frequently than the cerebellar plate. It usually takes many months to years for a cholesteatoma to cause this degree of bone erosion, cerebral prolapse, and dural erosion.

Traumatic encephalocele and CSF leakage are the most common pattern. Although a few of these cases occur secondary to temporal bone fracture, the great majority are the consequence of an operation in which portions of the tegmen or cerebellar plates were removed and the dura was exposed or traumatized. In these cases, both the disease process and the surgery contribute to the bony and dural trauma that facilitates the development of an encephalocele, CSF leakage, and intracranial complications.

Management of temporal bone encephalocele requires prompt surgical repair. Jackson and associates have written an excellent review of this subject.[20] The size of the bony defect and the volume of herniated brain are two of the important factors in choosing a surgical approach. Small defects usually can be adequately managed through the mastoid, but multiple and large defects are best repaired by a combination of transmastoid and middle cranial fossa approaches. Jackson et al recommended a combined transmastoid intradural middle cranial fossa approach for large encephaloceles. To reduce the risk of infection, the defects are usually repaired with autogenous graft materials.

INTRACRANIAL COMPLICATIONS

Meningitis

Meningitis is by far the most common intracranial complication of acute and chronic otitis media. In Gower and McGuirt's series of 100 consecutive patients with intracranial complications of acute and chronic otitis, 76 had meningitis, and 53 of those were under 2 years of age.[21] In infants and toddlers, the great majority of cases occur by hematogenous dissemination of infection during acute otitis media. Sometimes these patients are not included in reports that document the complications of otitis, and the frequency of otogenic meningitis appears to be low.[1]

Conditions that allow CSF to gain access to the middle ear cavity predispose the patient to developing or suffering a recurrence of bacterial meningitis. Cerebrospinal fluid otorrhea is common after lacerations of the dura by longitudinal temporal bone fractures and occasionally is seen in patients who spontaneously develop defects of the tegmen of the middle ear or mastoid with associated encephaloceles. Unrecognized injury to the dura of the posterior or middle cranial fossa during mastoid surgery is an overlooked source of CSF leakage that can lead to meningitis. Patients with syndromic or nonsyndromic congenital stapes fixation, Mondini's dysplasia, and enlarged vestibular aqueducts are also at risk for a bacterial middle ear infection spreading to the CSF. The pathway can be obvious, as, for example, when a "perilymph gusher" is discovered during stapes surgery, or it can be more subtle, as when there are microscopic disruptions of the continuity of the oval or round windows, internal auditory canal, and cochlear aqueduct. Meningitis can occur even in individuals with normal middle ear and labyrinthine anatomy if acute otitis media complicates conditions such as traumatic stapes dislocation or perilymphatic fistula and then progresses to acute suppurative labyrinthitis.

The predominant symptom of meningitis is a generalized, severe headache. The patient will tend to lie quietly, vomit frequently, and exhibit photophobia and general hyperesthesia. The patient's level of consciousness may be normal, somnolent, stuporous, or unresponsive. Fever, often high and sustained, is nearly universal. Nuchal rigidity and pain when the clinician attempts to flex the patient's neck are ominous signs that should compel the clinician to perform other neurologic tests. In Brudzinski's sign, passive flexion of the head on the chest is followed by involuntary flexion of both thighs and both legs. In Kernig's sign, when the patient is in the supine position and one thigh is flexed to a right angle, attempts at passive extension at the knee are accompanied by pain and resistance owing to spasm of the hamstring muscles. Fundoscopic examination may reveal the presence of papilledema.

Whenever meningitis is suspected, a CT scan is critical to rule out the presence of brain abscess, cerebritis, or subdural empyema and to determine that it is safe to perform lumbar puncture (ie, intracranial pressure is not inordinately increased). Cerebrospinal fluid from the lumbar puncture is examined for intrathecal pressure, cells, bacteria, glucose, protein, chloride, and other factors. Cerebrospinal fluid pressure increases early in the course of the disease and protein and glucose concentrations increase in comparison to serum values as the disease progresses, but bacteria are not present until late in the disease.

When meningitis is secondary to acute otitis media or acute otitis media with suppurative labyrinthitis, antibiotic treatment of the meningitis is also adequate treatment for the otitis, and no sur-

gery is indicated. Tympanometry and otoscopy should be performed after treatment is complete to document that the middle ear has returned to normal. However, when meningitis develops from chronic otitis media and mastoiditis, the mastoid must be exenterated, and all disease must be excised. Even though the patient has already had a CT scan before the lumbar puncture, an enhanced MRI should be obtained to rule out the presence of any additional intracranial complications. The neurologic condition of the patient determines the timing of the surgery, and the mastoidectomy should be performed as soon as the patient is stable. Meningitis is one of the gravest complications of acute or chronic otitis. Further, meningitis appears to be more lethal when secondary to chronic otitis media and mastoiditis than when it is a hematogenous complication of acute otitis media in the first 2 years of life.

Brain Abscess

The incidence and mortality of intraparenchymal brain abscesses have decreased considerably. Nearly all reports in the past two decades come from centers outside North America and Western Europe. However, recent reports from Africa and Asia, where new imaging techniques are used, have documented substantial improvements in diagnosis and therapy and clarified some aspects of surgical treatment. Yen and associates' recent series of 122 consecutive patients seen in a Taiwan hospital between 1981 and 1994 revealed that otitis was the third most common cause of intraparenchymal brain abscess, exceeded only by those associated with cyanotic congenital heart disease and those secondary to head injury or neurosurgery.[22] Seventy-five percent of abscesses occurred in males, principally from the lower socioeconomic classes. In the 1990s, four reports from India, Turkey, and South Africa discussed 149 patients with otogenic brain abscesses.[1,2,23,24] Abscesses from acute and chronic otitis media occurred nearly always on the same side as the otitis and nearly equally in the temporal lobe or cerebellum. Almost three-fourths were secondary to cholesteatoma, half occurred in patients in the second decade of life, and two-thirds affected males.

A brain abscess begins when bacteria propagate in and around venous channels leading from the mastoid into the adjacent brain parenchyma. As soon as the bacteria arrive in the cortex or white matter, polymorphs migrate into local capillaries, producing endothelial swelling and focal cerebritis. The tissue becomes edematous, hemorrhagic, and necrotic. The abscess may vary greatly in size. Often it has an irregular shape and is multilocular. At first, the capsule is poorly defined, but over time it firms and can be easily stripped from the underlying edematous brain.[25]

In addition to manifesting symptoms and signs of general intracranial sepsis, patients with cerebellar abscesses often exhibit coarse horizontal nystagmus, dysmetria, dysdiadochokinesia, or action tremor. Temporal lobe abscess may cause homonymous visual field defects, contralateral hemiparesis, and other focal signs listed below for subdural empyema. The physician should carefully examine the patient and the imaging studies because up to two-thirds of patients with intraparenchymal abscesses will have more than one intracranial complication.[23]

The physician should immediately begin broad-spectrum intravenous antibiotics directed at both aerobic and anaerobic microorganisms. If imaging studies suggest no other complications and the patient's condition is stable, neurosurgical drainage or abscess excision is performed within the first 24 hours of admission, followed immediately by mastoidectomy performed through a separate surgical field. Mortality has decreased to approximately 10%.

Subdural Empyema

Subdural empyema is a fulminating purulent infection that develops between the dura and the pia-arachnoid membranes. It is one of the most immediate of neurosurgical emergencies. It comprises only about 20% of all cases of localized intracranial bacterial infection, and an otitic origin is uncommon compared to bacterial contamination (from trauma or a neurosurgical operation), suppurative paranasal sinus disease, or meningitis. At least two-thirds of the cases occur in men, and most occur during the second decade of life. The abscess usually begins by direct spread from adjacent infected bone or by retrograde venous propagation. Once infection enters the subdural space, pus forms rapidly and spreads widely. Thrombophlebitis of cortical veins is virtually guaranteed. Swelling, necrosis, and infarction of the cortex account for many of the clinical features and explain how a

barely detectable, thin layer of subdural pus can cause such devastating consequences as raised intracranial pressure, focal signs, and seizures.[26]

Clinically, patients exhibit a cascade of symptoms that begins with a severe headache, the most persistent symptom. The temperature rises dramatically as the disease progresses. General malaise, chills, and nuchal rigidity indicate that the patient is becoming seriously ill. After an unpredictable period of time, the patient's level of consciousness may abruptly decline and focal signs and symptoms may develop. Most patients with collections of pus on the left side of the brain develop aphasia and a progressive contralateral hemiparesis. Paralysis of conjugate gaze to the contralateral side and deviation of the eyes toward the side of the lesion are also common. Jacksonian seizures are common at this stage, and papilledema may be evident. In patients with subdural empyema in the posterior fossa, localizing signs are often absent, but marked neck stiffness and papilledema are always present. The entire clinical picture of subdural empyema may develop in as little as a few hours or as long as 10 days.

In a well-developed subdural empyema, a contrast-enhanced CT scan will show a crescent-shaped, low-density collection of pus displacing the brain from the inner table of the skull, with enhancement of the adjacent edge of cortex. Unfortunately, the scan may appear normal early in the course of the disease. When this happens, the clinician should obtain an enhanced MRI or repeat the CT scan after a suitable time interval. Lumbar puncture is dangerous and is contraindicated because it may precipitate transtentorial coning; however, it may be unavoidable when the CT scan fails to detect the lesion and focal signs are absent because meningitis and brain abscess must also be ruled out.

Emergency neurosurgical drainage and appropriate antibiotic therapy are necessary. The patient's likelihood of survival will be related to his or her level of consciousness at the time of surgery. The presence of localizing signs in any patient with chronic otitis media whose level of consciousness is declining requires immediate and decisive action.

Subdural empyema can also develop in neonates as a complication of meningitis. Up to 30% of infants and children with *H. influenzae* meningitis develop a collection of fluid in the subdural space (usually bilaterally) that varies from 5 to 100 mL.[27] This problem is less common after meningitis from other microorganisms. Its frequency has also decreased since universal immunization against *H. influenzae* has dramatically reduced the incidence of meningitis. Uninfected subdural collections after meningitis are called "subdural effusions." If bacteria are evident on direct examination of the fluid, the collections are called "infected subdural effusions," and if the fluid is grossly purulent, they are classified as "subdural empyemas." In this age group, subdural collections from meningitis secondary to chronic otitis media that require otologic attention are rare.

Epidural Abscess

Epidural abscess results from bone erosion caused by cholesteatoma, granulation tissue, or coalescence. In addition to the bone erosion, an intense inflammatory response to the infection results in granulation tissue formation and forms an abscess between the temporal bone and the dura mater. The involved dura thickens in response to its contact with granulation tissue on its surface (pachymeningitis). Although it can occur as the only complication of chronic mastoiditis, epidural abscess is frequently associated with lateral sinus thrombophlebitis, meningitis, and cerebritis or brain abscess. Rarely, epidural abscess can result from acute otitis and mastoiditis in the same way as will be described for lateral sinus thrombosis.

Most patients experience deep mastoid pain, but there are no signs or symptoms specifically attributable to epidural abscess, and many small epidural abscesses without associated complications are discovered only at surgery. If the abscess is large enough, it can be detected on enhanced CT or MRI scans as a fluid-filled cavity between the temporal bone and the enhanced dura. Surgery is the only recommended treatment. The surgeon performs the type of mastoidectomy that is appropriate for the underlying otitis and mastoiditis. After removing the cortex, he or she should progressively exenterate the air cells, proceeding from lateral to medial and from regions of less disease to those with more disease. When the locus of pathology is the cerebellar plate, the surgeon should first smooth the tegmen mastoideum and tegmen tympani and then thin the posterior aspect of the bony

external auditory canal and remove the tip cells to maximize exposure and hemostasis. Then the surgeon approaches the diseased area along the cerebellar plate from the superior, inferior, lateral, and, finally, medial directions, removing the overlying bone until only a thin plate remains. Because the dura has been thickened by the infection, blunt flat instruments, such as a Freer elevator or large curettes, can safely scrape the granulation tissue from the abscess cavity and help the surgeon identify the most centripetal extent of the abscess. The surgeon should remove all of the bone overlying the abscess using diamond burs, rongeurs, or curettes until healthy dura without granulation tissue is evident on all margins of the abscess. When the abscess involves the tegmen, the surgeon should change the order of the surgical steps to complete the surgery in the other areas first and address the tegmental abscess last.

Lateral Sinus Thrombosis

Thrombosis of the lateral sinus usually forms as an extension of a perisinus abscess that develops after mastoid bone erosion from cholesteatoma, granulation tissue, or coalescence.[27] The perisinus abscess exerts pressure on the dural outer wall of the sinus, and necrosis results. The necrosis extends to the intima and attracts fibrin, blood cells, and platelets, and a mural thrombus forms. The mural thrombus becomes infected, enlarges, and occludes blood flow through the sinus. Fresh thrombus can propagate in a retrograde direction to the transverse sinus, the torcular Herophili, and even to the superior sagittal sinus. In the opposite direction, clot can extend via the jugular bulb into the internal jugular vein in the neck or via the inferior and/or superior petrosal sinuses to the cavernous sinus. The infected clot frequently showers the bloodstream with bacteria, which give rise to the signs and symptoms of septicemia and the possibility of metastatic abscesses (most commonly to the lungs).

Because of the relative frequencies of the various types of mastoiditis and the ages of those affected, lateral sinus thrombosis today is most common in adults or older children with cholesteatoma and develops less commonly from the other types of chronic mastoiditis. In addition to those cases arising by direct extension from bone erosion, however, lateral sinus thrombosis can also be caused by osteothrombophlebitis during acute otitis media and mastoiditis. In this situation (and in occasional cases of chronic mastoiditis), the bony sinus plate is intact at the time of surgical exploration.

Because the sinus is a continuation of the cerebellar dura mater, extension of an infection by only a few millimeters can result in the equally grave complications of meningitis, epidural abscess, subdural empyema, cerebritis, or cerebellar abscess. In two recent series, 14 of 19 patients with lateral sinus thrombosis had one or more of these additional complications.[28,29]

Although its frequent association with other complications obscures its clinical picture, a diagnosis of lateral sinus thrombosis can be strongly suspected when there are signs and symptoms of septicemia such as chills and spiking fever or blockage of blood flow through the sinus. A picket-fence fever pattern with diurnal temperature spikes that exceed 103°F has been described with this condition for decades. Some recent articles have suggested that this pattern is not seen as frequently, in part because many patients present with previous or current antibiotic therapy; however, Singh reported that the majority of his 36 patients were on intravenous antibiotics and still exhibited that fever pattern.[30] Unless the patient has been transferred from another hospital, only the patient's temperature on admission is available; nonetheless, the clinician who observes such a high fever spike should suspect the possibility of sigmoid sinus thrombophlebitis. Singh also noted that the neck pain in lateral sinus thrombosis may be mistakenly attributed to the neck stiffness of meningitis, when it is actually pain and tenderness occurring along the anterior border of the sternocleidomastoid muscle. Singh further identified percussion tenderness of the mastoid tip as another important sign.

More ominous are the signs of sudden intracranial hypertension secondary to decreased venous drainage from the skull (see "Otitic Hydrocephalus" below). The most prominent sign is the occurrence or sudden worsening of a severe headache. This is most likely to occur when the patient's dominant venous drainage system is obstructed (the right side in 60% of individuals). More grave is progressive obtundation. This may herald the development of cerebral edema from increased intracranial pressure secondary to involve-

ment of the superior sagittal sinus or the cavernous sinus. It carries a very high mortality rate.

Imaging studies and lumbar puncture are essential to sort out the possible types of complications in obtunded patients. If the patient's condition permits, both enhanced CT scan and enhanced MRI should be performed. The diagnosis can be made preoperatively when the clinician identifies enhancement of the wall of the sigmoid sinus on CT scan, the Delta sign. Sinus wall enhancement is more sensitively evaluated by MRI, which also can document abscess formation within the sinus and can exclude nearby subdural empyema, cerebritis, or cerebellar abscess.

All patients with sigmoid sinus thrombosis will require mastoidectomy to treat the underlying mastoid disease adequately. Treatment controversies exist in the timing of the mastoid surgery and the management of the infected clot and sinus wall. When a brain abscess is also present, surgical drainage of that abscess takes precedence over any mastoid procedure (see "Brain Abscess"). If the patient's condition permits, neurosurgical drainage of the brain abscess should be followed in the same anesthetic by expeditious mastoidectomy. When meningitis is present in conjunction with lateral sinus thrombosis, mastoid surgery should be completed as soon as the patient is neurologically stable.

Management of the clot can include anticoagulation, ligation of the jugular vein in the neck, and opening the sinus and evacuating the infected clot. Use of anticoagulants is rarely indicated, but it should be strongly considered when extension of the clot to the transverse sinus or cavernous sinus is documented or suspected. Other considerations in these deteriorating patients include neurosurgical decompression and corticosteroids. Similarly, ligation of the jugular vein in the neck should not be routine, but it should be considered whenever there is evidence of extension of the clot into the neck and should be strongly considered when septic emboli are present.

The sinus is usually addressed only after the mastoidectomy portion of the operation has been completed. The surgeon then widely exposes the sinus and the adjacent dura by removing the overlying bone with rough diamond burs. The surgeon passes an 18- or 20-gauge needle through the sinus wall and, if there is no free blood on aspiration, makes a linear incision through the sinus wall and evacuates the abscess and any infected clot. Some surgeons recommend extending the exposure, incision, and clot evacuation proximally and distally until free bleeding occurs. The author recommends a more conservative approach, extending the surgery proximally and distally only as far as a grossly infected clot is found to remove as much infection as possible. The mastoid and cerebellar walls of the sinus are usually thickened and relatively tough as a result of the chronic infection. The surgeon can safely scrape them gently with a broad flat instrument such as a Freer elevator to remove the infected tissue. Syms and associates reported treating six patients with lateral sinus thrombosis and concomitant additional intracranial complications with mastoidectomy without opening and evacuating the lateral sinus clot.[29] All patients survived, but the average hospital stay was more than 49 days—38 days when the patient with the longest stay was omitted. In contrast, Kaplan and Kraus's 13 patients underwent surgery that included incision of the sinus and evacuation of the clot or abscess.[28] Their mean hospital stay was only 17 days.

Otitic Hydrocephalus

Seventy years ago, Symonds described otitic hydrocephalus as a disease in which signs and symptoms of acute hydrocephalus followed some form of otitis.[31] Controversy and disagreement about the entity still exist today. Some authors view otitic hydrocephalus as a distinct clinical entity consisting of intracranial hypertension and resolved or resolving (usually acute) otitis media without any relationship to lateral sinus thrombophlebitis.[21] Others view it as a pathophysiologic consequence of lateral sinus thrombosis from any cause. The latter viewpoint holds that whether symptoms will develop following lateral sinus occlusion depends on (1) the size of the involved lateral sinus relative to that of the opposite side, (2) the adequacy of the collateral venous network including the cavernous sinus and the opposite inferior petrosal sinus, and (3) whether propagation of the process has occluded additional venous outflow. The physician should recognize the possible relationship among the symptoms, ear disease, lateral sinus, and emergent nature of the condition.

In the early stages, a diffuse, severe headache dominates the clinical picture. In uncomplicated cases with (or occasionally without) medical management, the continuous headache gradually ame-

liorates and dissipates over 3 to 7 days as collateral venous outflow increases to accommodate arterial inflow. If collateral venous drainage is inadequate and intracranial pressure remains elevated, signs of the sensorium dulling or decreased visual acuity from retinal vein occlusion will appear. The physician must recognize that these developments are an extreme emergency, obtain immediate neurosurgical consultation, and begin high doses of intravenous corticosteroids and diuretics. Rapid progression to coma and death is possible.

In assessing any patient suspected of having otitic hydrocephalus, the physician must make three determinations: (1) whether the patient has increased intracranial pressure, (2) whether the patient has active acute or chronic mastoid infection, and (3) whether there is free flow of blood through the involved lateral sinus. The physical examination and the imaging studies are paramount in making these determinations. The results of these inquiries will allow the physician to formulate the appropriate treatment of the mastoid, lateral sinus, and increased intracranial pressure.

REFERENCES

1. Singh B, Maharaj TJ. Radical mastoidectomy: its place in otitic intracranial complications. J Laryngol Otol 1993;107:1113–8.
2. Osma U, Cureoglu S, Hosgoglu S. The complications of chronic otitis media: report of 93 cases. J Laryngol Otol 2000;114:97–100.
3. Gates GA. Acute otitis media and otitis media with effusion. In: Cummings CW, editor. Otolaryngology-head and neck surgery. 3rd ed. St. Louis: Mosby-Year Book; 1998. p. 461–77.
4. Harker LA, Koontz FP. The bacteriology of cholesteatoma. In: McCabe BF, Sade J, Abramson M, editors. Cholesteatoma: first international conference. Birmingham (AL): Aesculapius; 1977. p. 264–7.
5. Latchaw RE, Hirsh WL, Yock DH. Imaging of intracranial infection. Neurosurg Clin North Am 1992;3:303–22.
6. Dobben GD, Raofi B, Mafee MF, et al. Otogenic intracranial inflamations: role of magnetic resonance imaging. Top Magn Reson Imaging 2000;11:76–86.
7. Dew LA, Shelton C. Complications of temporal bone infections. In: Cummings CW, editor. Otolaryngology-head and neck surgery. 3rd ed. St. Louis: Mosby-Year Book; 1998. p. 3047–75.
8. Neely JG. Complications of temporal bone infection. In: Cummings CW, editor. Otolaryngology-head and neck surgery. 2nd ed. St. Louis: Mosby-Year Book; 1993. p. 2840–64.
9. Go C, Bernstein JM, de Jong AL, et al. Intracranial complications of acute mastoiditis. Int J Pediatr Otorhinolaryngol 2000;52:143–8.
10. Samuel J, Fernandez C. Otogenic complications with an intact tympanic membrane. Laryngoscope 1985; 95:1387–90.
11. Kerrison PD. Diseases of the ear. Philadelphia & London: JP Lippincott; 1913.
12. Lempert J. Complete apicetomy (mastoido-tympano-apicetomy), new technique for complete apical exenteration of apical cartoid portion of petrous pyramid. Arch Otol 1937;25:144.
13. Gersdorff MC, Nouwen J. Labyrinthine fistula after cholesteatomatous chronic otitis media. Am J Otol 2000;21:32–5.
14. Manolidis S. Complications associated with labyrinthine fistula in surgery for chronic otitis media. Otolaryngol Head Neck Surg 2000;123:733–7.
15. Soda-Merhy A, Betancourt-Suarez M. Surgical treatment of labyrinthine fistula caused by cholesteatoma. Otolaryngol Head Neck Surg 2000;122:739–42.
16. Bak-Pedersen K, Ostri B. Labyrinthine fistula in chronic otitis media with cholesteatoma. In: Tos M, editor. Cholesteatoma and mastoid surgery. Amsterdam: Kugler & Ghedini; 1989. p. 1049–50.
17. Harker LA, Pignatari S. Facial nerve paralysis secondary to chronic otitis media without cholesteatoma. Am J Otol 1992;13:372–4.
18. May JS, Mikus JL, Matthews BL, Browne JD. Spontaneous cerebrospinal fluid otorrhea from defects of the temporal bone: a rare entity? Am J Otol 1995;16:765–71.
19. Lundy LB, Graham MD, Kartush JM, La Rouere MJ. Temporal bone encephalocele and cerebrospinal fluid leaks. Am J Otol 1996;17:461–9.
20. Jackson CG, Pappas DG, Manolidis S, et al. Brain herniation into the middle ear and mastoid: concepts in diagnosis and surgical management. Am J Otol 1997;18:198–206.
21. Gower D, McGuirt WF. Intracranial complications of acute and chronic infectious ear disease: a problem still with us. Laryngoscope 1983;93:1028–33.
22. Yen P-T, Chan S-T, Huang T-S. Brain abscess: with special reference to otolaryngologic sources of infection. Otolaryngol Head Neck Surg 1995;113: 15–22.

23. Kurien M, Job A, Mathew J, Chandy M. Otogenic intracranial abscess: concurrent craniotomy and mastoidectomy—changing trends in a developing country. Arch Otolaryngol Head Neck Surg 1998; 124:1353–6.
24. Murthy P, Sukumar R. Otogenic brain abscess in childhood. Int J Pediatr Otorhinolaryngol 1991;22:9–17.
25. Reid H, Fallon RJ. Bacterial Infections. In: Adams JH, Duchen LW, editors. Greenfield's pathology. New York: Oxford University Press; 1992. p. 302–34.
26. Kerr RSC, Mitchell RG. Abscess. In: Swash M, Oxbury J, editors. Clinical neurology. Edinburgh & London: Churchill Livingstone; 1991. p. 886–98.
27. Bell WE, McCormick WF. Neurologic infections in children. Philadelphia: WB Saunders; 1981. p. 141–5.
28. Kaplan DM, Kraus M. Otogenic lateral sinus thrombosis in children. Int J Pediatr Otorhinolaryngol 1999; 49:177–83.
29. Syms MJ, Tsai PD, Holtel MR. Management of lateral sinus thrombosis. Laryngoscope 1999;109: 1616–20.
30. Singh B. The management of lateral sinus thrombosis. J Laryngol Otol 1993;107:803–8.
31. Symonds CP. Otitic hydrocephalus. Brain 1931;54:55.

CHAPTER 12

Otosclerosis

Herman A. Jenkins, MD, Omid Abbassi, MD, PhD

The term otosclerosis is derived from the Greek words for "hardening of the ear." Although first recognized as a pathologic entity by Valsalva early in the eighteenth century,[1] it was not until the mid-nineteenth century, with the observations of authors such as Magnus,[2] von Tröltsch,[3] Politzer,[4] and Toynbee,[5] that the clinical presentation of the disease was clearly defined. Today, otosclerosis is recognized as an alteration in bony metabolism of the endochondral bone of the otic capsule. The ongoing process of resorption and redeposition of bone results in fixation of the ossicular chain and conductive hearing loss.

PREVALENCE

Otosclerosis occurs most commonly among Caucasians,[6] with an incidence of 1%, followed by Asians at 0.5%. It is far less common in African Americans. In postmortem examinations of temporal bones, Guild found evidence of a higher prevalence rate histologically, with 8.3% of Caucasians and 1% of African Americans manifesting the disease.[7] Although encountered in all age groups, usually the clinical presentation occurs between the second and fifth decades of life. The disease process has a female predominance of 2 to 1, and its progression tends to accelerate with hormonal changes, particularly during pregnancy or the use of birth control pills. Bilateral disease is present in 80% of patients.

HISTOPATHOLOGY

The otosclerotic process is divided into two phases histologically. Bone resorption and increased vascularity characterize the early phase. When present near the periosteum of the middle ear, this hyperemia is responsible for Schwartze's sign. As the mature collagen content diminishes, the bone acquires a spongy appearance (otospongiosis).[8] On hematoxylin-eosin staining, it assumes a bluish coloration, referred to as the blue mantles of Manasse.[9] In the late stage, the reabsorbed bone is replaced with dense sclerotic bone—thus the name otosclerosis.

When involving the stapes, otosclerosis often starts from the fissula ante fenestram, although focal lesions involving the posterior annular ligament are also seen. In general, it progresses from an anterior focal lesion to complete footplate involvement and, in more advanced cases, may fill the oval window niche entirely with new bone (obliterative otosclerosis). In contrast, the round window is less frequently involved, and complete obliteration is a rare finding. Involvement of the cochlea can result in sensorineural hearing loss.

ETIOLOGY

Evidence for inflammatory and infectious causes has been reported. A genetic component has long been recognized, and transmission has generally been accepted to be autosomal dominant with incomplete penetrance. The gene for otosclerosis has not been clearly identified, but authors in one study narrowed its location to chromosome 15q25-26.[10] Others have related otosclerosis to the *COL1A1* gene that encodes for type 1 collagen. Investigators have noted its similarities to osteogenesis imperfecta[11] and Paget's disease,[12] both sharing lesions in the otic capsule nearly identical to that of otosclerosis. A role for autoimmunity as been suggested, but the data remain inconclusive.[13]

Recent investigations on the role of infectious agents have implicated the measles virus as having, at least, an inciting role in patients with a genetic predisposition for otosclerosis.[14] Elevated titers of

immunoglobulin G specific for measles virus antigens have been found in the perilymph of patients with otosclerosis. Immunohistochemical evidence of measles antigen has also been demonstrated in active otosclerotic foci, using reverse transcription polymerase chain reaction amplification of measles virus ribonucleic acid.[15] However, the actual role of the virus in producing the disease is not established.

CLINICAL PRESENTATION

The clinical presentation of otosclerosis is a progressive conductive hearing loss in an adult. It is thought to occur in less than 20% of genetically affected individuals. Some patients report improved speech understanding in a noisy environment (known as paracusis of Willis). Tinnitus is the second most common complaint reported. Vestibular symptoms are uncommon. In women, an increase in the rate of progression of the hearing loss is reported with hormonal changes such as those associated with pregnancy. Sensorineural hearing loss may be associated with the conductive changes in the disease. However, an isolated sensorineural hearing loss owing to otosclerosis is rare.

Physical examination shows a normal appearance of the external auditory canal and tympanic membrane. Schwartze's sign, a reddish hue over the promontory caused by increased vascularity of the bone immediately under the periosteum, may be seen in the early stages of the disease but is not present in all patients.

LABORATORY TESTING

Depending on the stage of the disease, audiometric studies typically show a mild-to-moderate conductive hearing loss. The air–bone gap is wider at the lower frequencies. Carhart's notch is characteristic of otosclerosis and appears as a sensorineural hearing loss at 2 kHz that is spurious since the bone conduction in the mid-frequency range is not reliable. Stapes fixation interferes with the bone conduction of the acoustic signal, and up to 15 dB of the apparent sensorineural hearing loss disappear after surgery for otosclerosis. With progression of the conductive loss, the Rinne test demonstrates greater bone conduction than air conduction, and Weber's test lateralizes to the affected side. The tympanogram is either depressed (A_s) or normal. The stapedial reflex may be normal in the early stages but cannot be elicited as stapes fixation proceeds. Speech discrimination is often normal, except in patients with cochlear involvement. Clinically, other diagnostic studies such as high-resolution computed tomography (CT) and magnetic resonance imaging (MRI) are of little use in evaluation of otosclerosis. With high-resolution CT, however, one may be able to identify the sclerotic lesion.

MEDICAL MANAGEMENT

Medical management of otosclerosis remains controversial and is primarily directed at maturing the involved bone and decreasing osteoclastic activity. Shambaugh and Scott introduced use of sodium fluoride as treatment, based on its success in osteoporosis.[16] However, this required high doses, and the efficacy has not been clearly established. Bisphosphonates that inhibit osteoclastic activity and cytokine antagonists that inhibit bone resorption may offer hope for the future. At present, no medical treatment is recommended on a consistent basis. Hearing aids, however, do offer an effective means of nonsurgical management of hearing loss in otosclerosis.

SURGICAL MANAGEMENT

History

The history of surgery for otosclerosis dates back to Kessel in 1877 with early attempts at mobilization and extraction of the stapes.[17] Other otologists subsequently attempted similar procedures, notably Miot in 1890, who described extensive experience with surgery and suggested various surgical techniques.[18] His successes were many. However, the report by Seibenmann at the turn of the twentieth century was less encouraging and condemned the surgery.[19] This effectively ended these early ventures into stapes surgery for conductive hearing loss.

Interest in this field of ear surgery lay dormant until 1938, when Lempert described the fenestration procedure that became popular with otologists of the era, rekindling interest in surgery for otosclerosis.[20] The focus returned to the stapes with Rosen's report of stapes mobilization and the restoration of hearing.[21] Shea introduced the stapedectomy, a technique of stapes extraction with tissue coverage of the oval window and polyethylene strut reconstruction of the stapes, in the 1950s, ushering in the modern

approach to stapes surgery.[22] In the intervening years, techniques have changed, and new materials for reconstruction have been introduced; however, the principles still remain basically the same.

The surgical goal in otosclerosis is restoration of the transmission mechanism for sound from the tympanic membrane, going through the ossicular chain to the oval window, bypassing the resistance of the fixed stapes footplate. Today, a variety of techniques are used to correct for stapes footplate fixation. Generally, the stapes arch is removed, and either a perforation or a partial to complete removal of the footplate is performed. A prosthetic implant is employed to connect the incus to the oval window.

Patient Selection

Selection of patients for operation is based on audiologic findings and physical examination. Preferred are patients with normal middle ear aeration, free of any infection or tympanic membrane perforation and with a Rinne test that demonstrates bone conduction to be greater than air conduction. When bilateral disease presents, the worse hearing ear is treated first followed by the other ear several months later. Experienced otologists should perform surgery on the only hearing ear, exclusively and with great trepidation.

Informed Consent

Preoperative consent is obtained, informing the patient of the risks of loss of hearing, vertigo, injury to the facial nerve, loss or alteration of taste, tympanic membrane perforation, prosthesis extrusion, prosthesis mobilization, and residual conductive hearing loss. Singers and musicians are informed of a possible change in quality of sound perception that may affect their professional performance. Individuals who are exposed to or plan to be involved in activities associated with rapid and/or considerable change in pressure, such as scuba diving and piloting nonpressurized airplanes, are advised to refrain from stapedectomy.

Perioperative Treatment

The operation may be performed under either general or local anesthesia. With improvements in anesthesia, more otologists are now performing the operation under general anesthesia. Use of any anticoagulants during the 2 weeks prior to the operation should be avoided, including anti-inflammatory agents. Muscle relaxants in conjunction with anesthetic agents are not recommended because of their effect on facial nerve activity. Perioperative antibiotics are at the surgeon's discretion, but antiemetic agents are recommended.

Whether a general or local technique is employed, injections of local anesthetics should be administered in such a manner as to avoid unintentional involvement of the facial nerve medial to the mastoid tip. Sterility of the operative field is of paramount importance since a direct connection to the labyrinth is established during parts of the operation. Routine operative site scrub technique, including installation of antibacterial preparation solution into the external auditory canal, is recommended. Facial nerve activity is monitored by direct vision of the face through a transparent occlusive drape.

Surgical Technique

Stapes surgery may be performed via an endaural or anterior incisural incision. The endaural incision requires use of a speculum held stationary with the left hand or a speculum holder assembly. In contrast, the anterior incisura incision is held open using one or two self-retaining retractors, eliminating the need to operate through a speculum. The cosmetic result of the incision is rarely noticeable. The latter also provides direct access to the tragal cartilage if a perichondrial graft is desired.

Regardless of the approach, a tympanomeatal flap is raised, the annulus is identified, and hemostasis is established prior to entering the middle ear space. A 1% lidocaine with 1:100,000 epinephrine or an 1:1,000 epinephrine-soaked piece of Gelfoam is used in this step. The tympanomeatal flap is elevated anteriorly, and the chorda tympani nerve is dissected free toward the malleus. A portion of the scutum should be curetted to expose the incus and the incudostapedial joint. Fixation of the stapes is determined by palpation of the malleus while viewing the suprastructure and footplate of the stapes.

The distance between the mid-shaft of the long process of the incus and the footplate is measured (Figure 12–1, A). The incudostapedial joint is separated, and the stapedial tendon is severed. The stapes suprastructure is fractured inferiorly and removed from the middle ear (Figure 12–1, B).

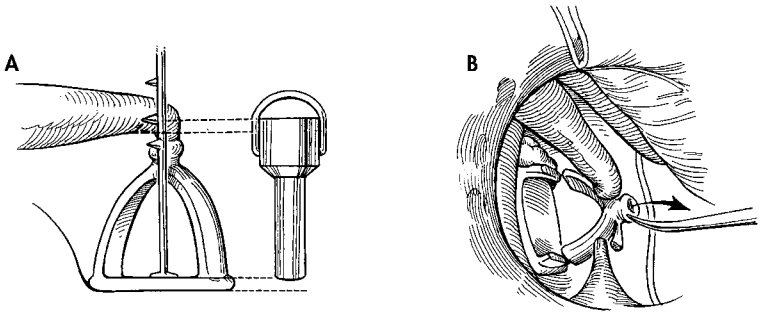


FIGURE 12–1. *A*, Proper measuring technique for a stapes prosthesis. *B*, Fracturing the superstructure of the stapes. Reproduced with permission from Coker NJ, Jenkins HA. Atlas of otologic surgery. Philadelphia: WB Saunders; 2001.

Establishing contact with the perilymphatic space may be done in several ways. The trend within the last decade has moved to smaller fenestra to protect the inner ear as much as possible. Typically, a small fenestra is created, or the footplate is partially removed. When the surgeon removes the footplate, a stapedectomy is performed. Depending on the training and preference of the otologist, the posterior half of the footplate is removed (Figure 12–2). In the technique demonstrated, the footplate is fractured in half, with the posterior portion being removed. A perichondrial graft from the tragus or a vein graft is placed over the defect, and the prosthesis is positioned over it and secured to the incus.

In the stapedotomy technique, a perforation in the footplate is made, just large enough to allow passage of the prosthesis. Figure 12–3 illustrates the technique popularized by Fisch in which a perforation is gradually enlarged with a handheld drill to 0.6 mm in diameter. The stapes replacement prosthesis of choice is placed in the perforation and attached to the incus. The length of the prosthesis used is longer (by 0.25 mm) than the distance between the medial aspect of the long process of the incus and footplate, measured earlier, since the tip of the prosthesis has to go through the footplate. The addition of fresh clotted blood to the area also helps reduce the risk of a perilymphatic fistula.

Many otologists have advocated use of a laser in performing a stapedotomy. The advantage of the laser is decreased manipulation of the suprastructure and footplate. The thermal effect is negligible.[23] The disadvantage is the extra time, expense, and instrumentation needed. Perkins and Curto popularized a combination laser stapedotomy with tissue coverage of the perforation.[24] Figure 12–4 demonstrates this

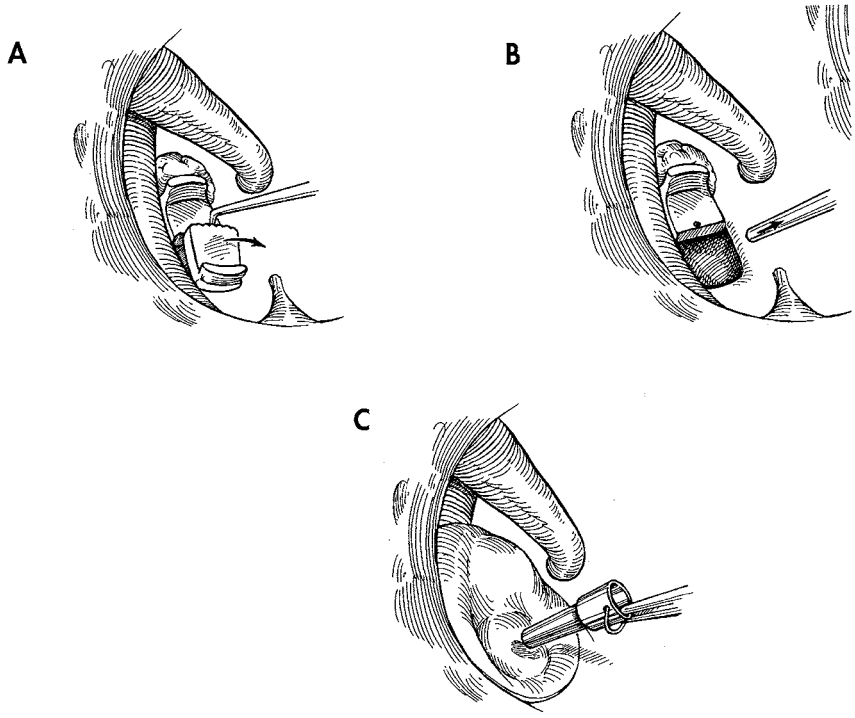


FIGURE 12–2. *A*, Partial stapedectomy with extraction of the footplate. *B*, Aspiration technique: place suction to the side of the oval window and avoid suctioning directly over the footplate. *C*, Perichondrial graft in place. Prosthesis being brought into position. Reproduced with permission from Coker NJ, Jenkins HA. Atlas of otologic surgery. Philadelphia: WB Saunders; 2001.

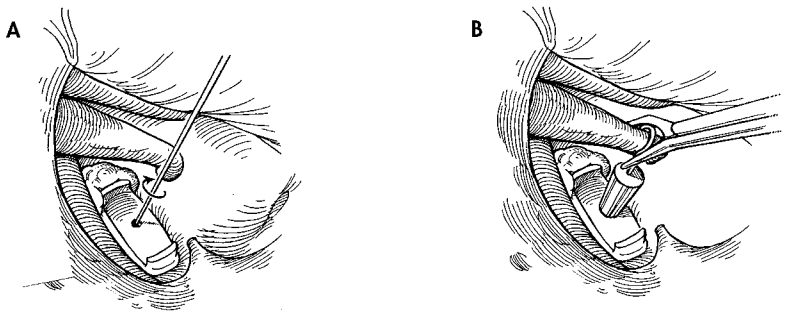


FIGURE 12–3. Stapedotomy technique. *A*, Fenestration of footplate by a graduated series of perforators. *B*, Placement of prosthesis. Reproduced with permission from Coker NJ, Jenkins HA. Atlas of otologic surgery. Philadelphia: WB Saunders; 2001.

technique. A vein graft is placed over a drilled hole in a Teflon block. A prosthesis is placed in the hole, and the vein graft is allowed to dry and become adherent to the prosthesis. Rosettes of charred bone are created by the laser and gently removed with a pick. The prosthesis and graft are positioned over the fenestra with the tip projecting into the vestibule, and the prosthesis is positioned under the incus.

Postoperative Management

Postoperatively, the patient is sent home to bed and to remain on light activity for several days. Pain management requires mild oral medication, and antibiotics are not necessary. Postoperative vertigo is treated with ondansetron (Zofran) or promethazine (Phenergan). Stool softeners help reduce straining and decrease the chance of a perilymphatic fistula formation. Patients are cautioned against blowing their noses and sneezing with their mouths closed.

Postoperative follow-up is scheduled in 1 week to remove the suture and packing and assess the integrity of the tympanic membrane. Hearing testing is done 4 to 6 weeks following the operation.

Pitfalls

The pitfalls of the surgery include inadequate exposure and anatomic variations. The endaural incision and speculum produce a narrower opening when compared to the anterior incisura approach. The scutum can cover the long process of the incus and the posterior half of the stapes. Adequate removal of the scutum ensures proper visualization of the foot-

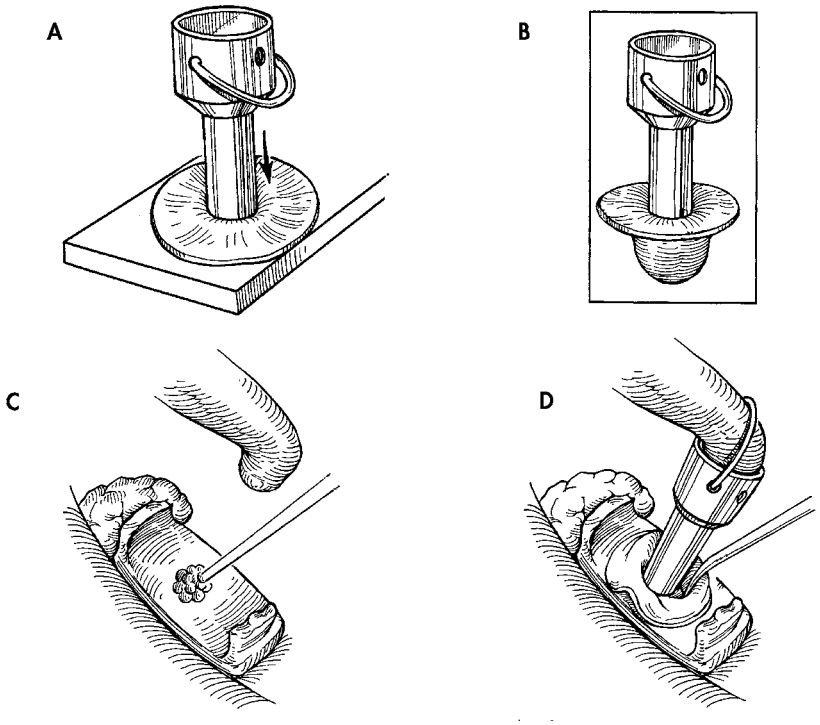


FIGURE 12–4. Stapedotomy technique with vein graft. *A*, Graft is adhered to the prosthesis by drying in a well. *B*, Adherent graft. *C*, Laser stapedotomy with rosette formation. *D*, Prosthesis and graft in place. Reproduced with permission from Coker NJ, Jenkins HA. Atlas of otologic surgery. Philadelphia: WB Saunders; 2001.

plate in the crucial stage of footplate perforation and prosthesis placement. An aberrant facial nerve may prohibit a stapedectomy, and relocation of the nerve, which carries a risk of facial weakness and paralysis, is not recommended. A dehiscent nerve has a higher chance of facial nerve injury.

Advanced otosclerosis with obliteration of the oval window requires drilling of the footplate and significant experience with temporal bone anatomy. Sclerotic obstruction of the round window is of less importance since only a small opening over the round window is necessary to allow proper cochlear function. Drilling out the round window often results in a severe hearing loss and is avoided.

If the footplate should drop into the vestibule, attempts should be made to remove it. However, this should be done only if one can easily grasp the edge and gently remove it. Never go fishing for bony fragments that descend into the vestibule, away from the annular rim.

During the stapedectomy, the protective function of the stapedius muscle is destroyed. A new technique in which the stapedius muscle is left in place and the posterior crus of the fractured suprastructure is shaped and used as an autologous stapes replacement graft has been proposed and used. The efficacy of this technique in preserving the acoustic reflex is controversial.

Results

Otosclerosis surgery has withstood the test of time since its reinstitution in the 1950s. Shea, in his review of 40 years of stapes surgery, reported closure to within 10 dB of the preoperative bone-conduction level in over 95% of patients.[25] Glasscock et al reported over 91% closure to within 5 dB.[26] Both groups reported significantly less success in revision surgery. Persson et al contrasted stapedectomy and stapedotomy in the review of their series.[27] They demonstrated that partial and total stapedectomy had better results at all frequencies with the exception of 4 kHz. However, the hearing results in this group tended to deteriorate more quickly than the stapedotomy patients. Others have reported similar preservation of high frequencies with stapedotomy.[28] Outcome in the training situation demonstrates significantly poorer results.[29] A definite learning curve in stapes surgery is present after training.[30] The accepted overall rate of anacusis following stapes surgery is in the range of 1 to 2%.[28]

Complications

The complications of the stapes surgery are immediate and delayed in nature. Immediate complications are those occurring during the operation, such as facial nerve paralysis secondary to infusion of the local anesthetic or injury, vertigo, and/or hearing loss caused by suctioning of the perilymphatic fluid from the oval window during the operation or persistent postoperative perilymphatic leakage from the oval window. Bed rest and light activities are recommended for vertiginous patients, and most recover shortly after the operation. Metallic taste and loss of taste caused by manipulation of or injury to the chorda tympani nerve are not uncommon. Labyrinthitis, although possible, occurs rarely under sterile conditions.

Delayed postoperative complications, including perilymphatic fistula,[31] granuloma,[32] and prosthesis dislocation, have been reported. Immediate treatment with antibiotics and rest is recommended. Re-exploration should be entertained to correct for the fistula should persistent vertigo occur. Granuloma formation can occur for totally unknown reasons and in the best of settings. Keeping all foreign material, such as the talc powder of the glove and bone dust, away from the footplate area may decrease the chances of granulomas. A dislocated prosthesis requires re-exploration and revision.

REFERENCES

1. Valsalva AM. Opera, hoc est, tractatus de aure humana. Venice: Pitteri; 1735.
2. Magnus A. Über Verlauf und Sektionsbefund eines Falles von hochgradiger und eigenthumlicher Gehorstorung. Arch Ohrenheilk 1876;11:244–51.
3. von Tröltsch A. Treatise on the diseases of the ear including the anatomy of the organ. New York: William Wood; 1869.
4. Politzer A. Über primäre Erkrankung der knöchernen Labyrinthkapsel. Z Ohrenheilk 1894;25:309–27.
5. Toynbee J. Diseases of the ear. Philadelphia: Blanchard and Lea; 1860.
6. Levin G, Fabian P, Stahle J. Incidence of otosclerosis. Am J Otol 1988;9:299–307.
7. Guild SR. Histologic otosclerosis. Ann Otol Rhinol Laryngol 1944;53:246–66.

8. Siebenmann F. Demonstration mikroscopischer und makroscopischer Paraparate von Otospongiosis progressiva. Papers Int Otol Congr 1912;9:207.
9. Manasse P. Neue Untersuchungen zur Otosklerosenfrage. Z Ohrenheilk 1922;82:76–96.
10. Tomek MS, Brown MR, Mani SR, et al. Localization of a gene for otosclerosis to chromosome 15q25-26. Hum Mol Genet 1988;7:285–90.
11. McKenna MJ, Kristiansen AG, Bartley ML, et al. Association of *COL1A1* and otosclerosis: evidence for a shared genetic etiology with mild osteogenesis imperfecta. Am J Otol 1998;19:604–10.
12. Khetarpal U, Schuknecht HF. In search of pathologic correlates for hearing loss and vertigo in Paget's disease. A clinical and histopathologic study of 26 temporal bones. Ann Otol Rhinol Laryngol Suppl 1990; 99:1–16.
13. Yoo TJ, Tomoda K, Stuart JM, et al. Type II collagen-induced autoimmune otospongiosis. A preliminary report. Ann Otol Rhinol Laryngol 1983;92:103–8.
14. Niedermeyer HP, Arnold W. Otosclerosis: a measles virus associated inflammatory disease. Acta Otolaryngol (Stockh) 1995;115:300–3.
15. McKenna MJ, Mills BG. Immunohistochemical evidence of measles virus antigens in active otosclerosis. Otolaryngol Head Neck Surg 1989;101:415–21.
16. Shambaugh GE, Scott A. Sodium fluoride for arrest of otosclerosis. Arch Otolaryngol 1964;80:263.
17. Kessel J. Über das Ausschneiden des Trommelfelles und Mobilisieren des Steigbügels. Arch Ohrenheilk 1877;12:237.
18. Miot C. De la mobilisation de l'étrier. Rev Laryngol 1890;7:225.
19. Siebenmann F. Sur le traitement chirurgical de la sclerose otique. Congr Int de Med Sect d'Otol 1900;13:170.
20. Lempert J. Improvement in hearing in cases of otosclerosis: a new, one stage surgical technic. AMA Arch Otolaryngol 1938;28:42.
21. Rosen S. Palpation of the stapes for fixation. AMA Arch Otolaryngol 1952;56:610.
22. Shea JJ. Fenestration of the oval window. Ann Otol Rhinol Laryngol 1958;67:932–51.
23. Coker NJ, Ator GA, Jenkins HA, et al. Carbon dioxide laser stapedotomy, thermal effects in the vestibule. Arch Otolaryngol 1985;111:601–5.
24. Perkins R, Curto FS. Laser stapedotomy: a comparative study of prostheses and seals. Laryngoscope 1992;102:1321–7.
25. Shea JJ. Forty years of stapes surgery. Am J Otol 1998;19:52–5.
26. Glasscock ME, Storper IS, Haynes D, Bohrer PS. Twenty-five years of experience with stapedectomy. Laryngoscope 1995;105:899–904.
27. Persson P, Harder H, Magnuson B. Hearing results in otosclerosis surgery after partial stapedectomy, total stapedectomy and stapedotomy. Acta Otolaryngol (Stockh) 1997;117:94–9.
28. Kursten R, Schneider B, Zrunek M. Long-term results after stapedectomy versus stapedotomy. Am J Otol 1994;15:804–6.
29. Backous DD, Coker NJ, Jenkins HA. Prospective study of resident-performed stapedectomy. Am J Otol 1993;14:451–4.
30. Hughes GB. The learning curve in stapes surgery. Laryngoscope 1991;101:1280–4.
31. Goodhill V. Variable oto-audiologic manifestations of perilymphatic fistulae. Rev Panam Otorhinolaringol Bronchoesofagol 1967;1:100–9.
32. Harris I, Weiss L. Granulomatous complications of oval window fat grafts. Laryngoscope 1962;72:870–85.

CHAPTER 13

Hereditary Hearing Impairment

Richard J. H. Smith, MD, Stephen W. Hone, MD

Hearing impairment may be classified etiologically as either inherited or acquired and temporally as either prelingual (congenital) or postlingual (late in onset). Although useful, these types of classification belie the complex interaction of genetics and environment that make the study of hearing impairment, particularly late-in-onset deafness, difficult. For example, although presbycusis and noise-induced hearing loss might be dismissed as the end result of accumulated age and environmental trauma, animal studies have shown that genetic background is an important determinate of final outcome. By focusing on congenital hereditary hearing loss, the confounding effects of environment are minimized, and our understanding of the genetics of aural development and auditory function is simplified.

Approximately 50% of congenital deafness is inherited,[1,2] and among school-aged children, 1 child in 650 to 2,000 has some form of hereditary deafness.[3,4] As the incidence of deafness owing to infectious and iatrogenic causes diminishes, and as our ability to diagnose abnormalities improves, the relative importance of hereditary factors as causes of deafness increases.

Patterns of Inheritance

Genetic information is passed from one generation to the next encoded in the human genome. The human genome is comprised of 46 chromosomes, 22 pairs of autosomes, and the sex chromosomes, XY in males and XX in females. The autosomes vary in size and can be arranged by karyotype from largest to smallest (Figures 13–1, A and B). Variances in shape are caused by the centromere, which divides a chromosome into two arms. A chromosome is described as metacentric if the centromere is in the center, submetacentric if it is off center, or acrocentric if it is near an end. The shorter of the two arms is designated "p" (for petite), and the longer arm is designated "q." Chromosomes are further labeled by the banding pattern that is produced by staining (see Figures 13–1, A and B). Band patterns are distinct for each chromosome and are individually numbered (7q31 refers to chromosome 7, long arm, band three, one).

Individuals inherit half of their autosomal chromosomes from their father and half from their mother. Therefore, every gene exists as a pair, with one copy of paternal origin and the other copy of maternal origin. Each copy is referred to as an allele. The alleles of a gene pair may be identical, or subtle differences may be present. If the alleles are identical, an individual is said to be homozygous for that gene pair; alternatively, if the alleles are different, an individual is said to be heterozygous. For example, if a gene has two possible alleles, A and A′, and an AA′ by AA′ mating occurs, the progeny will have genotypes AA, AA′, or A′A′. If normal function of this gene is essential for normal hearing, and A′ encodes an allele of the gene that is associated with hearing impairment, deafness could result. If progeny with genotypes AA′ or A′A′ are hearing impaired, one can assume that A′ is dominant over A. Alternatively, if all progeny have normal hearing except those with genotypes A′A′, one can assume that A′ is recessive with respect to A. In the first case, both parents will be hearing impaired, whereas in the second case, only A′A′ progeny will be hearing impaired. These patterns of allele segregation are referred to as autosomal dominant and autosomal recessive inheritance, respectively (Figures 13–2, A and B).

Deafness caused by genes on the X chromosome is usually inherited as an X-linked recessive trait (Figure 13–2, C). The deafness is rarely penetrant in a carrier female, but half of all sons are affected, and half of all daughters are carriers. If the affected gene is inherited through the father, there are no affected offspring, although all daughters are carriers. Mitochon-

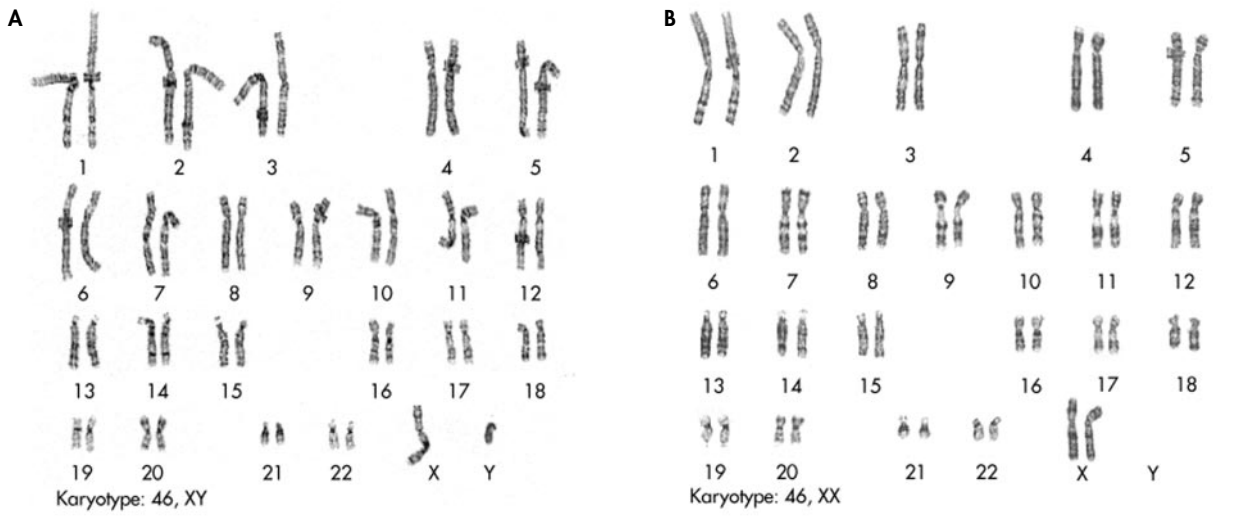


FIGURE 13–1. *A*, Male karyotype, 23,XY. *B*, Female karotype, 23,XX.

drial deafness is inherited only through the mother,[5,6] and sons and daughters may be affected equally.

When hereditary deafness is classified by mode of inheritance, 60 to 70% of cases are autosomal recessive, 20 to 30% are autosomal dominant, and 2% are X-linked.[2,7] In nearly 33% of cases, other phenotypic characteristics cosegregate with the hearing loss. These types of hearing impairment are known as "syndromic" and make the unequivocal diagnosis of hereditary deafness much easier. Typically, a wide range of phenotypes occurs, even in individuals carrying the same deafness-causing genetic mutation, a phenomenon known as variable expressivity. An affected individual may exhibit a few, some, or all of the phenotypic manifestations typically associated with a particular genetic abnormality. On rare occasions, an individual may have no abnormal physical findings, and the genetic mutation is said to be nonpenetrant. In the absence of cosegregating physical findings, inherited deafness is said to be "nonsyndromic." Nonsyndromic deafness is subclassified by mode of inheritance as DFNA, DFNB, or DFN for dominant, recessive, or X-linked, respectively (DFN for *deafness*). Loci are numbered in order of discovery. Mitochondrial deafness is designated by mutation type.

SYNDROMIC HEARING IMPAIRMENT

Hearing loss has been described in over 400 syndromes.[8] Although there are many classifications for these syndromes, one of the more useful is based on the involved organ system (Table 13–1). Some of the more common forms of syndromic hearing impairment are discussed in more detail with particular reference to recent advances in the understanding of their genetic basis (Table 13–2).

Autosomal Dominant Syndromic Hearing Impairment

Branchio-Oto-Renal Syndrome Although the association of branchial arch anomalies with hearing impairment has long been recognized, it was not until 1975 that Melnick and colleagues described branchio-oto-renal (BOR) syndrome.[9] Affected persons have branchial clefts or fistulae, otologic abnormalities, and renal anomalies (Figure 13–3). The pattern of inheritance is consistent with autosomal dominant transmission, and although gene penetrance approaches 100%, there is considerable variation in expression. Disease prevalence approximates 1 in 40,000 in the general population, but in the Deaf community, BOR syndrome may be responsible for over 2% of profound deafness.[10]

Hearing impairment is the single most common trait of BOR and is found in over 90% of affected individuals. The loss can be conductive, sensorineural, or mixed, with differences occurring even within families. Age of onset varies from early childhood to young adulthood, and both stable and progressive hearing loss have been reported.[11–14]

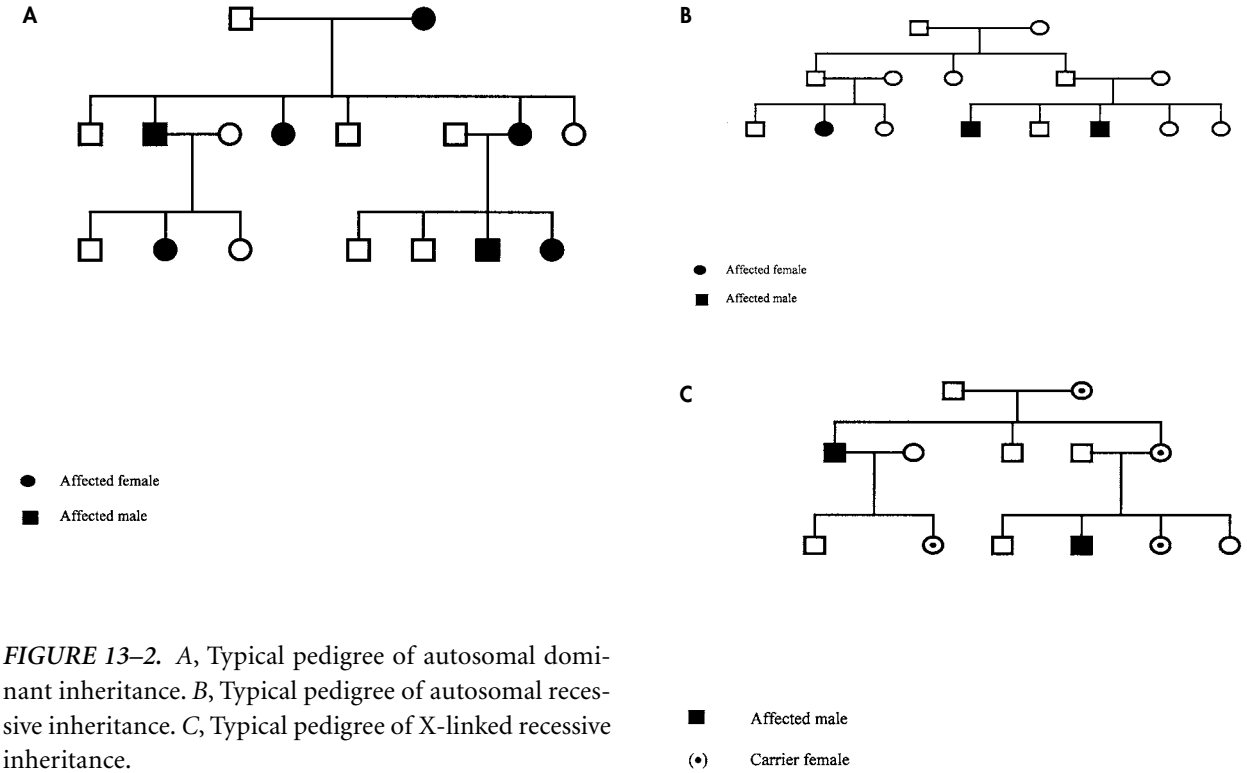


FIGURE 13–2. *A*, Typical pedigree of autosomal dominant inheritance. *B*, Typical pedigree of autosomal recessive inheritance. *C*, Typical pedigree of X-linked recessive inheritance.

Progressive loss is likely to be attributable to associated temporal bone abnormalities such as a dilated vestibular aqueduct (DVA).[13] Other associated temporal bone abnormalities include cochlear hypoplasia, bulbous internal auditory canals, deep posterior fossae, acutely angled promontories, and anomalous facial nerve course. Table 13–3 outlines the most common features of BOR syndrome. Early studies probably underestimated the true incidence of renal anomalies, which occur in up to 75% of cases and include renal agenesis, renal hypoplasia, renal dysplasia, pelvic ureteral junction obstruction, and polycystic kidneys. Glomerular lesions and vesicoureteral reflux progressing to renal failure have been described as well.

In the early 1990s, the BOR gene was mapped to chromosome 8q, and in 1997, the causative gene was cloned.[15–17] This gene is a human homologue of the *Drosophila* eyes absent gene (*eya*) and has been called *EYA1*. Its expression pattern suggests a role in the development of all components of the inner ear. In the kidney, murine *Eya1* plays a role in the development of the metanephric cells surrounding the "just-divided" ureteric branches.

A variety of missense and nonsense mutations, as well as insertions and deletions, have been identified in families with BOR syndrome; however, in about 70% of persons with a BOR syndrome phenotype, *EYA1* mutations cannot be found. This finding, together with marked phenotypic variability, has suggested to some investigators that BOR syndrome is a heterogeneous disease, a hypothesis recently confirmed with the identification of a second BOR syndrome locus on chromosome 1p31.[18]

Crouzon's Disease Crouzon's disease (CD) is another autosomal dominant type of syndromic hearing impairment, although one-third of cases are attributable to new mutations. It is diagnosed in about 5% of newborns with craniosynostosis and occurs with an estimated incidence of 1.65 in 100,000 births.[19] In addition to craniosynostosis, it is characterized by hypertelorism, midface hypoplasia, and exophthalmos.[20] One-third of affected persons have conductive hearing loss secondary to external or middle ear abnormalities. Frequently, there is an associated sensorineural component.[21]

Mutations in fibroblast growth factor receptor (*FGFR*) genes are implicated in a number of craniosynostosis syndromes including CD.[22] These genes

TABLE 13–1. Classification of Syndromic Hearing Impairment

System Involved	*Examples*	*Phenotypic Features*
Cardiac	Jervell and Lange-Nielsen syndrome	Prolonged QT interval, sudden death
Craniofacial/cervical	Apert's syndrome	Craniosynostosis, syndactyly
	Goldenhar's syndrome	Oculoauriculovertebral dysplasia
	Crouzon's disease	Craniosynostosis
	Treacher Collins syndrmoe	Mandibulofacial dysostosis
	Pierre Robin syndrome	Micro-/retrognathia, cleft palate
Chromosomal abnormalities	Down syndrome (trisomy 21)	Flat facial profile, oblique palpebral fissures, mental retardation
	Turner's syndrome	Short stature, amenorrhea, webbed neck
Endocrine	Pendred's syndrome	Goiter
	Diabetes mellitus and deafness	Diabetes mellitus
Integumentary	Waardenburg's syndrome	Dystopia canthorum, pigmentary anomalies of hair and skin
	Neurofibromatosis 1	Café-au-lait spots, neurofibromas
	Neurofibromatosis 2	Bilateral acoustic schwannomas
Metabolic	Hurler's syndrome	Dwarfism, hepatosplenomegaly, corneal clouding
	Hunter's syndrome	Dwarfism, hepatosplenomegaly
Nervous	Charcot-Marie-Tooth disease	Progressive neuropathic (peroneal) muscular atrophy
Ocular	Usher's syndrome	Retinitis pigmentosa
	Alström syndrome	Retinal degeneration, diabetes mellitus, infantile obesity
Renal	Alport's syndrome	Hereditary nephritis
	Branchio-oto-renal syndrome	Branchial and renal anomalies
Skeletal abnormalities	Osteogenesis imperfecta	Blue sclera, fractures

are tyrosine kinases that span the cell membrane and are important in mitogenesis and cell migration, development, and differentiation. *FRGR2* mutations cause CD.[23]

Neurofibromatosis 2 Neurofibromatosis 2 (NF2) is a central form of neurofibromatosis characterized by bilateral vestibular schwannomas. Other features include meningiomas, spinal cord dorsal root schwannomas, and posterior subcapsular cataracts. Criteria for diagnosis include one of the following[24]: (1) bilateral internal auditory canal/cerebellopontine angle tumors; (2) a first-degree relative with NF2 and a unilateral eighth nerve tumor; (3) a first-degree relative with NF2 and two of the following: neurofibroma, meningioma, glioma, schwannoma, or juvenile posterior subcapsular lenticular opacity.

The causative gene encodes a protein called merlin that shows similarity to a group of cell membrane-cytoskeleton protein linkers that regulate cell adhesion and morphogenesis. Inactivation of merlin in the mouse by targeted mutagenesis produces a variety of malignant tumors with a high rate of metastasis, suggesting that merlin also functions as a

TABLE 13–2. Selected Causes of Syndromic Hearing Impairment with Genetic Features

Syndrome/Disease	*Locus*	*Gene*	*Function of Encoded Protein*
Alport's	Xq22	*COL4A5*	Specific components of glomerular basement membrane within the kidney; in the cochlea, they are found in the basilar membrane, parts of the spiral ligament, and the stria vascularis
	2q36-2q37	*COL4A3*	
	2q36-2q37	*COL4A4*	
Branchio-oto-renal			
BOR1	8q13.3	*EYA1*	Human homologue of *Drosophila* eyes absent gene; plays a role in the development of all components of the inner ear
BOR2	1q31	Unknown	
Crouzon's	10q25-26	*FGFR2*	Member of the tyrosine kinase receptor superfamily; has high affinity for peptides that signal the transduction pathway for mitogenesis, cellular differentiation, and embryogenesis
Jervell and Lange-Nielsen			
JLNS1	11p15.5	*KVLQT1*	Subunits of a voltage-gated potassium channel protein; important for endolymph homeostasis
JLNS2	21q22.1-q22.2	*KCNE1*	
Neurofibromatosis			
NF2	22q12	*NF2*	Merlin, a tumor suppressor
Pendred's	7q21-34	*PDS*	Chloride-iodide transporter
Stickler's			
STL1	12q13.11-13.2	*COL2A1*	Fibrillar collagens arrayed in a quarter-staggered fashion; extracellular matrix components
STL2	6p21.3	*COL11A2*	
STL3	1p21	*COL11A1*	
Treacher Collins	5q32-q33.1	*TCOF1*	Highly phosphorylated nucleolar protein; a nuclear transcription factor
Usher's			
USH1A	14q32	Unknown	
USH1B	11q13.5	*MYO7A*	Unconventional myosin—moves actin filaments using actin-activated adenosine triphosphatase; maintains stereocilia integrity; present in inner and outer hair cells
USH1C	11p15.1	*USH1C*	Harmonin, which may function as a rafting protein in gating complexes in the stereocilia
USH1D	10q	*CDH23*	Important for the formation of tight junctions

Continued

TABLE 13–2. Selected Causes of Syndromic Hearing Impairment with Genetic Features—Continued

Syndrome/Disease	*Locus*	*Gene*	*Function of Encoded Protein*
USH1E	21q	Unknown	
USH1F	10	Unknown	
USH2A	1q41	*USH2A*	Usherin a possible component of basal lamina and extracellular matrices; may be involved in cellular adhesion
USH2B	3p23-24.2	Unknown	
USH2C	5q14.3-q21.3	Unknown	
USH3	3q21-q25	Unknown	
Waardenburg's			
WS1	2q35	*PAX3*	A DNA-binding protein that is believed to regulate the expression of other genes; mutations result in neural crest–derived melanocyte deficiency
WS2	3p14.1-p12.3	*MITF*	Homodimeric transcription factor
WS3	2q35	*PAX3*	As above
WS4	13q22	*EDNRB*	A receptor involved in the formation of an endothelin signaling pathway
WS4	20q13.2-q13.3	*EDN3*	A transcription factor
WS4	22q13	*SOX10*	Required for development of early neural crest–derived progenitor cells

Adapted from Van Camp G and Smith RJH.[25]

tumor suppressor.[26,27] A variety of mutations have been described, and it appears that there is some genotype-phenotype correlation as truncating or inactivating mutations lead to severe phenotypes with an earlier age of onset, whereas missense mutations are associated with milder disease and a later age of onset.[28,29]

Stickler's Syndrome Stickler (STL) syndrome, also known as hereditary arthro-ophthalmopathy, is characterized by marfanoid features, spondyloepiphyseal dysplasia, joint hypermobility, midface hypoplasia, severe myopia, and varying degrees of Robin sequence (cleft palate, micrognathia, and glossoptosis). Because of the possibility of retinal detachment, ophthalmologic assessment is mandatory. Approximately 15% of affected individuals have a mixed hearing loss.[30–32] Gene linkage studies have demonstrated considerable genetic heterogeneity, with mutations in *COL2A1*, *COL11A2*, and *COL11A1* implicated in STL1, STL2, and STL3, respectively.[33–35] Because *COL11A2* is not expressed in the eye, persons affected with STL2 do not have myopia.

Treacher Collins Syndrome Treacher Collins syndrome (TCS) is a disorder of craniofacial development affecting structures derived from the first branchial arch. It is characterized by midface hypoplasia, micrognathia, macrostomia, colobomas of the lower eyelids, downward slanting palpebral fissures, cleft palate, and conductive hearing loss owing to external and middle ear abnormalities.[36] Inner ear abnormalities are rare, although enlargement of the utricle and aplasia of the horizontal canal have been reported.[37] The reported incidence of TCS is about 1 in 50,000 live births.[38]

The causative gene, *TCOF1*, encodes a protein called treacle that is structurally related to nucleolar phosphoproteins and may play a role in nucleolar-cytoplasmic transport.[39] Up to 60% of affected per-

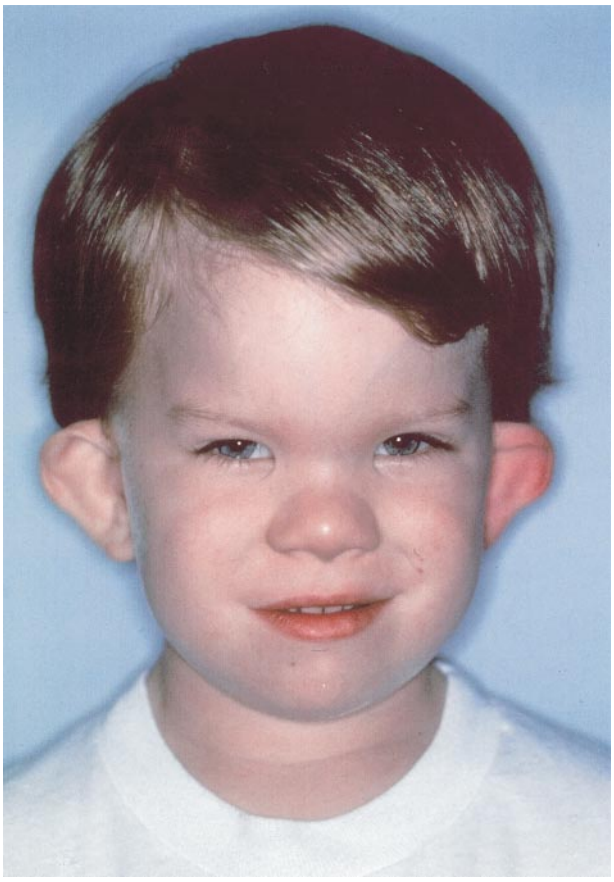


FIGURE 13–3. Child with branchio-oto-renal syndrome. Note failure of development of the antihelical folds resulting in prominent ears.

TABLE 13–3. Phenotypic Manifestations of Branchio-Oto-Renal Syndrome

Phenotypic Anomaly	*Percentage Affected*
Hearing loss	73–93
Conductive	23–33*
Sensorineural	17–29*
Mixed	35–52*
Preauricular pits	70–82
Branchial fistulae	49–84
Renal anomalies	9–75
Pinna anomalies	36–62
External auditory canal stenosis	2–29
Preauricular tags	8–13
Lacrimal duct aplasia	5–11
Retrognathia	4–16
Cleft palate	2–5

*Percentage affected in those with hearing loss.

sons have *TCOF1* mutations, the majority of which result in premature stop codons.[40] Its role in the inner ear is unknown.

Waardenburg's Syndrome Waardenburg's syndrome (WS), an auditory-pigmentary syndrome caused by the scattered absence of melanocytes from the skin, hair, eyes, and stria vascularis, was first described by Petrus Waardenburg in 1951.[41] He estimated the prevalence of this syndrome in the general population to be 1 in 42,000 and 1.43% among the congenitally deaf. Four clinical types are recognized: WS1, WS2, Klein-Waardenburg syndrome (WS3), and Shah-Waardenburg syndrome (WS4) (see Table 13–2). Klein-Waardenburg syndrome is a severe variant of WS1 with associated limb and skeletal abnormalities. The features of WS2, together with Hirschsprung's disease, characterize Shah-Waardenburg syndrome.[42]

The presence of dystopia canthorum in WS1 distinguishes it from WS2. Formal clinical criteria have been adopted to diagnose WS1, and criteria have been suggested for WS2, although its clinical definition covers any auditory-pigmentary syndrome that cannot be clearly classified (Table 13–4).[43,44] As such, WS2 includes a mixed collection of melanocyte defects and is likely to exhibit considerable genetic heterogeneity. Apart from dystopia canthorum, all features of WS1 and WS2 show marked interfamilial and intrafamilial variability.

Dystopia canthorum is the most common feature of WS1 and results from fusion of the eyelids medially leading to a reduction in the visible sclera medial to the iris (Figure 13–4). Hearing loss occurs in over 60% of cases of WS1 and in over 80% of cases of WS2.[45] The loss is typically sensorineural, prelingual, and nonprogressive and varies from mild to profound; profound bilateral loss is most common.

Both WS1 and WS3 are caused by mutations in *PAX3*, some cases of WS2 are associated with mutations in the microphthalmia gene (*MITF*), and the WS4 phenotype results from mutations in *EDNRB*,

its ligand *EDN3*, or *SOX10* (see Table 13-2). Recent studies have confirmed a complex interaction between these genes, with SOX10 and PAX3 proteins synergistically activating *MITF* expression.[46–51] Numerous different allele variants in these genes have been reported as causing deafness.

Autosomal Recessive Syndromic Hearing Impairment

Jervell and Lange-Nielsen Syndrome Jervell and Lange-Nielsen syndrome (JLNS) is characterized by profound prelingual sensorineural hearing loss, syncope, and sudden death owing to a prolonged QT interval. The syndrome was first described in Norway, where its estimated prevalence is 1 in 200,000. Diagnostic criteria include a QTc > 440 ms in males and > 460 ms in females.[52] Syncopal attacks are usually associated with exertion or emotion, and with prompt diagnosis and antiarrhythmic treatment, the high mortality rate can be significantly reduced.

Mutations in *KCNQ1* and *KCNE1,* which encode subunits of a voltage-gated potassium channel protein, have been shown to cause JLNS. However, there is likely to be considerable genetic heterogeneity as mutations have not been identified in a number of families. Some heterozygous individuals may have a prolonged QT interval in the absence of hearing loss and are prone to life-threatening arrhythmias. Genetic counseling should be offered to affected persons and their families to minimize potential morbidity and mortality.[52]

Pendred's Syndrome Pendred's syndrome (PS) is characterized by congenital sensorineural hearing loss and goiter. Prevalence estimates range from 1 to 7.5 per 100,000 newborns, suggesting that it may account for up to 7.5% of all childhood deafness.[53,54] The degree of hearing loss is variable but is most frequently profound and is associated with temporal bone abnormalities ranging from a DVA to Mondini's dysplasia (Figure 13–5). Goiter may be apparent at birth but typically presents in mid-childhood. The thyroid defect involves organification of iodine and can be diagnosed by administering perchlorate, which releases unbound iodide from thyroid follicular cells. Despite this abnormality, affected individuals usually remain euthyroid.

The disease is caused by mutations in *PDS*, a member of the solute carrier 26 gene family. The encoded protein pendrin transports chloride and iodide and mediates the exchange of chloride and formate, properties that suggest tissue-specific function.[55] In the thyroid gland, pendrin has been immunolocalized to the apical membrane of the thyrocyte, where it may allow intracellular iodide to pass into the colloid space to be bound to thyroglobulin. In the kidney, pendrin probably functions as a chloride/formate exchanger, which is important for chloride transport in the proximal tubule. The role of pendrin in the inner ear is unknown.

TABLE 13-4. Diagnostic Criteria for Waardenburg's Syndrome Type 1

*Major Criteria**	*Minor Criteria*
Sensorineural hearing loss	Congenital leukoderma (areas of hypopigmented skin)
Pigmentary disturbance of the iris	Synophrys or medial eyebrow flare
Complete heterochromia	
Partial or segmental heterochromia	
Hypoplastic blue eyes	
Hair hypopigmentation (white forelock)	Broad, high nasal root
Dystopia canthorum (W>1.95)	Alae nasi hypoplasia
Affected first-degree relative	Premature graying of hair (before 30 y)

*An affected individual must have at least two major criteria or one major criterion and two minor criteria. Criteria for Waardenburg's syndrome type 2 have been suggested that include premature graying as a major criterion instead of dystopia canthorum.
W = Waardenburg index.

FIGURE 13–4. Mother and child with Waardenburg's syndrome (WS). Note the characteristic dystopia canthorum of WS1. The child exhibits heterochromia iridis (see color figure on the CD-ROM).

Over 45 different deafness-causing *PDS* mutations have been reported, and although many families have "private mutations," in about half of affected persons, one of three different mutations is found (L236P, T416P, 1001+1G→A).[56,57] *PDS* allele variants can be identified in approximately 80% of multiplex families segregating for either DVA or Mondini's dysplasia, suggesting that mutations in this gene are the major genetic cause of these two temporal bone anomalies.

Usher's Syndrome Usher syndrome (USH) is the most common autosomal recessive syndromic form of hearing impairment. Characterized by blindness caused by retinitis pigmentosa and sensorineural hearing loss, it is responsible for half of all deaf-blindness in the United States and an estimated 3 to 10% of all congenital deafness.[58] It can be classified into three different types based on clinical presentation (Table 13–5). USH1 and USH2 are most common, whereas USH3 is quite rare, accounting for only 5 to 15% of all USH.[59,60]

Vestibular dysfunction differentiates USH1 from USH2. Persons with USH1 have vestibular areflexia and, as infants, fail to achieve normal motor developmental milestones.[61] They are late in sitting and walking and have absent ice-water responses to caloric testing. Hearing loss is typically profound, and hearing aids are frequently ineffectual as a habilitation option. Accordingly, most persons with USH1 integrate into the Deaf community. Persons with USH2, in contrast, have normal vestibular function and usually have a moderate-to-severe hearing loss. They use hearing aids effectively and communicate orally. Persons with USH3 have progressive vestibular and auditory dysfunction.

Initial visual problems with USH begin as nyctalopia or night blindness. This problem can occur in the preschool years, although visual acuity usually remains good until the third decade. By the fifth decade, 40% of persons with USH are blind; by the seventh decade, this figure increases to 75%.[62,63] Constriction of the visual fields accompanies the loss in visual acuity. Studies of multiplex families have documented considerable intrafamilial variation in the rate and degree of visual deterioration. Electroretinography (ERG) may uncover early retinitis pigmentosa and should be considered in all cases of congenital deafness to identify early USH.[64]

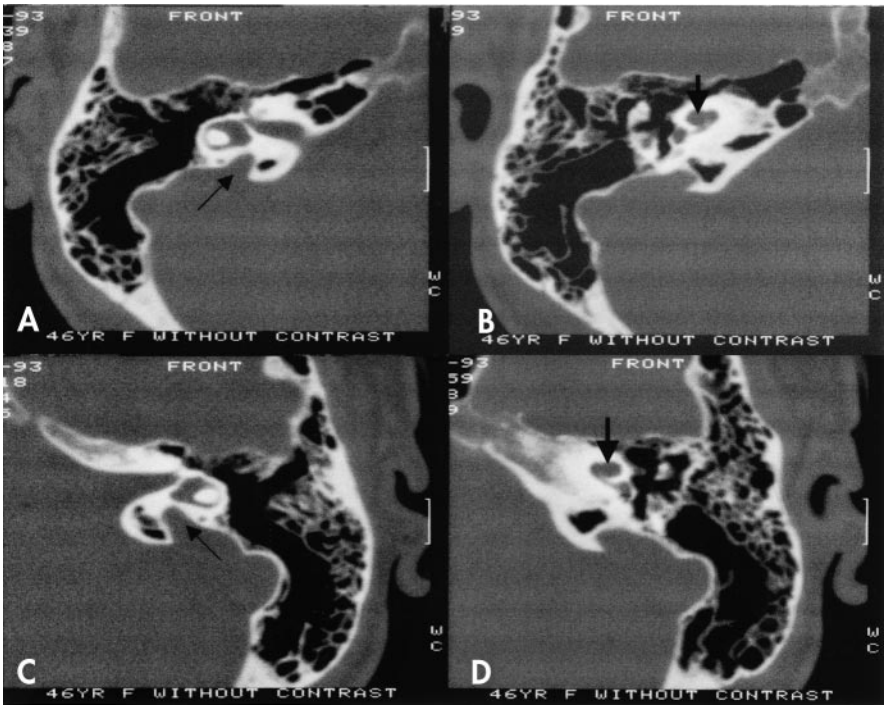


FIGURE 13–5. Axial computed tomograms of temporal bones in a patient with Pendred's syndrome (*A* and *B*, left temporal bone; *C* and *D*, right temporal bone), illustrating dilated vestibular aqueducts (*small arrows*) and Mondini's dysplasia (*large arrows*). The normal two and a half turns of the cochlea are replaced by one and a half turns in Mondini's dysplasia.

Usher's syndrome demonstrates considerable genetic heterogeneity. To date, seven USH1 loci (USH1A-1G), three USH2 loci (USH2A-C), and one USH3 locus have been identified.[65–70] Five of the relevant genes have been cloned, *MYO7A*, *USH1C*, *USH2A*, and *CDH23*, and PCDH1S mutations in which cause USH1B, USH1C, USH2A, and USH1D, and USH1F respectively.[71–73] *MYO7A* is an unconventional myosin expressed mainly in inner and outer hair cells. In the eye, it is localized to microvilli projections in retinal pigmentary epithelial cells and photoreceptor cells. Because there are no mutational hotspots, mutation detection in at-risk families requires considerable work. Mutations in this gene also cause DFNB2 and DFNA11.

The *USH1C* gene encodes a PDZ domain–containing protein called harmonin, which may function as a rafting protein in gating complexes in the stereocilia.[72] Several different mutations have been found in USH1C patients. The *USH2A* gene encodes a protein designated usherin, which has both laminin epidermal growth factor and fibronectin type III domains.[74] The 2299delG-allele variant is the most frequent mutation and is found in about 15% of persons with USH2A. The USH1D gene *CDH23* is a member of the cadherin gene family, and its encoded protein is important for the formation of tight junctions. Mutations in this gene also cause DFNB12.

X-Linked Syndromic Hearing Impairment

Alport's Syndrome Progressive glomerulonephritis, sensorineural hearing loss, and specific eye findings characterize Alport's syndrome (AS), which can be inherited as an X-linked or autosomal disorder. To facilitate diagnosis, four clinical criteria were established in 1988.[75] In the presence of unexplained hematuria, a person can be considered affected if three of the following criteria are met: (1) a positive family history of hematuria or chronic renal failure; (2) electron microscopic renal biopsy evidence of AS; (3) characteristic eye signs of anterior lenticonus, white macular flecks, or both; (4) high-frequency sensorineural hearing loss.

Since light microscopy of renal biopsy specimens is usually normal in children and because findings are nonspecific in adults, electron microscopy is essential. This degree of resolution reveals splitting of segmental areas of glomerular basement membrane, accompanied by thickening and electron-lucent areas containing dense granulations. Serial biopsies demonstrate progressive deterioration.[76] The eye findings characteristic of AS are rarely noted in child-

TABLE 13–5. Clinical Characteristics of Usher's Syndrome

Type	*Hearing Loss*	*Vestibular Response*	*Onset of Retinitis Pigmentosa*
I	Profound HI from birth	Absent	First decade
II	Moderate HI from birth	Normal	First or second decade
III	Progressive HI	Variable	Variable

HI = hearing impairment.

hood and may become apparent only with renal failure. These include anterior lenticonus (conical projection of the anterior surface of the lens), macular flecks, and peripheral coalescing flecks. The progressive myopia caused by anterior lenticonus is considered by some authors to be sufficient to diagnose AS.[77] Hearing loss is postlingual, progressive, and sensorineural.

The typical male with X-linked AS presents with hematuria at age 3 to 4 years, often following an upper respiratory tract infection. Toward the end of the first decade, hearing loss is detectable, and in the mid-teens, hypertension develops. By 25 years of age, over 90% of affected males have abnormal renal function.[76] The clinical course in female carriers is much more variable. Most are clinically asymptomatic through life. Although nearly all have evidence of microscopic hematuria, about one-third will have macroscopic hematuria. One-third will develop hypertension, whereas the risk of chronic renal failure may be as high as 15%.

X-linked AS is caused by mutations in *COL4A5*, a member of the type IV collagen gene family.[78] In a comprehensive review of the type IV collagen mutations, Lemmink and colleagues reported more than 160 different AS-causing mutations in *COL4A5*.[79] The mutation spectrum is broad; about 15% of affected males have large *COL4A5* deletions, and in 30%, a variety of missense and nonsense mutations are found. In many patients, however, no mutations are identified.

Mitochondrial Syndromic Hearing Impairment

Hearing loss may be associated with a number of syndromic mitochondrial diseases. Most frequent are the acquired mitochondrial neuromuscular syndromes and maternally inherited diabetes mellitus associated with deafness.

Kearns-Sayre Syndrome Kearns-Sayre syndrome is a multisystem disorder characterized by progressive ophthalmoplegia, pigmentary retinopathy, cardiac conduction abnormalities, and cerebellar ataxia. It usually presents in the second decade. Mixed hearing loss may be an additional feature.[25]

Maternally Inherited Diabetes and Deafness Studies of diabetic patients from different populations have confirmed an association between diabetes mellitus, sensorineural hearing loss, and mitochondrial mutations.[80,81] The hearing loss and the diabetes are felt to result from an impairment of oxidative phosphorylation. The most frequently found mutation, a 3243 A-to-G transition, is present in 2 to 6% of diabetic patients in Japan.[81] Over 60% of these patients have hearing loss, which usually develops after the onset of their diabetes. The same mutation is also found in MELAS.

MELAS Syndrome (Mitochondrial Encephalopathy, Lactic Acidosis, and Stroke-Like Episodes) MELAS syndrome is a childhood disease characterized by intermittent vomiting, proximal limb weakness, and recurrent cerebral insults that resemble strokes and cause hemiparesis and cortical blindness. MELAS syndrome is frequently associated with short stature. Hearing loss occurs in approximately 30% of affected individuals.[25]

MERRF Syndrome (Myoclonic Epilepsy and Ragged Red Fibers) Myoclonus, epilepsy, and ataxia characterize MERRF. Dementia, optic atrophy, and hearing loss frequently occur. The degree of hearing loss is variable.[25]

NONSYNDROMIC HEARING IMPAIRMENT

Over 70 different nonsyndromic hearing impairment loci have been mapped, and a number of the relevant genes have been cloned. The protein prod-

ucts of these genes include ion channels, membrane proteins, transcription factors, and structural proteins.

Autosomal Dominant Nonsyndromic Hearing Impairment

Autosomal dominant modes of inheritance account for 15% of cases of nonsyndromic deafness (ADNSHI).[82] The typical phenotype is one of postlingual hearing loss that starts in the second to third decades of life and progresses until it is moderate to severe in degree. However, the frequencies that are initially affected vary. For example, DFNA1, DFNA6, DFN14, and DFNA15 are characterized by a low-frequency hearing loss that progresses to involve the remaining frequencies. With other loci, hearing loss starts in the mid- or high frequencies before progressing. The DFNA3, DFNA12, and DFNA23 phenotypes are exceptional as they are congenital hearing losses on which age-related changes become superimposed. Many of the genes for ADNSHI have now been cloned (Table 13–6).[25]

Autosomal Recessive Nonsyndromic Hearing Impairment

Up to 85% of cases of nonsyndromic hearing loss are inherited in a recessive mendelian fashion (ARNSHI).[83] The typical phenotype is more severe than in ADNSHI and accounts for the majority of cases of congenital profound deafness. However, DFNB8 is an exception as it presents as a postlingual progressive hearing loss. To date, 30 ARNSHI loci have been identified and 8 genes cloned (see Table 13–6).[25] The first locus was published in 1994, 6 years after the first X-linked locus and 2 years after the first dominant locus were reported.[84] Three years later, the gene responsible for DFNB1, *GJB2* (gap junction beta 2), was discovered.[85] This gene encodes connexin 26 (Cx26), one of a class of proteins involved in gap junction formation. A group of six connexins oligomerize to form a torus-like structure called a connexon. Two connexons dock to form a transmembrane channel that links neighboring cells and facilitates the exchange of molecules up to 1 kD in size.[86] Immunohistochemical studies have demonstrated Cx26 expression in two groups of cochlear cells, supporting cells in the sensory epithelium and fibroblasts. It is postulated that Cx26 gap junctions allow potassium ions that enter hair cell stereocilia during mechanoelectrical transduction to be recirculated into the stria vascularis.[85,87]

The most significant discovery in the field of genetic deafness to date has been the finding that mutations in Cx26 are responsible for over half of moderate-to-profound congenital deafness in many world populations.[88,89] In certain regions of the Mediterranean, the prevalence may be as high as 79%, although studies in India and Pakistan reveal a much lower incidence.[90,91] Numerous different deafness-causing allele variants have been identified. In the United States and much of northern Europe, the most prevalent mutation is the deletion of a single guanine nucleotide from a sequence of six guanines at positions 30 to 35. Referred to as the 35delG mutation, this shift in codon reading frame results in premature termination of translation.[88] In the midwestern United States, the carrier rate for this mutation is 2.5%, whereas the carrier rate for all deafness-causing Cx26 mutations is 3%.[92] In the Mediterranean population, the carrier rate is even higher and approaches 5%. Other "common" mutations are found in different populations, such as the 167delT mutation in Ashkenazi Jews.[93]

These discoveries have had immediate application to clinical practice. In many populations, a definitive diagnosis can now be made in 50% of cases of suspected hereditary hearing loss. This ability to establish causality affects recurrence risk estimates and makes genetic counseling an important part of the evaluation of hereditary deafness. However, the degree of hearing loss in a child with Cx26-related deafness cannot be used to predict the degree of deafness in another offspring. There can be enough intrafamilial variability to make one affected child a candidate for cochlear implantation, whereas a sibling effectively uses hearing aids[94] (Figure 13–6).

X-linked Nonsyndromic Hearing Impairment

X-linked nonsyndromic hearing impairment is rare and makes up only 1 to 3% of nonsyndromic hearing loss. It exhibits considerable phenotypic heterogeneity, but most affected males have a congenital hearing loss, which can vary from severe to profound. Hearing loss is mild to moderate in carrier females. The losses associated with DFN1 and DFN6 may be progressive.[83]

TABLE 13–6. Nonsyndromic Hearing Impairment Genes Cloned to Date

Locus	*Location*	*Gene*	*Function of Encoded Protein*
DFNA1	5q31	*HDIA1*	Regulates polymerization of actin, a major component of the cytoskeleton of inner ear hair cells
DFNA2	1p34	*GJB3*	Connexin 31, a gap junction protein important for intercellular communication
		KCNQ4	Forms a functional potassium channel; found only in outer hair cells
DFNA3	13q12	*GJB2* *GJB6*	Gap junction proteins important for intercellular communication
DFNA5	7p15	*DFNA5*	Function unknown
DFNA9	14q12-q13	*COCH*	Probable extracellular matrix protein; function unknown
DFNA11	11q12.3-q21	*MYO7A*	Moves actin filaments using actin-activated adenosine triphosphatase, maintains stereocilia integrity; present in inner and outer hair cells
DFNA12	11q22-q24	*TECTA*	Interacts with β-tectorin to form noncollagenous matrix of tectorial membrane
DFNA13	6p21	*COL11A2*	An α-chain polypeptide subunit of type XI collagen, a minor fibrillar collagen; cochlear function unknown
DFNA15	5q31	*POU4F3*	Transcription factors; developmental regulators for determination of cellular phenotypes; expressed only in hair cells
DFNB1	13q12	*GJB2*	As above
DFNB2	11q13.5	*MYO7A*	As above
DFNB3	17p11.2	*MYO15*	Constitute a new class of myosins designated myosin XV; necessary for actin organization in hair cells
DFNB4	7q31	*PDS*	Chloride-iodide transport protein
DFNB9	2p22-p23	*OTOF*	Involved in calcium ion–triggered synaptic vesicle-plasma membrane fusion
DFNB10	21q22.3	*TMPRSS3*	Transmembrane serine protease; cochlear function unknown
DFNB21	11q	*TECTA*	See above
DFN1	Xq22	*DDP*	Probably evolutionarily conserved novel polypeptide necessary for normal human neurologic development
DFN3	Xq21.1	*POU3F4*	Transcription factors; expressed in the mesenchyma of the inner and middle ear

Adapted from Van Camp G and Smith RJH.[25]

Mitochondrial Nonsyndromic Hearing Impairment

Mitochondrial mutations may play a role in age-related hearing loss and have been implicated in a type of nonsyndromic deafness associated with increased susceptibility to aminoglycoside ototoxicity.[95]

Age-Related Hearing Loss Mitochondrial mutations may be a contributing factor to age-related

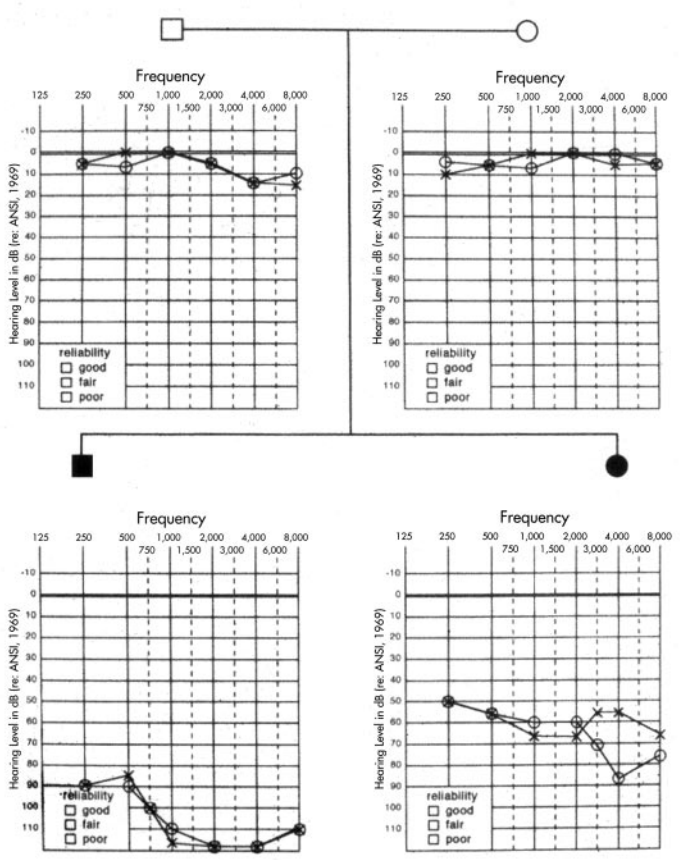


FIGURE 13–6. A pedigree of two siblings homozygous for the 35delG mutation in *GJB2*. Both parents have normal hearing. One sibling exhibits a profound sensorineural hearing loss and the other only a moderate loss.

hearing loss (presbycusis). The number of mutation-carrying mitochondria per cell increases with age, resulting in progressive reduction in oxidative phosphorylation capacity, the impact of which is tissue dependent. Ueda and colleagues detected high rates of mitochondrial DNA deletions in lymphocytes of persons with idiopathic sensorineural hearing loss and noted that the number of deletions increased with increasing hearing loss.[96] They posited that at least some cases of sensorineural hearing loss should be categorized as a mitochondrial phosphorylation disease. Temporal bone studies have also shown an increased load of mitochondrial mutations in persons with age-related hearing loss when compared with controls.[97]

Familial Aminoglycoside Ototoxicity A point mutation at base pair 1555 (adenine-to-guanine) in the 12s ribosomal ribonucleic acid (rRNA) causes both maternally inherited nonsyndromic deafness and maternally inherited susceptibility to aminoglycoside ototoxicity. This mutation makes the 12s rRNA more similar to bacterial rRNA and thus a target for aminoglycoside binding.[98] In countries such as China, where aminoglycosides are used frequently, aminoglycoside-induced ototoxicity is a major cause of hearing loss. One study in Shanghai found that 22% of the profoundly deaf had aminoglycoside-induced ototoxicity.[99] In the United States, over 15% of persons with aminoglycoside-induced hearing loss carry this mutation, a point of major clinical relevance for the prevention of aminoglycoside hearing loss.[100] Based on these data, it is prudent to inquire about any family history of aminoglycoside ototoxicity prior to drug administration. Individuals who develop ototoxicity should be tested for the 1555 mutation and offered family counseling if the mutation is detected.

CLINICAL DIAGNOSIS OF HEARING IMPAIRMENT

History

A complete history is an important element in the diagnosis of hereditary hearing impairment. Directed

questions should focus on prenatal, perinatal, and postnatal history, specifically reviewing maternal illnesses, drug use, alcohol intake, and smoking. Recognized risk factors for hearing loss in the perinatal period include low birth weight, prematurity, time spent in a neonatal intensive care unit, hyperbilirubinemia, sepsis, use of ototoxic medications, and birth hypoxia.[101–104] Risk factors in the postnatal period include viral illnesses such as mumps and measles as well as bacterial meningitis.[105] A record of speech and language milestones can establish whether the hearing loss is pre- or postlingual. However, even deaf infants coo and babble naturally up to the age of 6 months. A history of poor motor development may indicate vestibular dysfunction.

It is important to inquire about hearing loss in first- and second-degree relatives, especially if the loss started before the age of 30 years. Consanguinity or common origins from ethically isolated areas should increase suspicion of hereditary deafness.[106,107] If there are a number of family members with hearing loss, constructing a pedigree may delineate the mode of inheritance.

Physical Examination

Most cases of hereditary hearing impairment are nonsyndromic, so physical findings are absent. However, even in cases of syndromic hearing impairment, physical findings may be subtle. The physical examination should include a general inspection and systematic evaluation of all systems. Note hair color, the presence of a white forelock, facial symmetry, and skull shape. Fundoscopy, eye color, and relative position should be noted, taking specific biometric measurements in suspected cases of WS1. Examine the ears for auricular pits or sinuses and skin tags, noting the shape and size of the pinnae and checking for abnormalities of the external ear canal and tympanic membrane. Check the neck for branchial anomalies and thyroid enlargement and the oral cavity for clefts. Note the number, size, and shape of the digits and complete a thorough inspection of the skin for areas of pigmentation/hypopigmentation and café-au-lait spots. Do a complete neurologic examination, including tests of gait and balance to assess vestibular function.

Audiology

The test of choice for infants and young children with suspected hearing impairment is the auditory brainstem response, which gives accurate hearing thresholds from 1 to 4 kHz.[108] In older children or adults, a standard audiogram can be obtained. The presence or absence of hearing loss in other family members should be documented by formal audiometric testing.

Laboratory Testing

Specific laboratory tests should be ordered on the basis of the history, physical examination, and age of the patient (Table 13–7). For example, if PS is suspected, a perchlorate challenge test can be obtained to detect a defect in thyroid organification of iodide. If AS is being considered, urinalysis should be performed. An electrocardiogram should be ordered and the QT interval calculated if there is a history of syncopal episodes or a family history of sudden infant death syndrome (SIDS), as might be seen with JLNS. Serology can be used to rule out acquired causes of early deafness, such as cytomegalovirus infection, toxoplasmosis, and congenital rubella.

Chromosomal karyotyping is indicated in a child born to parents with a history of miscarriages, when the constellation of anomalies in a child is not recognizable as a previously reported syndrome, or if there are associated central nervous system or cardiac defects. Genetic testing for most of the types of deafness discussed in this chapter is not yet clinically available; however, it is likely that in the next 5 years, many new tests will be offered, making some older tests obsolete.

Radiology

Computed tomography of the temporal bones is the single best radiologic test for the evaluation of hearing impairment, with the reported incidence of anatomic abnormalities ranging from 6.8 to 28.4%.[109,110] Cochlear abnormalities such as Mondini's dysplasia and DVAs suggest the diagnosis of PS, although temporal bone anomalies are also seen with many other syndromes, including TCS and BOR syndrome. If the latter is being considered, renal ultrasonography should be performed.

Other Consultations

An ophthalmologic opinion should be obtained in all children with severe-to-profound hearing impairment as half will have ocular abnormalities.[111]

TABLE 13–7. Evaluation by Age of Onset

Type of Evaluation	*Neonate/Infant*	*Child/Adolescent*	*Adult*
To rule out acquired causes of HI	High-risk register Viral titers Viral cultures (urine, throat) Syphilis serology	Viral titers Syphilis serology	Viral titers Syphilis serology
Audiology and vestibular testing	Impedance testing ABR testing Otoacoustic emissions	Impedance testing Behavioral audiometry Conventional audiometry ABR Balancing testing	Impedance testing Conventional audiometry Cortical evoked response audiometry Electronystagmographic testing
Chromosomal karyotyping	Dysmorphic features History of miscarriage	Dysmorphic features Mental retardation	
Radiology	Temporal bone CT Renal ultrasonography	Temporal bone CT Renal ultrasonography	Temporal bone CT MRI of CP angle if hearing asymmetry
Ophthalmologic testing	Fundoscopy (CMV retinopathy)	Fundoscopy ERG	Fundoscopy ERG
Neurologic tests		Syndromes associated with mental retardation	Ataxia and neuropathies
Cardiac testing	ECG (Jervell and Lange-Nielsen syndrome)	ECG	
Dental evaluation		Osteogenesis imperfecta	
Renal evaluation	Ultrasonography (BOR)	Urine protein (Alport's syndrome) Urine dermatan sulfate, heparin suppurate (mucopolysaccharidoses)	Urine protein Ultrasonography

HI = hearing impairment; ABR = auditory brainstem response; CT = computed tomography; MRI = magnetic resonance imaging; CP = cerebellopontine; CMV = cytomegalovirus; ERG = electroretinography; ECG = electrocardiography; BOR = branchio-oto-renal syndrome.

Although most of these abnormalities are refractive errors, the correction of which is essential, ERG may uncover signs of early retinitis pigmentosa. Referral to a clinical geneticist should be requested to ensure that parents and patients adequately understand the issues such as recurrence risk.

MANAGEMENT OF HEARING IMPAIRMENT

Early identification of hearing impairment in infants and young children and early rehabilitation are essential for the development of age-appropriate speech and language skills. Children whose hearing losses are identified and in whom intervention is instituted before 6 months of age show significant advantage in communicative skill development compared with infants identified later.[112] To achieve this goal, newborn hearing screening programs have been implemented in most states and linked to rehabilitative programs. The level of habilitative intervention that is required depends on the degree of hearing impairment. Counseling of the family, proper hearing aid selection, hearing aid fitting, and continued

audiologic assessment are important. Schools that teach sign language, oral-aural communication, and a combination of both may be necessary. Cochlear implantation may be an option in persons with severe and profound hearing impairment.[113]

A variety of support systems exist, particularly on the World Wide Web. In the United States, the National Institute on Deafness and Other Communication Disorders established the Hereditary Hearing Impairment Resource Registry (HHIRR) (<www.boystown.org/hhirr>) to disseminate information on hearing impairment to professionals and families. The registry also collects information from individuals interested in supporting and participating in research projects and lists relevant resources and links to other organizations.

FUTURE DEVELOPMENTS

The rapid advances in the genetics and molecular biology of hearing and deafness that occurred during the last decade of the twentieth century are continuing at pace. Inexpensive genetic tests are becoming available for early detection of hearing impairment, and the use of these tests will ultimately impact management decisions by better defining therapeutic and habilitative options. It is also becoming increasingly possible to identify individuals at risk for environmental damage (noise, drugs, age) to their hearing. In the next decades, it is likely that physicians will be able to offer patients new, practical, and effective treatments for sensorineural hearing impairment.

REFERENCES

1. Morton NE. Genetic epidemiology of hearing impairment. Ann N Y Acad Sci 1991;630:16–31.
2. Marazita ML, Ploughman LM, Rawlings B, et al. Genetic epidemiological studies of early-onset deafness in the U.S. school-age population. Am J Genet 1993;46:486–91.
3. Sank D, Kallman FJ. The role of heredity in early total deafness. Volta Rev 1963;65:461.
4. Brown KS. The genetics of childhood deafness. In: McConnell F, Ward PH, editors. Deafness in childhood. Nashville (TN): Vanderbilt University Press; 1967. p. 177–202.
5. Anderson S, Bankier AT, Barrell BG, et al. Sequence and organization of the human mitochondrial genome. Nature 1981;290:457–65.
6. Jaber L, Shohat M, Bu X, et al. Sensorineural deafness inherited as a tissue specific mitochondrial disorder. J Med Genet 1992;29:86–90.
7. Grundfast KM. Hereditary hearing impairment in children. Adv Otolaryngol Head Neck Surg 1993;7: 29–43.
8. Gorlin A, Toriello H, Cohen M. Hereditary hearing loss and its syndromes. New York: Oxford University Press; 1995.
9. Melnick M, Bixler D, Silk K, et al. Autosomal dominant branchio-oto-renal dysplasia. Birth Defects Orig Artic Ser 1975;11:121–8.
10. Fraser FC, Sproule JR, Halal F. Frequency of the branchio-oto-renal (BOR) syndrome in children with profound hearing loss. Am J Genet 1980:7: 341–9.
11. Cremers CWRJ, Fikkers-van Noord M. The earpits—deafness syndrome. Clinical and genetic aspects. Int J Pediatr Otorhinolaryngol 1980;2: 309–22.
12. Fraser FC, Ling D, Clogg D, Nogrady B. Genetic aspects of the BOR syndrome—branchial fistulas, ear pits, hearing loss and renal anomalies. Am J Med Genet 1978;2:241–52.
13. Chen A, Francis M, Ni L, et al. Phenotypic manifestations of branchio-oto-renal syndrome. Am J Med Genet 1995;58:365–70.
14. Gimsing S, Dyrmose J. Branchio-oto-renal dysplasia in three families. Ann Otol Rhinol Laryngol 1986;95: 421–6.
15. Smith RJ, Coppage KB, Ankerstjerne JK, et al. Localization of the gene for branchiootorenal syndrome to chromosome 8q. Genomics 1992;14:841–4.
16. Kumar S, Kimberling WJ, Kenyon JB, et al. Autosomal dominant branchio-oto-renal syndrome—localization of a disease gene to chromosome 8q by linkage in a Dutch family. Hum Mol Genet 1992;1: 491–5.
17. Abdelhak S, Kalatzis V, Heilig R, et al. A human homologue of the *Drosophila* eyes absent gene underlies branchio-oto-renal (BOR) syndrome and identifies a novel gene family. Nat Genet 1997;15: 157–64.
18. Kumar S, Deffenbacher K, Marres HA, et al. Genomewide search and genetic localization of a second gene associated with autosomal dominant branchio-oto-renal syndrome: clinical and genetic implications. Am J Hum Genet 2000;66:715–20.
19. Cohen MM, Krieborg S. Birth prevalence studies of the Crouzon syndrome: comparison of direct and indirect methods. Clin Genet 1992;41:12–5.

20. Schiller JG. Craniofacial dysostosis of Crouzon: a case report and pedigree with emphasis on heredity. Pediatrics 1959;23:107.
21. Orvidas LJ, Fabry LB, Diacova S, McDonald TJ. Hearing and otopathology in Crouzon syndrome. Laryngoscope 1999;109:1372–5.
22. Hollway GE, Suthers GK, Haan EA, et al. Mutation detection in FGFR2 craniosynostosis syndromes. Hum Genet 1997;99:251–5.
23. Reardon W, Winter RM, Rutland P, et al. Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. Nat Genet 1994;8:98–103.
24. NIH Consensus Development Conference. Neurofibromatosis: conference statement. Arch Neurol 1988;45:475–8.
25. Van Camp G, Smith RJH. Hereditary hearing loss home page. Available at: http://dnalab-www.uia.ac.be/dnalab/hhh/ (accessed November 2000).
26. Trofatter JA, MacCollins MM, Rutter JL, et al. A novel moesin-, ezrin-, radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor. Cell 1993;75:826.
27. Gusella JF, Ramesh V, MacCollin M, Jacoby LB. Merlin: the neurofibromatosis 2 tumor suppressor. Biochim Biophys Acta 1999;1423:M29–36.
28. Kluwe L, Mautner VF. A missense mutation in the NF2 gene results in moderate and mild clinical phenotypes of neurofibromatosis type 2. Hum Genet 1996;97:224–7.
29. Evans DGR, Trueman L, Wallace A, et al. Genotype/phenotype correlations in type 2 neurofibromatosis (NF2): evidence for more severe disease associated with truncating mutations. J Med Genet 1998;35:450–5.
30. Knowlton RG, Struyk AF, Knobloch WH, et al. Genetic linkage analysis of hereditary arthro-ophthalmopathy (Stickler syndrome) and the type II collagen procollagen gene. Am J Hum Genet 1989;45:681–6.
31. Ahmad NN, Ala-Kokko L, Knowlton RG, et al. Stop codon in the procollagen II gene (COL2A1) in a family with Stickler syndrome (arthro-ophthalmology). Proc Natl Acad Sci U S A 1991;88:6624–7.
32. Winterpacht A, Hilbert M, Schwarze U, et al. Kniest and Stickler dysplasia phenotypes caused by collagen type II gene (*COL2A1*) defect. Nat Genet 1993;3:323–6.
33. Williams CJ, Ganguly A, Considine E, et al. A-2-->G transition at the 3′ acceptor splice site of IVS17 characterizes the *COL2A1* gene mutation in the original Stickler syndrome kindred. Am J Med Genet 1996;63:461–7.
34. Vikkula M, Mariman EC, Lui VC, et al. Autosomal dominant and recessive osteochondrodysplasias associated with the *COL11A2* locus. Cell 1995;80:431–7.
35. Richards AJ, Yates JR, Williams R, et al. A family with Stickler syndrome type 2 has a mutation in the *COL11A1* gene resulting in the substitution of glycine 97 by valine in alpha 1 (XI) collagen. Hum Mol Genet 1996;5:1339–43.
36. Rovin S, Dachi SF, Borenstein DB, Cotter WB. Mandibulofacial dysostosis, a familial study of five generations. J Pediatr 1964;65:215–21.
37. Sando I, Hemenway WG, Morgan WR. Histopathology of the temporal bones in mandibulofacial dysostosis. Trans Am Acad Ophthalmol Otolaryngol 1968;72:913–24.
38. Jahrsdoefer RA, Jacobson JT. Treacher Collins syndrome: otologic and auditory management. J Am Acad Audiol 1995;6:93–102.
39. Isaac C, Marsh KL, Paznekas WA, et al. Characterization of the nucleolar gene product, treacle, in Treacher Collins syndrome. Mol Biol Cell 2000;11:3061–71.
40. Edwards SJ, Gladwin AJ, Dixon MJ. The mutational spectrum in Treacher Collins syndrome reveals a predominance of mutations that create a premature-termination codon. Am J Hum Genet 1997;60:515–24.
41. Waardenburg PJ. A new syndrome combining developmental anomalies of the eyelids, eyebrows, and nose root with pigmentary defects of the iris and head hair and with congenital deafness. Am J Hum Genet 1951;3:195–253.
42. Read AP, Newton VE. Waardenburg syndrome. J Med Genet 1997;34:656–65.
43. Farrer LA, Grundfast KM, Amos J, et al. Waardenburg syndrome (WS) type I is caused by defects at multiple loci, one of which is near ALPP on chromosome 2: first report of the WS consortium. Am J Hum Genet 1992;50:902–13.
44. Liu XZ, Newton VE. Waardenburg syndrome type 1: phenotypic findings and diagnostic criteria. Am J Med Genet 1995;55:95–100.
45. Newton VE. Hearing loss and Waardenburg syndrome implications for genetic counselling. J Laryngol Otol 1990;104:97–103.
46. Tassabehji M, Read AP, Newton VE, et al. Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. Nature 1992;355:635–6.

47. Tassabehji M, Newton VE, Read AP. Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. Nat Genet 1994;8: 251–5.
48. Hoth CF, Milunsky A, Lipsky N, et al. Mutations in the paired domain of the human PAX3 gene cause Klein-Waardenburg syndrome (WS-III) as well as Waardenburg syndrome type I (WS-I). Am J Hum Genet 1993;52:455–62.
49. Attie T, Till M, Pelet A, et al. Mutation of the endothelin-receptor B gene in Waardenburg-Hirschsprung disease. Hum Mol Genet 1995;4:2407–9.
50. Edery P, Attie T, Amiel J, et al. Mutation of the endothelin-3 gene in the Waardenburg-Hirschsprung disease (Shah-Waardenburg syndrome). Nat Genet 1996;12:442–4.
51. Pingault V, Bondurand N, Kuhlbrodt K, et al. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat Genet 1998;18:171–3.
52. Tranebjaerg L, Bathen J, Tyson J, Bitner-Glindzicz M. Jervell and Lange-Nielsen syndrome: a Norwegian perspective. Am J Med Genet 1999;89:137–46.
53. Marres HA. Congenital abnormalities of the inner ear. In: Ludman H, Wright T, editors. Diseases of the ear. Bath (England): Arnold & Oxford University Press; 1998. p. 288–96.
54. Fraser GR. Association of congenital deafness with goitre (Pendred's syndrome). Ann Hum Genet 1965;28:201–48.
55. Scott DA, Wang R, Kreman TM, et al. The Pendred syndrome gene encodes a chloride-iodide transport protein. Nat Genet 1999;20:440–3.
56. Van Hauwe P, Everett LA, Coucke P, et al. Two frequent missense mutations in Pendred syndrome. Hum Mol Genet 1998;7:1099–104.
57. Coyle B, Reardon W, Herbrick JA, et al. Molecular analysis of the PDS gene in Pendred syndrome. Hum Mol Genet 1998;7:1105–12.
58. Rosenberg T, Haim M, Hauch AM, Parving A. The prevalence of Usher syndrome and other retinal dystrophy-hearing impairment associations. Clin Genet 1997;51:314–21.
59. Grondahl J, Mjoen S. Usher syndrome in four Norwegian counties. Clin Genet 1986;30:14–28.
60. Kimberling WJ, Moller CG, Davenport SL, et al. Usher syndrome: clinical findings and gene localization studies. Laryngoscope 1989;99:66–72.
61. Moller CG, Kimberling WJ, Davenport SL, et al. Usher syndrome: an otoneurologic study. Laryngoscope 1989;99:73–9.
62. Fishman GA, Kumar A, Joseph ME, et al. Usher's syndrome. Ophthalmic and neuro-otologic findings suggesting genetic heterogeneity. Arch Ophthalmol 1983;101:1367–74.
63. Cherry PM. Usher's syndrome. Ann Ophthalmol 1973;5:743–52.
64. Young NM, Mets MB, Hain TC. Early diagnosis of Usher syndrome in infants and children. Am J Otol 1996;17:30–4.
65. Kaplan J, Gerber S, Bonneau D, et al. A gene for Usher syndrome type I (USH1A) maps to chromosome 14q. Genomics 1992;14:979–87.
66. Smith RJ, Lee EC, Kimberling WJ, et al. Localization of two genes for Usher syndrome type I to chromosome 11. Genomics 1992;14:995–1002.
67. Sankila EM, Pakarinen L, Kaariainen H, et al. Assignment of an Usher syndrome type III (USH3) gene to chromosome 3q. Hum Mol Genet 1995;4:93–8.
68. Wayne S, Der Kaloustian VM, Schloss M, et al. Localization of the Usher syndrome type ID gene (Ush1D) to chromosome 10. Hum Mol Genet 1996;5:1689–92.
69. Kimberling WJ, Weston MD, Moller C, et al. Localization of Usher syndrome type II to chromosome 1q. Genomics 1990;7:245–9.
70. Hmani M, Ghorbel A, Boulila-Elgaied A, et al. A novel locus for Usher syndrome type II, USH2B, maps to chromosome 3 at p23-24.2. Eur J Hum Genet 1999;7:363–7.
71. Weil D, Blanchard S, Kaplan J, et al. Defective myosin VIIA gene responsible for Usher syndrome type 1B. Nature 1995;374:60–1.
72. Verpy E, Leibovici M, Zwaenepoel I, et al. A defect in harmonin, a PDZ domain-containing protein expressed in the inner ear sensory hair cells, underlies Usher syndrome type 1C. Nat Genet 2000;6:51–5.
73. Eudy JD, Weston MD, Yao S, et al. Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type IIa. Science 1998;280: 1753–7.
74. Weston MD, Eudy JD, Fujita S, et al. Genomic structure and identification of novel mutations in usherin, the gene responsible for Usher syndrome type IIa. Am J Hum Genet 2000;66:1199–210.
75. Flinter FA, Cameron JS, Chantler C, et al. Genetics of classic Alport's syndrome. Lancet 1988;2:1005–7.
76. Flinter F. Alport's syndrome. J Med Genet 1997;34: 326–30.
77. Govan JAA. Ocular manifestations of Alport's syndrome: a hereditary disorder of basement membrane? Br J Ophthalmol 1983;67:493–503.

78. Barker DF, Hostikka SL, Zhou J, et al. Identification of mutations in the *COL4A5* collagen gene in Alport syndrome. Science 1990;248:1224–7.
79. Lemmink HH, Schroder CH, Monnens LA, Smeets HJ. The clinical spectrum of type IV collagen mutations. Hum Mutat 1997;9:477–99.
80. Alcolado JC, Majid A, Brockington M, et al. Mitochondrial gene defects in patients with NIDDM. Diabetologia 1994;37:372–6.
81. Kadowaki T, Kadowaki H, Mori Y, et al. A subtype of diabetes mellitus associated with a mutation of mitochondrial DNA. N Engl J Med 1994;330:962–8.
82. Van Laer L, McGuirt WT, Yang T, et al. Autosomal dominant nonsyndromic hearing impairment. Am J Med Genet 1999;89:167–74.
83. Lalwani, AK, Castelein CM. Cracking the auditory genetic code: nonsyndromic hereditary hearing impairment. Am J Otol 1999;20:115–32.
84. Guilford P, Ben Arab S, Blanchard S, et al. A non-syndrome form of neurosensory, recessive deafness maps to the pericentromeric region of chromosome 13q. Nat Genet 1994;6:24–8.
85. Kelsell DP, Dunlop J, Stevens HP, et al. Connexin 26 mutations in hereditary nonsyndromic sensorineural deafness. Nature 1997;387:80–3.
86. Kumar NM, Gilula NB. The gap junction communication channel. Cell 1996;84:381–8.
87. Kikuchi T, Kimura R, Paul D, Adams J. Gap junctions in rat cochlea: immunohistochemical analysis. Anat Embryol 1995;191:101–18.
88. Denoyelle F, Weil D, Maw MA, et al. Prelingual deafness: high prevalence of a 30delG mutation in the connexin 26 gene. Hum Mol Genet 1997;6: 2173–7.
89. Maw MA, Allen-Powell DR, Goodey RJ, et al. The contribution of the DFNB1 locus to neurosensory deafness in a Caucasian population. Am J Hum Genet 1995;57:629–35.
90. Gasparini P, Estivill X, Volpini V, et al. Linkage of DFNB1 to non-syndromic neurosensory autosomal-recessive deafness in Mediterranean families. Eur J Hum Genet 1997;5:83–8.
91. Fukushima K, Ramesh A, Srisailapathy CR, et al. Consanguineous nuclear families used to identify a new locus for recessive non-syndromic hearing loss on 14q. Hum Mol Genet 1995;4:1643–8.
92. Green GE, Scott DA, McDonald JM, et al. Carrier rates in the midwestern United States for GJB2 mutations causing inherited deafness. JAMA 1999; 281:2211–6.
93. Morell RJ, Kim HJ, Hood LJ, et al. Mutations in the connexin 26 gene (GJB2) among Ashkenazi Jews with nonsyndromic recessive deafness. N Engl J Med 1998;339:1500–5.
94. Lefebvre PP, Van De Water TR. Connexins, hearing and deafness: clinical aspects of mutations in the connexin 26 gene. Brain Res Brain Res Rev 2000;32: 159–62.
95. Fischel-Ghodsian N. Mitochondrial mutations and hearing loss: paradigm for mitochondrial genetics. Am J Hum Genet 1998;62:15–9.
96. Ueda N, Oshima T, Ikeda K, et al. Mitochondrial DNA deletion is a predisposing cause for sensorineural hearing loss. Laryngoscope 1998;108: 580–4.
97. Fischel-Ghodsian N, Bykhovskaya Y, Taylor K, et al. Temporal bone analysis of patients with presbycusis reveals high frequency of mitochondrial mutations. Hear Res 1997;110:147–54.
98. Prezant TR, Agapian JV, Bohlman MC, et al. Mitochondrial ribosomal RNA mutation associated with both antibiotic-induced and non-syndromic deafness. Nat Genet 1993;4:289–94.
99. Hu DN, Qui WQ, Wu BT, et al. Genetic aspects of antibiotic induced deafness: mitochondrial inheritance. J Med Genet 1991;28:79–83.
100. Fischel-Ghodsian N, Prezant TR, Chaltraw WE, et al. Mitochondrial gene mutation is a significant predisposing factor in aminoglycoside ototoxicity. Am J Otolaryngol 1997;18:173–8.
101. Gerber S. Review of a high-risk register for congenital or early onset deafness. Br J Audiol 1992;26: 77–90.
102. Sutton GJ, Rowe S. Risk factors for childhood sensorineural hearing loss in the Oxford Region. Br J Audiol 1997;31:39-54.
103. Bamiou DE, Macardle B, Bitner-Glindzicz M, Sirimanna T. Aetiological investigations of hearing loss in childhood: a review. Clin Otolaryngol 2000;25: 98–106.
104. Davis A, Wood S. The epidemiology of hearing impairment: factors relevant to planning of services. Br J Audiol 1992;26:77–90.
105. Linthicum FH. Viral causes of sensorineural hearing loss. Otolaryngol Clin North Am 1978;11: 29–33.
106. Basil A. Childhood sensorineural hearing loss in consanguineous marriages. J Audiol Med 1994;3:151–9.
107. Naem Z, Newton V. Prevalence of sensorineural hearing loss in Asian children. Br J Audiol 1996;30: 332–9.

108. Tomaski SM, Grundfast KM. A stepwise approach to the diagnosis and treatment of hereditary hearing loss. Pediatr Clin North Am 1999;46:35–47.
109. Zalzal GH, Shott SR, Towbin R, Cotton RT. Value of CT scan in the diagnosis of temporal bone diseases in children. Laryngoscope 1986;96:27–32.
110. Bamiou D, Phelps P, Sirimanna T. Temporal bone computed tomography findings in bilateral sensorineural hearing loss. Arch Dis Child 2000;82: 257–60.
111. Armitage IM, Burke JP, Buffin JT. Visual impairment in severe and profound sensorineural deafness. Arch Dis Child 1995;75:53–6.
112. Downs MP, Yoshinaga-Itano C. The efficacy of early identification and intervention for children with hearing impairment. Pediatr Clin North Am 1999;46:79–87.
113. Nikolopoulos TP, O'Donoghue GM. Cochlear implantation in adults and children. Hosp Med 1998; 59:46–9.

CHAPTER 14

Trauma to the Middle Ear, Inner Ear, and Temporal Bone

Mitchell K. Schwaber, MD

EXTERNAL AUDITORY CANAL

Two types of injury are likely to involve the external auditory canal: blunt and penetrating trauma and thermal and caustic burns.

Isolated blunt trauma to the ear canal is most often caused by the insertion of a foreign object into the ear to scratch the skin or to remove wax. The skin of the ear canal, particularly the anterior and inferior part of the canal, is quite thin, with a minimal subepithelial layer. The tender skin of this portion of the canal is easily abraded and will bleed readily, particularly if the patient is on anticoagulant or antiplatelet therapy. In most instances, the pain and the sight of blood from the ear canal cause the patient to seek medical attention. In some cases, however, a secondary infection develops, and pain, hearing loss, or infected drainage causes the patient to seek help. The canal should be gently cleaned using microscopic technique, and blood clots, debris, and wax should be removed. When the bleeding site has been identified, placing a small pledget of Gelfoam coated with antibiotic ointment over it can readily control the bleeding. Alternatively, Gelfoam soaked with topical thrombin can be applied to the bleeding site. In rare cases, the site must be cauterized and then packed with a Merocel wick. In contrast, the skin of the posterosuperior part of the canal is much thicker and is more resistant to abrasions and injuries. Posterior lacerations usually stop bleeding because the subepithelial layer is more developed; therefore, vessels in this area readily contract and clot off.

After microscopic cleaning of the canal, the tympanic membrane should be inspected to determine the extent of injury. If an injury to the middle ear or temporal bone is discovered in this process, a complete neurotologic examination should be performed to evaluate the patient. Audiometric assessment is also obtained at this point to assess the function of the ossicular chain and the cochlea. Radiographs are not usually indicated unless an extremely severe injury is found. Ciprofloxacin and hydrocortisone otic drops are prescribed to prevent infection and to dissolve the Gelfoam pack.

Blunt injuries to the temporal bone typically result from the head being forced against a stationary object in a deceleration injury or from an object being thrown directly at the head. Most injuries involving this region occur as a result of a glancing blow to the temporal region. Although soft tissue injuries to the auricle and canal occur, they are often accompanied by fractures of the external auditory canal, middle ear structures, otic capsule, or surrounding structures. Mandibular injuries, particularly those that drive the mandible posteriorly into the jaw joint, will occasionally fracture the anterior wall of the ear canal, resulting in laceration of the skin and exposure of bone. Following blunt trauma to the ear and temporal bone, the external auditory canal should be carefully cleaned and bleeding controlled as described earlier. If exposed bone is found, it should not be débrided at this point but rather assessed later when the canal has healed. Radiographs, including facial bone computed tomographic (CT) scans and pantomographic mandibular views, should be obtained in these patients to define the injuries. Occasionally, a direct blow to the auricle results in an isolated fracture of the external auditory canal and mastoid process, that is, a fracture not involving the deeper parts of the temporal bone that are in contact with the dura mater. Regardless of where the fracture has occurred, the clinician should be aware that squamous cell epithelium can be entrapped by the fracture fragments, leading to the development of a canal

cholesteatoma. Canal fractures can also lead to chronic infection, bone sequestration, and stenosis of the canal. The development of any of these sequelae may necessitate surgical débridement, grafting, reconstruction, or meatoplasty to ensure a healthy open ear canal.

Penetrating injuries of the external auditory canal are usually caused by gunshot or stab wounds. Penetrating injuries in which the missile enters anteriorly through the parotid gland often involve the external auditory canal; the mechanism is thought to involve dissipation of the missile's energy on and reflection of the missile off the anterior aspect of the mastoid process. Other factors that influence the injury include the force and trajectory of the missile and, if a bullet, its type, hardness, and caliber. As a consequence, facial nerve injury, tympanic membrane perforation, and ossicular dislocation can result from gunshot wounds of the external auditory canal even when the apparent course of the missile would not traverse these structures (Figure 14–1). The facial nerve is most likely to be injured at the stylomastoid foramen apparently because it is relatively fixed at that point. In the absence of any of these additional injuries, gunshot wounds of the ear canal require cleaning, a light dressing, and prophylactic antibiotics. Occasionally, the canal must be stented following a gunshot injury, and in this case, a soft Merocel wick coated with ciprofloxacin and hydrocortisone otic drops is usually satisfactory.

Lacerations of the external auditory canal can occur either anteriorly or posteriorly and are often accompanied by partial avulsion of the auricle. These patients should be carefully evaluated for injury to the facial nerve and the great vessels; radiographic studies including arteriography may be indicated in these patients. Most external auditory canal lacerations require cleaning, gentle débridement, and suturing to realign the various parts of the ear canal and auricle. Surprisingly, stenosis does not usually occur in these patients.

Burns and caustic injuries to the ear canal often represent a potentially complicated situation in that severe burns can lead to circumferential scarring and stenosis of the canal. Most of these injuries are attributable to one of three mechanisms: a thermal burn, a caustic burn, or a welding injury. Thermal and caustic burns of the ear canal are usually associated with additional injury to the auricle, which may in itself lead to loss of cartilage, cicatrix formation, and stenosis of the canal. Most thermal burns of the ear canal are caused by flash injuries, fires, lightning strikes, or hot liquids such as oil. Like burns elsewhere on the body, the depth and extent of the burn should be determined and documented. Superficial thermal burns of the ear canal are usually treated with the application of antibiotic ointment. If more than half of the ear canal is involved or has third-degree burns, in addition to the application of antibiotic ointment, the canal is stented with soft Silastic tubing. Canal stenting is performed in an effort to prevent stenosis of the canal, which leads to the trapping of squamous debris and ultimately a destructive ear canal cholesteatoma. Stenosis of the canal is treated aggressively with corticosteroid injections, frequent dilations, and, in some cases, skin grafting or even meatoplasty.

Caustic burns are usually caused by a chemical spill or a foreign object such as an alkaline battery. Thermal and acid burns cause coagulation necrosis, whereas alkaline burns cause liquefaction necrosis, which leads to much more extensive injury over time. Kavanagh and Litovitz reported a series of battery-related injuries to the ear canal that were much more frequent and severe than expected, including tympanic membrane perforation, exposed bone of the ear canal, sensorineural hearing loss, ossicular destruction, and facial paralysis.[1] They also noted that otic drops must be withheld in these patients as they provide an external electrolyte bath for the battery, enhancing leakage and generation of an external current with subsequent tissue electrolysis and hydroxide formation. The foreign body should be removed as soon as possible, under general anesthesia if needed.[2] Once the injury has been assessed, caustic burns are treated much like thermal burns described earlier, with microscopic cleaning, antibiotic eardrops, and stenting if indicated. Because of the high incidence of exposed bone, skin grafting of the ear canal is often indicated in these patients. Skin grafting, tympanoplasty, and ossicular reconstruction should be delayed for 4 to 6 months, if possible, to allow the full extent of the injury to be determined.

Welding injuries occur when hot slag or molten metal enters the meatus, usually resulting in either a small and localized burn of the ear canal or a tympanic membrane perforation. In most patients, microscopic cleaning, ciprofloxacin and hydrocortisone otic drops, and observation are the only meas-

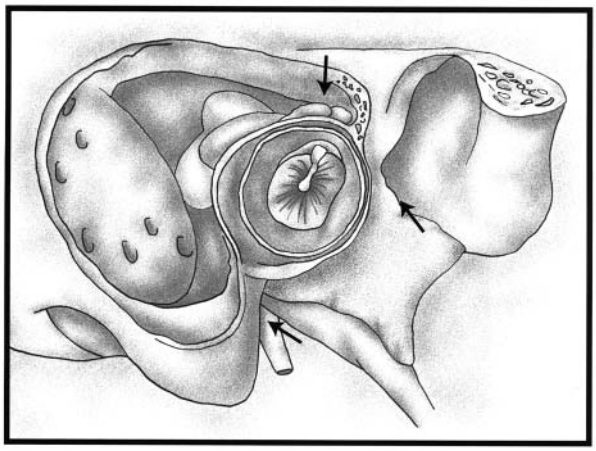


FIGURE 14–1. Drawing depicts the anatomy of the temporal bone from a lateral perspective. The *arrows* indicate sites likely to be injured in a gunshot wound, including the stylomastoid foramen, ossicular heads, and internal carotid artery.

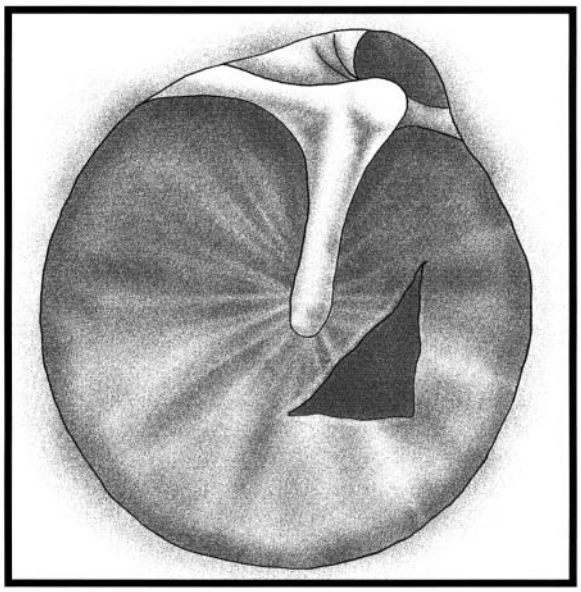


FIGURE 14–2. Drawing illustrates a tympanic membrane perforation in the anteroinferior portion of the drumhead.

ures indicated for welding injuries of the external auditory canal.

TYMPANIC MEMBRANE AND THE MIDDLE EAR

Trauma to the tympanic membrane and the middle ear can be caused by (1) overpressure, (2) thermal or caustic burns, (3) blunt or penetrating injuries, and (4) barotrauma. *Overpressure* is by far the most common mechanism of trauma to the tympanic membrane. The major causes of overpressure include slap injuries and blast injuries. Slap injuries are extremely common and can be a result of either a hand or water slap. Slap injuries usually result in a triangular or linear tear of the tympanic membrane (Figure 14–2). Most of these perforations cause mild hearing loss, aural fullness, and mild tinnitus. Blast injuries, although much less common, are potentially much more serious. Blast injuries may be caused by bomb explosions, gasoline explosions, and air-bag deployment in automobile accidents. Blast injuries from bomb explosions not only disrupt the tympanic membrane but also can cause temporal bone fracture, ossicular discontinuity, or high-frequency sensorineural hearing loss owing to cochlear injury. In addition, blast injuries can cause a perilymphatic fistula (PLF), with progressive and fluctuant hearing loss, vertigo, and dysequilibrium.

In a report by Hallmo, air- and bone-conduction audiometry in the frequency ranges of 0.125 to 18 kHz and 0.25 to 16 kHz, respectively, was performed on 38 patients with unilateral traumatic tympanic membrane perforation, mostly caused by overpressure injuries.[3] Bone-conduction threshold elevations were found in 16 ears. Both bone-conduction threshold elevations and tinnitus diminished with time, but in 9 patients it was permanent. Closure of the tympanic membrane perforation resulted in a 7 to 20 dB improvement of air-conduction thresholds, somewhat less in the upper than in the lower frequencies. A 3 dB mean final conductive hearing loss was found approximately 5 months after injury, probably owing to scars at the sites of the former perforations.

Following an overpressure injury, blood, purulent secretions, and debris should be carefully suctioned from the ear canal, and the perforation size and location should be recorded. Irrigation and pneumatic otoscopy should be specifically avoided in these patients. The ability to hear a whisper as well as tuning fork tests should be documented, and an audiogram should be obtained as soon as the patient's condition allows. A complete neurotologic examination should also be performed in these patients to document the status of the cranial nerves

including the facial nerve and the vestibular nerve as well as the central nervous system. If the tympanic membrane perforation is dry, it should be observed (ie, drops are not indicated). If there is drainage through the tympanic membrane perforation, the clinician should determine and note if the drainage is consistent with cerebrospinal fluid (CSF). If a CSF leak is suspected, immediate CT scan of the temporal bone should be obtained to rule out a fracture. If the drainage is not consistent with CSF, oral antibiotics and ciprofloxacin and hydrocortisone otic drops should be prescribed. A history of vertigo or nausea and vomiting and an audiogram showing a conductive hearing loss of more than 30 dB suggest disruption of the ossicular chain. Profound sensorineural loss also may signify oval window or cochlear damage.

Thermal injuries to the tympanic membrane include welding and lightning injuries. Welding injuries occur when hot slag enters the ear canal and passes through the tympanic membrane. Most of these injuries result in an inflammation in the middle ear with drainage. Panosian and Dutcher reported two patients with facial paralysis caused by hot slag in the middle ear.[4] One of their patients also sustained a sensorineural hearing loss. Welding injuries often result in nonhealing perforations, either as a result of infection or possibly because the slag acts to cauterize or devascularize the tympanic membrane as it passes through it. If infection occurs, the patient is treated with ciprofloxacin and hydrocortisone otic drops and an oral antibiotic. If the perforation is dry, it should be observed for a period of 12 weeks for spontaneous healing. If the drumhead does not heal, tympanoplasty should be performed.

Lightning and electrocution injuries are not rare, and the most frequent ear injury is perforation of the tympanic membrane. The most common vestibular disturbance is transient vertigo. Other clinical findings include sensorineural hearing impairment, conductive hearing loss, tinnitus, temporal bone fracture, avulsion of the mastoid process, burns of the ear canal, and facial nerve paralysis.[5] Jones et al reported one patient with bilateral oval window PLF following a lightning strike.[6] The initial management of the patient struck by lightning consists of life support measures. Thereafter, the patient should have a thorough audiovestibular assessment. Tympanic membrane perforations caused by lightning injury often do not heal, perhaps as a result of cauterization or devascularization of the tympanic membrane, much like welding injuries. These injuries are treated much like that described earlier for welding injuries. Tympanoplasty should be delayed in these patients for 12 weeks because spontaneous healing may take that long.

Caustic injuries to the tympanic membrane can cause a perforation. With alkaline caustics, the tympanic membrane is damaged by liquefaction necrosis, that is, the alkaline caustic penetrates the tympanic membrane, causing occlusion of the vasculature that may extend farther than the visible perforation. As a result, the size of the perforation may not be fully appreciated until all of the inflammation resolves. Furthermore, after caustic injuries, the middle ear can develop an extensive granulation reaction with scarification, ossicular fixation, and a chronic infection. Caustic injuries also can lead to canal blunting as the raw surfaces that surround the canal form a cicatrix, leading to narrowing of the ear canal and loss of the vibratory surface of the tympanic membrane. Similarly, following a caustic injury, chronic myringitis can develop on the surface of the tympanic membrane, creating a raw weeping surface with granulation on the surface of the drumhead. Caustic injuries are initially treated with ciprofloxacin and hydrocortisone otic drops, oral antibiotics, and analgesics. Audiologic assessment and a complete neurotologic evaluation are indicated in caustic injuries to determine the extent of injury. When the ear has stabilized, and preferably when drainage has diminished, the middle ear and tympanic membrane can be reconstructed.

Tympanic membrane perforations historically have a healing rate that approaches 80%. Kristensen reviewed more than 500 texts regarding the matter and found that the spontaneous healing rate appeared to be 78.7% in 760 evaluable cases of traumatic tympanic membrane perforations of all sorts seen within 14 days after injury.[7] A relative, causally related variation of spontaneous healing could be demonstrated in this review, with air pressure/hand slap injuries the most likely to heal. Ruptures induced by heat or corrosives, foreign objects, and water pressure were less likely to heal, perhaps because they were larger or more likely to be infected. Rybak and Johnson also reported that water slap injuries were less likely to heal as a result of infection.[8]

Griffin reported 227 traumatic perforations treated in his practice from 1969 to 1977.[9] He concluded that larger perforations, lightning and welding injuries, and infected ears were less likely to heal. Good hearing results were found regardless of the method of tympanoplasty, although spontaneous healing resulted in the best hearing outcome.

Regardless of the method used, successful tympanoplasty requires adequate exposure, débridement of middle ear granulation and scar tissue, de-epithelialization of the perforation, and careful graft placement including support of the graft until healing occurs. The author prefers to perform transcanal underlay tympanoplasty only on small perforations confined to the posterior portion of the tympanic membrane. In this circumstance, perichondrium harvested from the tragus is most often used as a graft material because of ease of handling and access. In larger perforations or perforations that extend anterior to the malleus, the author prefers to perform postauricular underlay tympanoplasty, using temporalis fascia as the graft material. In this technique, large "swinging door" flaps, based on the canal skin and normal tympanic membrane, are created. These flaps can be advanced to cover the entire mesotympanum if needed, a technique that is particularly useful in welding injuries as the adequacy of the vasculature of the residual drumhead may be uncertain. More information regarding tympanoplasty can be found in Chapter 10.

Penetrating trauma to the middle ear can, of course, result in perforation of the tympanic membrane, but unlike overpressure and thermal injury, the incidence of ossicular disruption, facial nerve, and other middle ear injuries is much greater. The most common causes are low-velocity gunshots followed by injury with a foreign object such as a stick or instrument. This type of injury should be suspected in patients with a tympanic membrane perforation, blood in the middle ear or ear canal, and the presence of vertigo or dizziness, a conductive loss greater than 25 dB, a sensorineural hearing loss, or a facial paralysis. In these patients, the ear canal should be gently suctioned and cleaned under microscopic vision, and the tympanic membrane and middle ear should be carefully inspected. A complete neurotologic examination, including facial nerve evaluation and examination for nystagmus, gait stability, fistula test, Romberg's test, and Dix-Hallpike test, should be performed. Imaging studies including CT scans of the temporal bone, magnetic resonance imaging (MRI), and even arteriography may be indicated depending on the type of injury suspected.

TEMPORAL BONE FRACTURES

Fractures of the temporal bone are caused by blunt injuries, and depending on the force and direction of the blow delivered, different types of fractures occur. Blunt trauma can be delivered by an object striking the head or by the head being thrown against a solid object. Traditionally, temporal bone fractures are classified as either longitudinal (extracapsular) or transverse (capsular) with respect to the long axis of the petrous portion of the temporal bone (Figure 14–3). Both are basal skull fractures and are associated with ecchymosis of the postauricular skin (Battle's sign).

Longitudinal fractures are, by far, the most common, accounting for 70 to 90% of temporal bone fractures, and typically result from a direct lateral blow to the temporal or parietal aspect of the head. The longitudinal fracture begins in the external auditory canal and extends through the middle ear and along the long axis of the petrous pyramid. Characteristically, there is bleeding from the ear canal owing to laceration of its skin and from blood coming through the perforated tympanic mem-

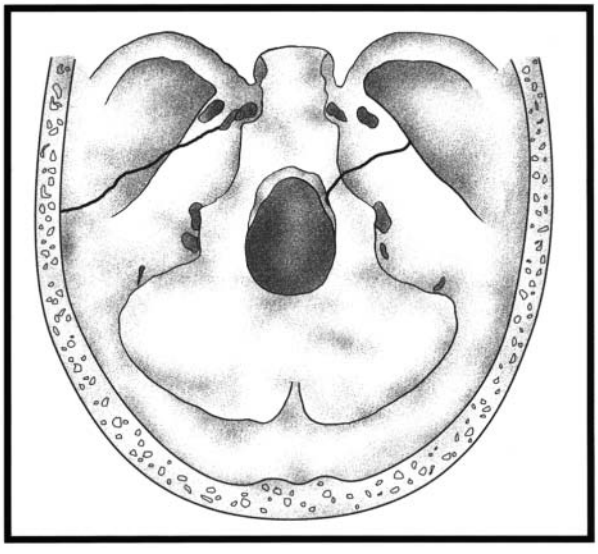


FIGURE 14–3. Drawing depicts the anatomy of the skull base. On the left is a longitudinal or extracapsular fracture. On the right is a transverse or capsular fracture.

brane. Facial paralysis occurs in 15%, and sensorineural hearing loss occurs in 35%.

Transverse fractures typically result from deceleration impacts in the occipital area. The fracture line transverses the long axis of the petrous portion of the temporal bone and usually extends through the cochlea and fallopian canal, resulting in sensorineural hearing loss and facial paralysis in most cases. There is bleeding into the middle ear, but the tympanic membrane remains intact and becomes blue-black owing to the hemotympanum.

Post-traumatic Ossicular Chain Disruption

Post-traumatic ossicular chain abnormalities include incudostapedial joint separation, dislocation of the incus, fracture and dislocation of the stapes, massive dislocation of the entire chain, and ossicular fixation owing to scarring or ossification. Incudostapedial joint separation is the most common ossicular abnormality and is more often seen with penetrating trauma as well as longitudinal or extracapsular temporal bone fractures (Figure 14–4). The forces that cause this type of fracture occur parallel to the long axis of the temporal bone and tend to displace the malleus and incus medially and inferior. As a consequence, incudostapedial joint separation is most common with longitudinal fractures, followed by dislocation of the incus. In patients in whom significant conductive hearing loss (ie, greater than 25 dB) is found, incudostapedial joint separation or incus dislocation should be suspected. In patients in whom mixed hearing loss is found or significant vertigo occurs, a fracture or dislocation of the stapes should be suspected. In this situation, the force may have sheared the crura off the footplate, or the entire stapes may be dislocated (Figure 14–5). This distinction is important in management in that incus subluxation or incudostapedial joint separation can be observed once the hemotympanum and swelling have resolved. In most cases, because of the tendency of the drumhead to adhere to the stapes, the residual conductive hearing loss is minimal after healing and requires no surgery. On the other hand, stapes dislocation with vertigo and/or progression of the sensorineural hearing loss is an indication for timely surgical exploration and repair to prevent anacusis.

If a significant conductive hearing loss is found after resolution of the hemotympanum and ventilation of the middle ear, the patient should be offered hearing rehabilitation either through surgery or the use of a hearing aid. Middle ear surgery in this case is usually performed through a transcanal approach, although a postauricular approach may be indicated if a large fracture or canal defect is present. After elevating a tympanomeatal flap, the middle ear and the ossicles are carefully visualized. Fibrous adhesions should be carefully dissected, and the fragility of the stapes should be respected. One of four procedures is used to correct the ossicular abnormality.

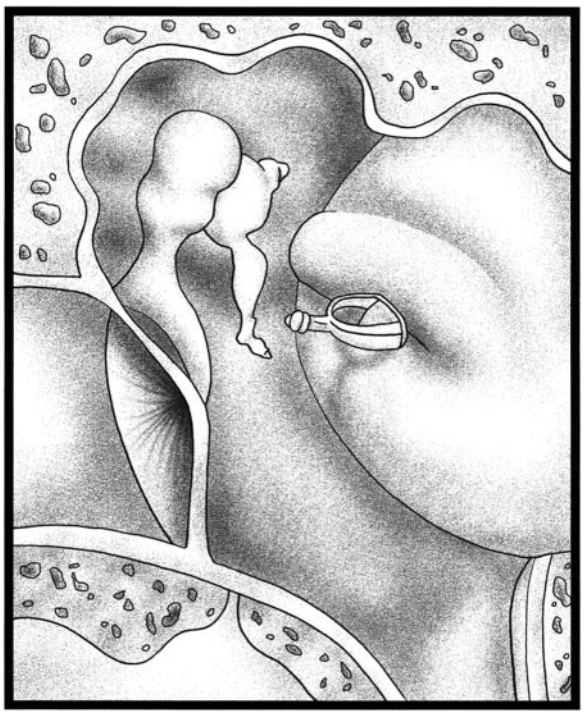


FIGURE 14–4. Drawing depicts incudostapedial joint disarticulation. Note that the incus is displaced inferiorly.

In the vast majority of patients, the incus is so displaced that it cannot be used efficiently. In these patients, the surgeon may choose to place a partial ossicular replacement prosthesis (PORP), linking the stapes directly to the drumhead. In these patients, the incus and the malleus handle must be removed, and the PORP is attached to the stapes capitulum and then covered with a thin wafer of cartilage (Figure 14–6, A). Alternatively, the surgeon may choose to link the stapes to the malleus, using either a commercially available prosthesis or a sculpted incus. These two techniques give comparable results, usually resulting in a 15 to 20 dB air–bone gap. In

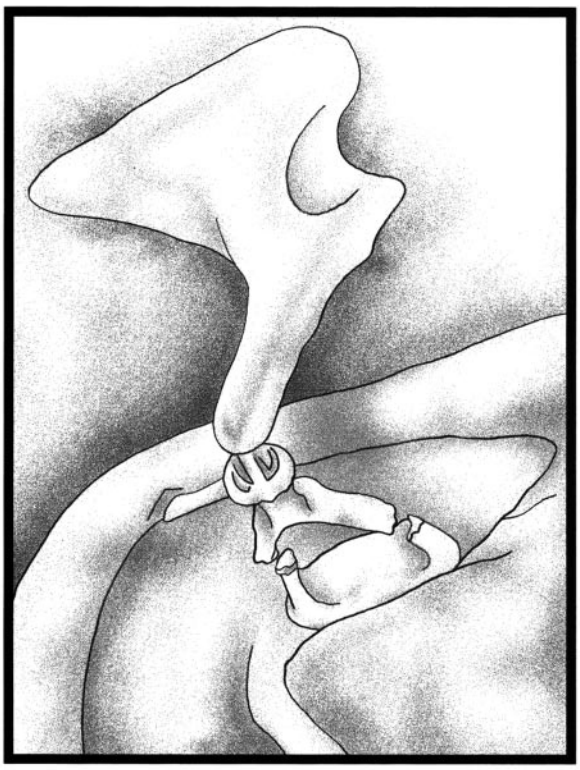


FIGURE 14–5. Drawing illustrates fracture through the stapes superstructure. In this case, the footplate is not displaced.

some cases, the surgeon may choose to place a total ossicular replacement prosthesis (TORP), with an intact stapes. This type of placement offers greater stability for the prosthesis, which can be placed between the tympanic portion of the facial nerve canal and the superstructure of the stapes (Figure 14–6, B). A small perichondrial graft placed on the footplate between the crura further stabilizes this assembly, as does linking the prosthesis to the malleus. Like the PORP technique, a thin wafer of cartilage should be placed over the TORP to link it to the drumhead.

Of course, a TORP can be used in patients in whom the stapes superstructure is disrupted or fractured. In these patients, the incus and malleus may not be usable because the displacement is so great or they are damaged. The stapes footplate should be covered with a perichondrial graft to stabilize the TORP and to ensure that any footplate fractures are covered. The TORP is placed on the perichondrial graft and linked to the tympanic membrane. In some patients, the stapes superstructure is damaged or the footplate is disrupted, in the presence of an otherwise normal malleus and incus. In these patients, a stapes prosthesis can be used. The first step in this circumstance is to ensure that a depressed stapes is carefully elevated out of the vestibule. Aggressive manipulation should be avoided, as should excessive suctioning. If the stapes can be restored to a more normal position, it can be reinforced with small bits of perichondrium and fibrin glue. Once the stapes footplate is in a satisfactory position, it is covered with perichondrium, and a stapes prosthesis can be linked from the incus to the footplate or from the malleus to the footplate. This technique should closely resemble the standard one used for otosclerosis except that the footplate is not removed. The prosthesis should be slightly shorter than in the standard situation because of the intact footplate. Although the author prefers to use a thin stemmed bucket-handle prosthesis in these patients, a crimped piston is also certainly appropriate.

Post-traumatic Perilymphatic Fistula

Severe head trauma, whether blunt or penetrating, is associated with a high incidence of sensorineural hearing loss. Nearly one-third of these patients develop a high-frequency hearing loss, which is thought to arise from a cochlear concussion, shearing of the basilar membrane, hair cell degeneration, and avulsion of auditory nerve fibers. Excessive stapes footplate excursion can also cause this same pathophysiologic consequence, so differentiating a concussion from a PLF can be difficult. Traumatic PLF can result from penetrating trauma with resultant fracture or subluxation of the stapes, from an overpressure injury such as a blast or a severe barotrauma event, or from a fracture of the temporal bone. The mechanics of these injuries are consistent with the implosive mechanism described by Goodhill.[10]

A PLF is a persistent, abnormal communication between the inner ear and the middle ear air space. In most cases, this communication occurs through the oval or round windows, although, occasionally, a fistula through a semicircular canal is discovered. As opposed to a cochlear concussion, the hearing threshold in PLF cases continues to fluctuate or to deteriorate. Tinnitus and fullness are also common complaints of these patients. However, the features that differentiate a PLF from a labyrinthine concussion are episodic vertigo, movement-trig-

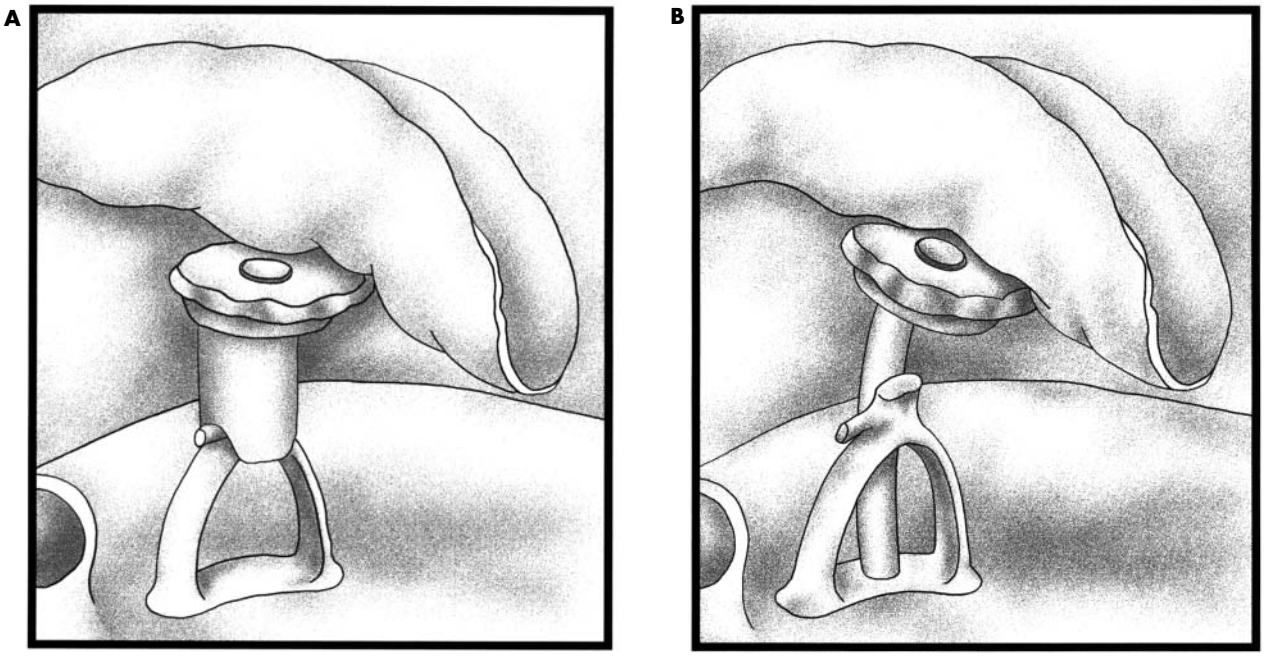


FIGURE 14–6. A, Drawing depicts the placement of a partial ossicular reconstruction prosthesis. *B,* Drawing depicts the placement of a total ossicular reconstruction prosthesis with an intact stapes superstructure. This technique provides better stabilization of the prosthesis, particularly if a graft is placed on the footplate.

gered vertigo, persistent unsteadiness, and often nausea. The vertigo in these patients lasts several minutes up to several hours. In the setting of trauma, with progressive hearing loss and episodic vertigo, a PLF should be suspected. The clinician should remember that there are no specific patterns to diagnose conclusively a PLF.

The evaluation of patients suspected of having PLF begins with a history and examination. The type of trauma suffered and the degree of injury are important features of the history. Also, the precipitating factors that lead to vertigo should be noted including ear manipulation, straining, and head positioning. Antecedent events such as increased tinnitus, ear fullness, or hearing fluctuation prior to vertigo are also significant. On examination, hemorrhage, laceration, and tympanic membrane perforation are all signs that point to the possibility of a PLF. In performing the neurotologic examination, special attention should be given to the evaluation for spontaneous nystagmus, an abnormal Romberg's test, a positive fistula sign, a positive Dix-Hallpike sign, an unsteady gait, or an abnormal step test. Audiometric assessment may show a mixed hearing loss or a pure sensorineural hearing loss. Additional audiovestibular testing is not usually obtained unless the drumhead and canal have healed, in which case an electronystagmographic (ENG) or a computerized dynamic posturographic test can add support for the abnormal physical examination findings. In addition, imaging studies should be reviewed with the radiologist to determine if a temporal bone fracture or some unsuspected central nervous system injury has occurred. In most cases, however, unless ossicular disruption is particularly severe, a PLF is not usually detected through imaging studies.

If a PLF is suspected, bed rest, corticosteroids, and even diuretics can be tried for a period of 2 to 3 days to determine if spontaneous healing might occur. If symptoms persist, the middle ear should be explored as soon as the patient's condition allows. Altered mental status and other injuries might delay middle ear exploration for a significant period of time, unfortunately. The middle ear in these patients can be explored either through a postauricular or transcanal approach, obtaining either temporalis fascia or tragal perichondrium for the repair. After elevating the tympanomeatal flap, the middle ear structures are carefully visualized and the necessary dissection is performed to visualize the stapes foot-

plate and the round window niche. A subluxed stapes should be carefuly elevated, as previously described, and secured in position using perichondrium and fibrin glue. If the oval window is indeed intact, the surgeon should carefully inspect the round window niche and membrane. At this point, several maneuvers are performed to visualize a PLF, including Trendelenberg positioning, neck compression, and Valsalva-like maneuver administered by the anesthesiologist. A PLF is likely present if fluid emanates from the oval window or pools in the round window niche with these maneuvers. Whether or not a PLF is found, most authors apply a soft tissue patch to the round and oval windows at this point. This should be done carefully to line the windows with tissue, not invaginate the tissue into the vestibule or the scala tympani. The soft tissue grafts are reinforced with Gelfoam, and the tympanomeatal flap is returned to its normal position. The canal is filled with antibiotic ointment or Gelfoam. Rest and avoiding heavy lifting are recommended for 5 to 7 days to allow the tissue seal to solidify. Following PLF repair, the episodic vertigo significantly improves in the vast majority of patients. However, the hearing results are much less predictable as only 15 to 20% of patients demonstrate substantial hearing improvement. The hearing threshold is usually stabilized, however.

Post-traumatic Vestibular Dysfunction

The most common type of post-traumatic vestibular dysfunction, by far, is benign paroxysmal positional vertigo (BPPV). This occurs in 50% of patients with temporal bone fractures and in 25% of patients with a head injury without fracture. In this circumstance, the dizziness usually begins within a few days of the injury, although it can be delayed for several months. Post-traumatic BPPV is thought to occur because deceleration forces disrupt the macula of the utricle, with release of the otoconia that ordinarily rest in a gel layer that covers the surface of the macula. The otoconia float into a dependent canal, most often the posterior semicircular canal, but the problem has been described in other canals. This mechanism is commonly called canalithiasis. The shifting of the mass of otoconia actually displaces the fluid in the canal on head tilting. As a consequence, a short burst of vertigo occurs, with a typical 5- to 7-second delay, lasting 20 to 30 seconds. Typically, post-traumatic BPPV is unilateral and is triggered by placing the affected ear in the down position. This condition can be diagnosed by performing a Nylen-Bárány or Dix-Hallpike maneuver and observing the nystagmus that is triggered. Although this should ideally be performed with Fresnel's glasses, the nystagmus can be observed without them in the vast majority of patients. Typically, in this maneuver, the nystagmus has a rotary motion, either clockwise or counterclockwise, toward the down and affected ear. Untreated, this condition usually lasts 3 to 4 months and gradually resolves. Treatment consists of vestibular suppressants such as alprazolam 0.50 mg or clonazepam 0.50 mg given twice daily in addition to antinauseants such as meclizine or promethazine as needed. Treatment also consists of vestibular rehabilitation or physical therapy, specifically using particle-repositioning maneuvers such as the modified Epley[11] or Semont[12] maneuver. In approximately 80% of patients, the symptoms are immediately alleviated by this method of treatment.

Post-traumatic vestibular dysfunction can also be caused by shearing of labyrinthine membranes as a result of concussive forces. In this circumstance, the patient may complain of chronic unsteadiness rather than true vertigo. In these patients, the nature of the injury should be carefully investigated with ENG as well as with computerized dynamic posturography. A vestibular deficit on either study adds support for this diagnosis as opposed to a postconcussion syndrome. In cases of chronic unsteadiness, a program of vestibular rehabilitation therapy, as well as low doses of clonazepam, can be extremely helpful. Unfortunately, since many of these cases follow automobile accidents, litigation issues often cloud the exact severity of symptoms.

Post-traumatic vestibular dysfunction also occurs after transverse temporal bone fractures. With a transverse fracture, a severe episode of vertigo occurs, which gradually improves, much like a patient who has undergone a labyrinthectomy. Following a transverse temporal bone fracture or a severe disruption of the stapes with total loss of hearing, the labyrinth repairs itself by a process called labyrithitis ossificans. This process is slow and takes several months before the semicircular canals ultimately ossify. Complete ossification, in effect, fixes the ampullae and maculae in solid bone so that they no longer respond to movement. If the vestibule or the posterior canal does not completely

ossify, symptoms such as movement-triggered vertigo or positional vertigo can be noted in addition to constant unsteadiness. Electronystagmographic examination usually reveals spontaneous nystagmus and absent caloric responses in patients with transverse fractures of the temporal bone. In symptomatic patients, caloric excitability and positional nystagmus indicate residual function in the traumatized labyrinth. In these rare patients, a case can be made for a transmastoid labyrinthectomy to eliminate all vestibular function.

Post-traumatic vestibular dysfunction can also occur if the central portion of the vestibular system is injured, such as that seen with cerebellar injury, brainstem nuclei injury, or even vestibular nerve avulsion. It should be noted here that these types of injuries nearly always involve multiple cranial nerve injuries, widely varied neurologic dysfunction, and prolonged unconsciousness following the injury. Careful evaluation with audiovestibular testing can determine if a predominantly central vestibular injury has occurred, as might be suggested by eye tracking abnormalities on ENG in the presence of normal symmetric caloric tests or abnormal auditory brainstem testing with normal otoacoustic emissions. Treatment considerations in these patients might include vestibular rehabilitation therapy, special glasses, and antinauseants.

Post-traumatic Cochlear Dysfunction

In addition to PLF, sensorineural hearing loss following trauma can be attributable to a variety of mechanisms including fracture of the otic capsule, concussion of the inner ear without fracture, noise- or blast-induced injury, and central auditory pathway injury. Transverse fractures of the temporal bone usually cross the vestibule or the basal turn of the cochlea, resulting in total, sudden sensorineural hearing loss. The exact nature of the loss remains a matter of conjecture, that is, it is possibly caused by loss of fluid mechanics, hair cell dysfunction, or vascular injury to the cochlea. It does not appear that most transverse temporal bone fractures result in significant cochlear neuronal injury. This is based on the information that three of four cochlear implants performed in the United States on patients with transverse fractures are functioning extremely well.

It is common to find high-frequency sensorineural hearing loss following head injury with or without fracture, blast injuries, and extreme noise exposure (ie, transient noise louder than 120 dB). These three mechanisms affect the cochlea in a similar fashion. Most often the hearing loss centers around 4 kHz, although more severe injuries can certainly affect all of the frequencies. According to Schuknecht, concussive forces on the cochlea result in a pressure wave that damages the outer hair cells in the basal turn of the cochlea, features identical to those seen with noise-induced hearing loss.[13] Alternatively, deceleration, blast injuries, and extreme noise transmit forces to the stapes footplate that result in disruption of the basilar membrane and the organ of Corti, with resultant degeneration of the spiral ganglion of the cochlea. Some common causes of this type of injury include air-bag detonation in automobile accidents, gunshot blasts, and loud music speakers in vehicles and in night clubs.

If high-frequency sensorineural hearing loss is found after a head injury, blast injury, or extreme noise injury, the author recommends bed rest, liberal fluid intake, a short course of oral corticosteroids, and avoidance of loud noise. In some patients, improvement in thresholds is observed 2 weeks after the onset. If not, these patients should receive auditory rehabilitation, including use of a hearing aid.

Post-traumatic hearing loss can also be found in injuries of the auditory nerve, brainstem, and temporal cortex. Because of the severity of the injury, most patients with this type of injury do not survive. However, the author has managed two documented and two likely cases of cochleovestibular nerve injury in the last 20 years. In one patient, the cochleovestibular nerve was avulsed from the brainstem, a situation discovered during the course of performing a translabyrinthine facial nerve exposure in which the facial nerve was followed from the stylomastoid foramen to the brainstem. This patient had bilateral temporal bone fractures and subsequently required a hypoglossal nerve to facial nerve crossover procedure to reanimate the face. This patient was operated after a prolonged period of altered mental status and was discovered to have a complete facial paralysis a year after the injury. The other three patients were seen after neurosurgical colleagues performed microvascular decompression of the facial and trigeminal nerves for hemifacial spasm or trigeminal neuralgia. In two of these patients, a complete hearing loss, vertigo, and facial palsy occurred, findings that may have been owing to either stretch-

ing of the cochleovestibular nerve or to occlusion of the vessels that feed the nerves. Neither of these patients recovered auditory or vestibular function, although both regained modest House-Brackmann grade III (see Chapter 24) facial function. One of these patients underwent a transmastoid labyrinthectomy a year after the injury, and the labyrinth in this case was not ossified, which suggests that the injury could not be attributable to vascular occlusion. In a third patient who postoperatively had a 90 dB sensorineural hearing loss, an abnormal auditory brainstem response, and normal otoacoustic emissions (findings that clearly suggest an auditory neuron injury), the hearing recovered to the 30 dB level at 6 months. In patients with temporal lobe injury following head trauma, typically there are complaints of difficulty understanding speech in noisy environments. In the few documented cases, speech understanding is indeed diminished as measured with hearing-in-noise tests or synthetic sentence intelligibility testing. Rehabilitation in these circumstances includes special listening devices and auditory training using neuroplasticity techniques (ie, fast forward therapy to improve temporal processing).

Vascular Injuries

Penetrating injuries of the temporal bone and skull base can cause injuries to the major vessels either in the temporal bone or immediately outside the temporal bone as these structures enter or exit from it. Vascular injury may include transection, laceration, thrombosis, and aneurysm formation. Vascular injuries can be quite alarming because of the volume of blood loss. Immediate management includes stabilization of the cardiovascular system with fluid and blood replacement, compression with dressings, and expedient evaluation with arteriography, MRI, magnetic resonance angiography, or high-resolution CT scans. Furthermore, since cranial nerves are in close proximity to the internal carotid artery and jugular vein, an evaluation for their function should be carried out as soon as possible. This specifically includes cranial nerves VI through XII, with particular attention to the vagus, spinal accessory, and hypoglossal nerves. Life-threatening hemorrhage can be quickly managed by an interventional neuroradiologist using balloon occlusion, although this procedure is usually avoided because of the possibility of acute or delayed stroke. Hemorrhage from the sigmoid sinus or the jugular bulb can be managed surgically by gaining transmastoid exposure of the sinus and placing extraluminal packing at the site of the injury.

The most common vascular injury in the temporal bone is to the lateral sinus with subsequent occlusion of the jugular bulb and lateral sinus. Venous occlusion rarely requires intervention, although delayed hydrocephalus is a concern in these patients. Occlusion of the petrous portion of the internal carotid artery is also common following injury. Arterial vascular occlusion is ordinarily managed through careful observation with radiologic evaluation. In some patients, occlusion of the distal part of the internal carotid artery is performed to prevent extension of a clot or an embolus into the downstream cerebral circulation. Occasionally, an aneurysm of the petrous or cavernous part of the internal carotid artery develops after a penetrating or blunt injury. Neurologic and radiologic consultation is recommended in this circumstance also. In these patients, the artery may be stented or may be bypassed using an external carotid or facial artery to middle cerebral artery bypass approach. In some cases, an aneurysm can be occluded using intra-arterial coils on the nondominant side.

In patients in whom the carotid artery is injured as it enters the skull, consultation with a vascular surgeon or neurosurgeon is recommended. In expanding lesions, exploration and repair may be indicated, and the otolaryngologist can provide additional exposure of the region through mobilization and retraction of the mandible, parotid gland, and facial nerve branches.

Eustachian Tube Injury

Penetrating injuries of the temporal bone occasionally cause blockage of the eustachian tube, usually through cicatrix formation. This problem may not be appreciated for a period of time after the injury, when persistent effusion in the middle ear space is noted. In some patients, a tympanostomy tube can provide sufficient ventilation to keep the middle ear open. In some patients, however, there is persistent infection and drainage through the tympanostomy tube. In these patients, exploration of the eustachian tube through a transmastoid facial recess approach is indicated. If the function of the eustachian tube cannot be restored, the patient usually prefers leaving the drumhead intact and the use of a hearing aid.

Injuries to Children in a Bomb Blast

As a direct result of the Oklahoma City truck bomb blast at the Alfred P. Murrah Federal Building, 816 adults and children were injured or killed. Among the 19 pediatric patients killed in this tragedy, 17 suffered skull fractures and 15 suffered cerebral evisceration. A variety of other injuries were found in this group, including abdominal and thoracic injuries, amputations, extremity fractures, burns, and contusions and lacerations.[14] Nonfatal injuries occurred in 47 pediatric patients, 7 of whom required hospitalization. Among the group hospitalized, 2 suffered depressed skull fractures, 2 closed head injuries, and 5 tympanic membrane perforations. A bomb blast is associated with a high incidence of cranial and temporal bone injury.

REFERENCES

1. Kavanagh KT, Litovitz T. Miniature battery foreign bodies in auditory and nasal cavities. JAMA 1986; 255:1470–2.
2. Capo JM, Lucente FE. Alkaline battery foreign bodies of the ear and nose. Arch Otolaryngol Head Neck Surg 1986;112:562–3.
3. Hallmo P. Extended high-frequency audiometry in traumatic tympanic membrane perforations. Scand Audiol 1997;26:53–9.
4. Panosian MS, Dutcher PO. Transtympanic facial nerve injury in welders. Occup Med (Lond) 1994; 44:99–101.
5. Ogren FP, Edmunds AL. Neuro-otologic findings in the lightning-injured patient. Semin Neurol 1995; 15:256–62.
6. Jones DT, Ogren FP, Roh LH, Moore GF. Lightning and its effects on the auditory system. Laryngoscope 1991;101:830–4.
7. Kristensen S. Spontaneous healing of traumatic tympanic membrane perforations in man: a century of experience. J Laryngol Otol 1992;106:1037–50.
8. Rybak LP, Johnson DW. Tympanic membrane perforations from water sports: treatment and outcome. Otolaryngol Head Neck Surg 1983;91:659–62.
9. Griffin WL. A retrospective study of traumatic tympanic membrane perforations in a clinical practice. Laryngoscope 1979;89:261–82.
10. Goodhill V. Sudden deafness and round window rupture. Laryngoscope 1971;81:1462–74.
11. Epley JM. Particle repositioning for benign paroxysmal positional vertigo. Otolaryngol Clin North Am 1996;29:323–31.
12. Semont A, Freyss G, Vitte E. Curing the BPPV with a liberatory maneuver. Adv Otorhinolaryngol 1988; 42:290–3.
13. Schukneckt HF. Mechanism of inner ear injury from blows to the head. Ann Otol Rhinol Laryngol 1969; 78:253–62.
14. Quintana DA, Parker JR, Jordan FB, et al. The spectrum of pediatric injuries after a bomb blast. J Pediatr Surg 1997;32:307–10.

CHAPTER 15

Occupational Hearing Loss

Peter W. Alberti, MB, PhD

Worldwide, exposure to excessive sound is the most common cause of acquired adult sensorineural hearing loss. It is incurable but should be entirely preventable. This chapter discusses the effect of noise on the ear, noise-induced hearing loss (NIHL), and its prevention.

Noise is any unwanted sound that at low intensity may be irritating and at high intensity may damage hearing. Noise has physical, physiologic, and psychological connotations, all of which differ. Physically, sound can be measured and is defined in terms of duration, frequency (measured in Hz), and intensity measured in sound pressure levels (SPLs) expressed as decibels (dB). Sound may be continuous, intermittent, impulsive, or explosive.

HISTORY

Although hearing loss produced by noise has been known for centuries, it became a major problem with the discovery of gunpowder and a more widespread problem with the Industrial Revolution. It was recognized in the United States, Germany, and the United Kingdom in the 1870s and 1880s. In 1886, Thomas Barr of Glasgow wrote, "It is familiarly known that boilermakers and others who work in very noisy surroundings are extremely liable to dullness of hearing. In Glasgow, we would have little difficulty finding hundreds whose sense of hearing has thus been damaged, by the noisy character of their work."[1] Indeed, the riveting associated with shipbuilding gave the generic title "boilermakers' deafness" to NIHL. Barr's work stands the test of time as an excellent, well-conceived, and well-executed epidemiologic survey.

The histology of noise-induced hearing loss in the organ of Corti was already described by Habermann in 1890.[2] Audiometrically, Fowler was the first to comment on the 4 kHz dip produced by noise,[3] and Bunch gave an excellent account of the audiometric features of NIHL in 1939.[4]

As a result of the technical advances of the Second World War, machinery became larger, more efficient, and noisier, and as developing countries have rapidly industrialized, NIHL has become a global problem. Currently, with the advent of amplified music and inexpensive motorized transportation, social, recreational, and community sound also plays a major disturbing role. So, although NIHL has been well recognized for a century, its evaluation, accurate quantification, and prevention have only recently become important. We each live in a personal soundscape.

EFFECT OF SOUND ON THE EAR

Sound, depending on its intensity, may produce (1) adaptation, (2) temporary threshold shift (TTS), and (3) permanent threshold shift (PTS) in hearing.

Adaptation, which occurs whenever the ear is stimulated by sound, is a physiologic phenomenon, and for sounds of 70 dB SPL or less, recovery occurs within half a second.

Temporary threshold shift is a short-term effect usually measured in minutes and hours rather than seconds or days, in which the hearing threshold elevates temporarily after exposure to noise. It occurs following intense sound stimulation. With a low-frequency stimulus, the maximum threshold shift may occur an octave above the center frequency of the stimulating tone; as the stimulating frequency increases, the point of maximum shift moves closer to the stimulating tone. Tones of higher frequency cause more TTS than tones of lower frequency of similar intensity. The amount of TTS increases for the duration and intensity of the sound exposure. It lasts from minutes to a few hours and is also a physiologic phenomenon. For further details on adaptation and TTS, the reader is referred to works by

Mills[5] and Kryter.[6] Permanent threshold shift is a permanent elevation of hearing threshold; the term is usually confined to post–noise exposure elevation of hearing threshold.

The relationship between TTS and PTS has been the subject of much investigation but remains unclear. Mills studied the effect of continuing sound stimulation and TTS on certain psychophysical responses of the ear such as temporal integration and simultaneous masking.[5] He found that, as long as the sound continued, TTS grew, but the rate of growth varied according to the frequency of the sound. Periodic rest periods are protective. At 4 kHz, no TTS occurs until the SPL reaches 74 dB. Thereafter, a 4 kHz sound produced an asymptotic threshold shift of 1.7 dB for each decibel increase in sound level. The lower the frequency, the higher the base level and the slower the growth of the TTS. Thus, at 1 kHz, there was no TTS below 82 dB. Figure 15–1 summarizes the findings of Mills and shows how difficult it is to define any one "safe" sound level in terms of a composite single intensity level. A safe sound varies according to its frequency components.

Most TTS recovers in a few hours. If the noise exposure is sufficiently intense, however, TTS may last for several days or weeks.[7] The cutoff point seems to be a TTS of approximately 40 dB. Below this, recovery is swift; above this, it is delayed. This is pathologic TTS, almost always accompanied by some residual PTS. One mechanism for TTS has been elegantly demonstrated by Puel and colleagues in a series of experiments that suggest that glutamate oversecretion at the synapses between hair cells and the auditory nerve fibers leads to synaptic overload, swelling of synaptic terminals, and damage both to the primary auditory nerve fibers and hair cells.[8] Both recovered. This led to much experimentation to produce biochemical blockage of the damage.

One of the practical problems of TTS is its effect on accurate quantification of hearing loss in both audiometric screening and pension evaluation.

PERMANENT THRESHOLD SHIFT

The main purpose of this chapter is to discuss PTS produced by exposure to intense sound. This is associated with pathologic changes in the cochlea, although the direct correlation between hearing loss and anatomic injury is not as clear-cut as was once thought.

HISTOPATHOLOGIC CHANGES RESULTING FROM EXCESSIVE SOUND EXPOSURE

The effect of sound stimulation on the ear has been studied by means of light microscopy and surface preparations, as well as scanning and transmission electron microscopy. An excellent review by Saunders et al pointed out that acoustic injury to the ear has a dynamic and a static phase.[9] The dynamic phase begins during acoustic stimulation, which results in cellular elements in the ear undergoing structural and functional changes that may be permanent or that may recover. Outer hair cells are more susceptible to damage. When the sound trauma ceases, the structure of the ear may recover completely or partially, scar, or disappear. This leads to a static phase in which hearing and the anatomic changes are stable. Cochlear changes following damaging noise exposure may thus be described as temporary/permanent/degenerative and reparative.

The difficulties of studying the cochlea should not be minimized. Human material is difficult to obtain and relatively unfresh; history regarding presbycusis and other ototoxic exposure such as infection and drug ingestion is usually unknown. Animal experimentation is essential. There is no absolute correlation between the results of animal experiments and NIHL or among electrophysiologic tests, behavioral audiograms, and histopathology in animals with noise-damaged hearing. Cytocochleograms performed for areas of hair cell loss frequently do not correlate with changes in behavioral hearing threshold. However, this "all or none" approach to cochlear damage—hair cells present or absent—is very crude. Current electron microscopy and molecular biology techniques allow detailed examination of individual hair cells and their components, which are able to show intermediate amounts of damage. Merely to look at hair cell destruction is too crude. When more subtle damage such as damage of the rootlet of the stereocilia or linkages between the cilia is looked for, there is much better correlation.[10,11] A critical level of noise exposure appears to exist. Below this level, there is biochemical and perhaps reversible change in the cochlea and, above it, major mechanical and irreparable damage.

In turn, physiologic thinking about the structure of the cochlea has moved from a macro- to a

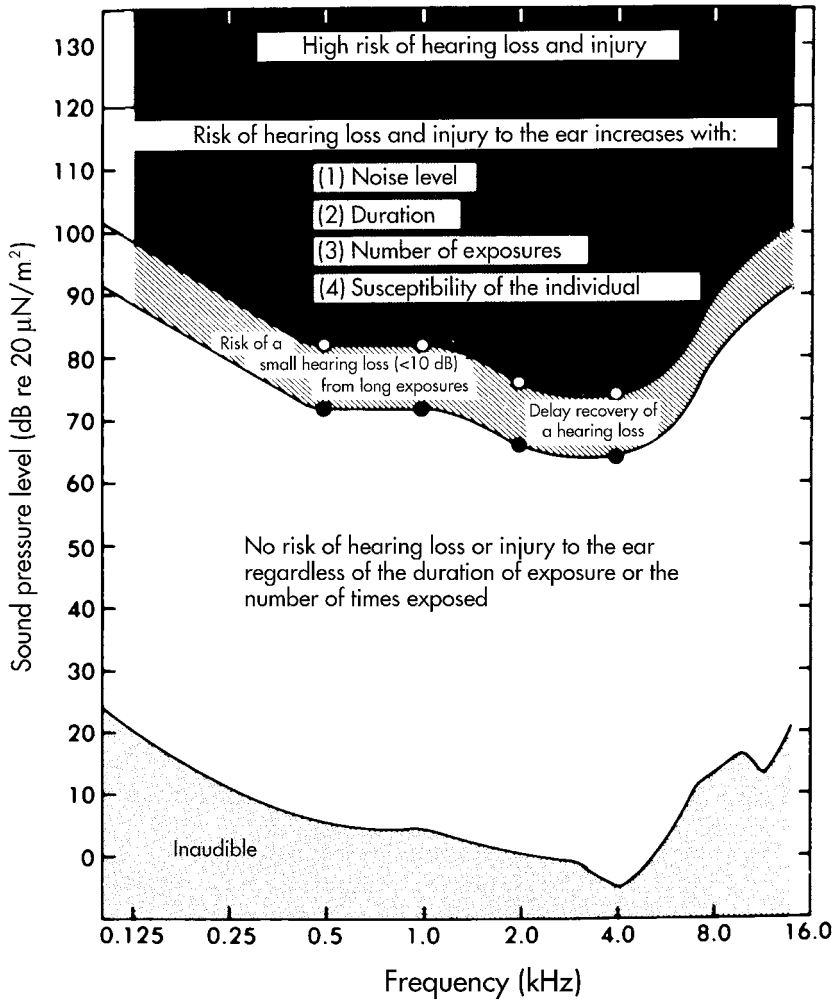


FIGURE 15–1. Most of the range of human audibility is categorized with respect to the risk of injury and hearing loss. Reproduced with permission from Mills JH.[5]

microdynamic model. No longer is the main focus of attention the traveling wave. Current studies emphasize the functional integrity of individual hair cells and their innervation, interaction between outer and inner hair cells, and knowledge obtained from single auditory nerve fibers and their characteristic frequencies, both normal and noise damaged. It is now apparent that outer hair cells contract in response to a sound stimulus; this enhances and sharpens the mechanical displacement transmitted to inner hair cells, which alone appear to transduce mechanical stimulation to initiate an eighth nerve afferent signal. The use of tracers such as horseradish peroxidase has allowed study of the concomitant central nervous system connection and damage. The reader is referred to work by Kiang et al[12] and to work by Schacht and Canlon for biochemical changes.[13]

HAIR CELL INJURY

Cochlear hair cells are damaged by exposure to excessive sound. The degree and type of damage are related to the amount of noise exposure. Newer electron microscopy techniques have allowed the demonstration of subtle changes that, when studied in conjunction with current views of cilia as containing actively contractile actin molecules, are beginning to lead to a better understanding of the damage that occurs. The first sign is shortening of the cilial rootlet.[11] This acute finding, produced by least damaging noise, is associated with floppy cilia and is partly reversible. The resistance of cilia to deformation after noise stress is reduced. The first signs of an irreversible change are damage to the intercilial connections and actual fracture of the

rootlet at the level of the cuticular plate. The cilia fall over and ultimately are absorbed. With these greater degrees of damage, intracellular changes also occur, including damage to lysosome granules, mitochondria, and the nucleus. As with other tissues in the body, there is a process of damage and repair. The repair with mild degrees of injury is within the cell itself; with greater degrees of injury, healing and scar formation occur by migration of other cells.[14] An elegant series of experiments by Gao et al demonstrated, by means of electron photomicrographs, responses of hair cells to intense and very intense stimuli; in the former, the cilia flop slightly but do not collapse, and they appear to recover completely; in the latter, they flop significantly and never recover; indeed, they are scavenged and disappear.[15] High levels of stimulation release excess of neurotransmitter in inner hair cells, which, in turn, leads to damage to associated nerve rootlets, which may be temporary or permanent.[8]

Over the past decade, the biochemical and molecular basis of NIHL has become clearer; this is leading to the possibility of fuller preventive and even curative strategies.[16,17] It appears that oxidative stress is a major cause of hair cell damage both in NIHL and antibiotic ototoxicity. Excessive stimulation by noise produces toxic levels of reactive oxygen species (ROS), which damage phospholipids in cell and nuclear membranes and deoxyribonucleic acid (DNA); it increases intracellular Ca^{++} and up-regulates cell death genes.[18] There may be ways of preventing or diminishing the formation of ROS, as, for example, by increasing endogenous antioxidant systems by administering substances such as *N*-acetylcysteine. The administration of glutathione, a powerful antioxidant, to guinea pigs prior to noise exposure leads to less cochlear damage and hearing loss than in matched noise-exposed animals that are not protected.[19] Antioxidants also protect the inner ear of chinchillas from the damaging effects of impulse noise.[20] Glial cell line–derived neurotrophic factor also provides protection from noise. It and heat shock protein (which protects hair cells against excessive stimulation) are produced when the ear is exposed to mild noise exposure and may account for "toughening" of the ear to noise by prior sound exposure. Although promising, none of these mechanisms currently have any practical application in humans.

It is also clear from the work of Puel and Pujol,[8] Duan et al,[21] and Canlon et al[22] that damage from noise and aminoglycosides affects nerves as well as hair cells. So protective strategies, although so far mainly based on hair cell protection from ROS, also include protection of axons by neurotrophins.

The development of NIHL is also related to feedback from the central nervous system (CNS). That nerve supply of the outer hair cells is by efferent fibers has been known for years, but their function was poorly understood. It now appears that there is an active feedback system, which may be responsible for depressing the contractile activity of outer hair cells, thereby diminishing the stimulus to the associated inner hair cells. This may function to sharpen frequency discrimination and perhaps to eliminate the effect of low masking sound.

This mechanism may also explain why exposure to enjoyable sound of equal intensity and duration produces less hearing loss than exposure to noise (eg, industrial of similar intensity and duration).

As early as 1988, Canlon et al showed that prior exposure to low-level sound "toughened" ears against subsequent exposure to harmful levels of sound.[22] Prior toughening appears to work by up-regulating the production of antioxidants such as glutathione.[21]

There has been much discussion in the literature about the role of other factors in noise damage, including potentiation by exposure to other industrial toxins such as carbon monoxide and various organic solvents. These will be described below.

AUDITORY NERVE AND CENTRAL NERVOUS SYSTEM DAMAGE AFTER ACOUSTIC INJURY

There is also evidence of eighth nerve and CNS injury after chronic traumatic acoustic exposure. There certainly is good evidence of damage to the dendrites and nerve endings surrounding the hair cells following acoustic overstimulation and, in chronic preparations, spiral ganglion cell degeneration spreading into the cochlear nucleus.[23]

PHYSIOLOGIC AND PSYCHOPHYSICAL CHANGES TO NOISE EXPOSURE

It has long been known, but only recently appreciated, that 90% of afferent nerve fibers from the

cochlea, which presumably carry information about sound stimulation to the CNS, arise from inner hair cells, and yet damage to outer hair cells is associated with hearing loss from noise exposure. Outer hair cells are overwhelmingly innervated by efferent fibers, and it is becoming clear that stimulation of efferent fibers can alter the sensitivity of the cochlea by 20 or 30 dB. In other words, outer hair cells have a tuning effect on inner hair cells. Outer hair cells themselves contract in response to sound stimulation, providing mechanical enhancement of the traveling wave motion transmitted to inner hair cells.

The contractions produce audible sounds detected as otoacoustic emissions (OAEs). These have been intensively studied to see if alteration of OAEs after sound exposure can be used to predict susceptibility to noise exposure. To date, there are no definite outcomes.[24]

Much psychophysical work has been directed toward single nerve fiber tuning curves, which have demonstrated a broadening of tuning curves of individual nerve fibers following cochlear damage. At their most sensitive, they are less frequency selective than normal, and this is associated with outer hair cell damage. In humans, the tuning curve of the cochlear nerve becomes broader with increasing hearing loss, and, certainly, damage from noise is associated with reduced suprathreshold hearing function.

NOISE-INDUCED HEARING LOSS

Intense acoustic overstimulation, whether acute or chronic, can produce TTS and PTS. The noise source may be continuous, intermittent, or mixed. Acoustic damage can also occur from a single intense sound such as an explosion or even a single gunshot. Typically, the first audiometric change in both TTS and PTS is a drop in hearing at 3, 4, or 6 kHz, with normal hearing below and above these levels. This is known as an acoustic notch. In the early phases, the subject usually recovers. It has long been known that the previously unexposed ear is more easily damaged by noise (green ears) than one accustomed to listening to loud sound,[25] and there is recent evidence of the "toughening of ears" by noise exposure.[22] This is a topic of great current interest and has been well reviewed by Hamernik and Ahroon[26] and Yoshida and Liberman.[27]

The recovery period following TTS is of interest. It has been studied in detail by Mills[28] in humans and in many animal studies. Most recovery takes place within the first 12 hours after noise exposure and is overwhelmingly complete within 48 hours, although in detailed animal studies, slight recovery continues for many weeks.[28] This recovery is so slight that it is within the confidence limits of normal clinical audiometric tests. However, after a single massive insult such as an explosion, recovery may continue for many weeks until the hearing loss is stable. This is usually accompanied by some PTS.

NOISE-INDUCED PERMANENT THRESHOLD SHIFT

Noise-induced permanent threshold shift (NIPTS) has many synonyms, one of the earliest being boilermakers' deafness and the currently popular one being NIHL. This is probably now the most common cause of acquired hearing loss in North American adults—indeed throughout the world—and is overwhelmingly caused by chronic exposure to excessive sound levels in the workplace, whether civilian or military. It may be potentiated by additional social and recreational noise exposure. The speed of onset and the degree of loss are related to the total quantity of noise exposure (ie, its intensity and duration) and to the individual's susceptibility to damage by the noise. In general terms, industrial noise exposure is less intense but more prolonged than military noise exposure.

It is difficult to quantify accurately the relationship between prolonged noise exposure and hearing loss because, over the years, workplace noise exposure changes, machinery is altered, jobs change, and hearing protection is provided. There are remarkably few long-term longitudinal studies equating hearing loss with noise exposure, and the likelihood of further studies in the future is becoming increasingly remote because, as the hazards of noise have been recognized, companies that are interested enough to take regular noise measurements also provide hearing protection. The well-studied group of jute workers in Dundee in the 1960s is the benchmark against which most of this work is measured.[29]

NATURAL HISTORY OF NOISE-INDUCED PERMANENT THRESHOLD SHIFT

There is interest in the progress of NIPTS over time. What is the rate of growth of hearing loss? Does it

spread into other frequencies? Does the rate of increase level off, accelerate, or remain the same? These are difficult questions to answer because of the lack of good longitudinal studies. It is unlikely that PTS will occur without TTS, and it is also unlikely that PTS will be greater than TTS. Other than that, there is little relationship between the two, and attempts to predict PTS from TTS have not been successful. Permanent threshold shift usually starts in the higher frequencies characteristically at 4 or perhaps at 3 or 6 kHz and gradually spreads into neighboring frequencies. To begin with, the hearing loss is asymptomatic, and only when it spreads into the lower frequencies such as 2 kHz does it give rise to complaints. As time goes on, the notch disappears for the hearing in the higher frequencies also worsens (perhaps as a result of aging rather than noise) so that the audiogram becomes indistinguishable from many other sensorineural hearing losses. It is stated, probably correctly, that after 10 years of exposure to an average workplace noise, say between 90 and 94 dB, the loss in the higher frequencies stops worsening, but gradually the loss spreads into the lower frequencies. The major part of the loss may occur quite early in the first 2 or 3 years,[30] and late losses are usually contaminated with presbycusis, which becomes the dominant cause of worsening hearing. The rate of spread and the degree of loss are related to both the intensity of sound and the individual's susceptibility to noise. This is as important with sound exposure as with other physical insults such as response to sunlight; individuals vary greatly, and, unfortunately, there is no good way of predicting the unduly susceptible individual. In mice, it has been shown that there is a genetic susceptibility to hearing loss from noise, and it is postulated that the same is true in humans, adding to predictive difficulties.[31] If the International Standards Organization (ISO) tables of risk are correct, it is probable that in later years too much hearing loss has been blamed on noise and too little on presbycusis. Furthermore, it appears that in mice, a recessive gene (*Ahl*) that is responsible for premature age-related hearing loss is also implicated in excessive susceptibility to NIHL.[32] If this is also true in humans, it may help explain some of the variability noted in response to sound exposure.

DAMAGE RISK CRITERIA

Many attempts have been made to define the relationship between noise exposure and hearing loss. It is currently accepted that, within a significant range with a lower boundary of safe sound and an upper boundary of sound so intense that it produces acute permanent damage, equal amounts of sound exposure produce equal amounts of damage. This is true whether sound exposure is spread over a short or long period. In other words, a given sound experienced for 8 hours does as much harm as twice that sound experienced for 4 hours or half that sound experienced for 16 hours. The reader must remember that decibels are a logarithmic scale. Doubling sound intensity is equivalent to a 3 dB increase in SPL. Thus, a 93 dB sound is twice as intense as a 90 dB sound, and a 103 dB sound is twice as intense as 100 dB sound. This equal energy concept, promulgated by Burns and Robinson,[33] has been widely adopted in Europe for many years and is gaining acceptance in North America. However, it is now suggested that if noise exposure is intermittent, as is usually the case in industry, the ear has time to recover and the 3 dB rule may be too strict. The best fit is probably a 4 dB halving and doubling, although no single rule fits all circumstances. This issue is well reviewed by Ward.[34]

The appropriate technique of evaluating the harmful effects of impact sound is still debated. Rifle fire is a good example of military impact sound; a drop forge, a pile driver, and a hammer blow are all examples of industrial impact noise. The characteristics of impact sound vary widely; they are measured by rise and decay time and peak intensity and are difficult to measure because of the transient nature of the sound. Hamernik and Hsueh give a good review of the nature and measurement of impulse sound.[35] It is probable that impulse noise is more damaging than steady-state noise of the same intensity.

Attempts have been made to codify risks into national and international standards for noise exposure, which are used both to define the upper limit of safe exposure and to quantify the risk of hearing damage at various levels of intense sound. International standards have been based on the equal energy (3 dB halving and doubling) concept and were initially embodied in ISO 1999 (New York, ISO, 1971) and revised as ISO 1999 (New York, ISO, 1984), and again as ISO 1999 (1990[E]),[36] which takes into account such factors as other ear disease and aging. In North America, tables based on 5 dB halving and doubling were introduced in the mid-1960s after a Committee On Hearing, Bioacoustics and Biomechanics

(CHABA) of the US National Research Council report in 1966 and codified in the Walsh-Healey Public Contracts Act of 1969 (Federal Register 34, no. 96, 1969). Here it is assumed that the trading relationship is 5 dB for halving and doubling of exposure, known as L_{OSHA} (US Occupational Safety and Health Administration), as opposed to the European 3 dB loudness equivalent known as Leq. There is a significant, but in practical terms frequently unnecessary, debate about these trading relationships. It is important to know that with a 90 dBA (dB on the A-weighted scale) sound exposure for 8 hours a day, 5 days a week, 15% of the population is at risk for significant hearing loss after 10 years of exposure, and that for 85 dBA exposure 8 hours a day, 5 days a week, after 10 years, only 7% of the population is at risk. It is prudent, therefore, to initiate hearing conservation measures from 85 dB upward,[37] and from a hearing conservation standpoint, little is to be gained from further dissection of these tables. Where they are significant is in attempting to equate total noise exposure in terms of medicolegal investigations. Few jobs have a steady noise exposure throughout the working day or working week, and attempts to equate actual noise measurements at work over a period of years in which several jobs have been held into a single risk figure for legal purposes are dramatically affected by which table of risk is used (Table 15–1). The tables of risk are based on human hearing threshold changes and are difficult to correlate with histologic damage and hair cell destruction in noise-exposed animals. Better correlation will be found as more subtle changes in hair cells are sought. The human studies are based on large samples of workers, including a large General Motors study and studies of Austrian industrial workers. They have been extensively evaluated in the literature in a series of publications by Passchier-Vermeer.[38]

TABLE 15–1. Comparison of Hours of Equivalent Sound Exposure*

Hours of Permitted Exposure	L_{OSHA}	L_{eq}
16	85	87
8	90	90
4	95	93
2	100	96
1	105	99
0.5	110	102
0.25	115	105

*Using an 8-hour 90 dBA baseline as allowed by L_{OSHA} and L_{eq}.

TINNITUS

Tinnitus is a distressing and frequent concomitant of NIHL. Transient tinnitus is experienced commonly after exposure to intense sound and almost invariably after a blast injury, but it usually clears. After years of exposure, however, it may become permanent and is present permanently in 50 to 60% of those with NIHL,[39] higher in those exposed to impact noise. The incidence and severity of tinnitus usually increase with increased hearing loss. This was not found by McShane et al in NIHL compensation claimants,[40] highlighting the difficulty of distinguishing the organic from the nonorganic components of this distressing but immeasurable complaint. There has been abundant literature about tinnitus in recent years: the reader is referred to a recent handbook edited by Tyler[41] and to Chapter 22.

WORSENING OF HEARING AFTER NOISE EXPOSURE ENDS?

It is frequently asked if hearing loss from noise exposure can progress after removal from noise. Although hearing loss improves rather than worsens after removal of chronic noise exposure, there is considerable controversy in the literature about the interaction among noise, presbycusis, and progression of hearing loss. In mice strains that show premature age-related hearing loss, which are also usually susceptible to the effects of excessive noise exposure, there is a suggestion that after cessation of noise exposure, hearing worsens more rapidly than in similar but non–noise-exposed mice.[31] The overwhelming body of opinion, however, strongly suggests that hearing loss that progresses after removal of noise exposure is from some other cause. Tschopp and Probst undertook a longitudinal study of hearing after exposure to a single sound; they found no worsening.[42] The matter remains controversial and is frequently raised in medicolegal claims.

COMBINED EFFECTS OF OTOTRAUMATIC AGENTS

Many things are known to injure the inner ear, for example, drugs, infections, trauma, age, and noise.

The ear is also affected by premature degenerative disorders such as familial hearing losses. Most of these agents have a final common pathway: they damage the hair cells and other structures of the cochlea. There is great interest in and concern about how they interact: whether they are synergistic, additive, or protective. Is the noise-exposed ear more or less affected by aging or is the effect additive? Is an old ear more likely to be damaged by noise than a young ear or is age irrelevant? Matters are compounded by species differences among animals and between animals and man, and it is not always reasonable to extrapolate from one to the other.

Unfortunately, the ears that are most readily studied and about which there is a large body of knowledge, those of the guinea pig and chinchilla, do not necessarily behave in the same way as human ears. With this caveat, in guinea pigs it has been shown that kanamycin given after noise exposure is more damaging than kanamycin alone, although the reverse is not true. Aminoglycoside antibiotics and excessive sound exposure both exert their effect on hearing by releasing free oxygen radicals, which explains their synergy; it also suggests ways of protection by exposing the hair cells to antioxidants. The ototoxic/cytotoxic drug cisplatin acts synergistically with other ototoxic drugs and noise in the same way.

Certain industrial solvents and chemicals are ototoxic. Jet fuel and styrene are predominantly vestibulotoxic but also appear to impair central auditory function. Rybak published an exhaustive review in which he suggested that there may be a synergistic effect between inhaling styrene and noise exposure. He draws similar conclusions about inhaled carbon disulfide.[43] Toluene is the most commonly used, most studied, and probably the most ototoxic industrial chemical. It produces hearing loss and potentiates the effect of noise exposure.[44] Morata et al have provided convincing epidemiologic studies demonstrating the synergistic effects of noise exposure and toluene[44,45] and between noise and petrochemicals inhaled in the oil refining industry. Noise and vibration act synergistically in those who suffer from white hand and work in the cold such as foresters using chain saws or shipbuilders.[46]

The interaction between noise exposure and aging is hotly debated. At most, it appears that presbycusis is additive to noise change, and even this is not demonstrated unequivocally. Studies by Macrea[47] and others before him (summarized by Robinson[48]) give some evidence for additivity, although they all say that there is no interaction. This matter is of considerable concern in terms of pension assessments, which are frequently made only at a time when presbycusis might already be present. ISO 1999[36] implies an additive mechanism, which has been modeled by Macrea (1991)[47] and Dobie (1992).[30] Robinson reviewed the subject at length.[48] It now seems that in later years, aging is a much more dominant cause of worsening hearing than further noise exposure; indeed, at the higher frequencies by the age of 80, hearing is the same, whether the ear is noise exposed or not! The best model known to the author is that of Corso, who gave a thoughtful discussion of the topic and concluded that, at best, there is limited additivity.[49] There is a good recent discussion by Dobie that concluded that except for high noise exposure levels, the effects are additive.[50]

Is an already noise-damaged ear more or less susceptible to damage from further noise exposure than an undamaged ear? This question is frequently raised when hearing-impaired workers attempt to move from one noisy industry to another only to fail to find employment in the belief that, because of previous damage to their ears, their hearing is more likely to degenerate more rapidly than in a normal ear. The little existing evidence suggests that this is wrong and that the ear will continue to degenerate at a steady rate with a natural history appropriate for it.

ACOUSTIC TRAUMA

Explosion and other single loud sounds can cause hearing loss. Early examples were "telephone ear," produced by static in early telephones. Other examples include workers inside metal tanks, the outside of which is struck by sledge hammers, quarrymen too close to a blasting accident, and armed forces personnel near a single explosion. The most recent examples are hearing loss produced by the inadvertent ringing of a cordless telephone into the ear of the listener[51] and the detonation of air bags in automobile accidents.[52] Perhaps, globally, the most pervasive examples at present are firecrackers and high-power weapons. An older but excellent review of the topic is given by Phillips and Zajtchuk.[53]

SOCIOACUSIS

The ear does not distinguish between occupational and social noise exposure. All noise is additive. Leisure pursuits may add significantly to total daily noise exposure and turn a marginally safe job, from a noise standpoint, into an acoustically hazardous one. The noise levels of everyday life are steadily increasing, and to this should be added the voluntary sound levels of recreation. Transportation is the main cause; automobiles, particularly on thruways, motorcycles, trains, planes, and the substitution of diesel for gasoline engines have all helped to produce sound levels that in some cities (eg, Bangkok) exceed 100 dB in the city center. There is voluminous literature about this, and the reader is referred particularly to *The Effects of Noise on Man* by Kryter[6] and the recent World Health Organization (WHO) document on community noise.[54] The worker exposed to sound levels of 88 dBA for an 8-hour workday who then experiences 94 dBA for 2 hours traveling to and from work is at risk for hearing loss from noise because of the additive effects of the sounds. At home, too, the blender, vacuum cleaner, and lawnmower all add to noise. In North America, ambient sound level ranges from 20 dBA on the north rim of the Grand Canyon to over 80 dBA in city centers near motorways. Noise levels within public transportation are high; some subway systems have sound levels above 90 dBA. Public pressure for noise reduction can be effective, and the quieting of airplanes is a major result of such activity. Environmental noise is a particular hazard in the megacities of Asia, with transportation noise being the dominant sound source, although by no means the only one. Two-stroke engine noise from motorcycles and motorized rickshaws (frequently unsilenced); old, poorly maintained diesel-powered trucks and buses; air horns; bazaar noise; and sounds from small roadside metal-forming shops all add to produce a cacophony that not only makes living unpleasant but also puts hearing at risk.[55]

Recreational noise also is hazardous. Firecrackers can produce sudden deafness in children, and their use in public festivities, such as national day celebrations, or religious ceremonies, may produce permanent hearing loss.[56,57]

Recreational vehicles such as snowmobiles and motorcycles are frequently considered by their users to be more fun if loud. Chain saws and radial-arm saws produce damaging sound levels for the hobbyist. There is much concern about the use of personal radios and cassette and compact disc players and the potential harmful effect of pop, rock, and disco music, although it would appear that they are not harmful. Clark reviewed current North American sources of potentially harmful social noise and highlighted the sound output of power tools and horticultural equipment.[58] Children from rural areas appear particularly at risk from firearms, tractors, and other farm equipment.[59] Indeed, farm workers are a major at-risk group for occupational hearing loss, working long hours, driving diesel tractors and other equipment, and being exposed to grain dryers, chain saws, and pigs (which may squeal in excess of 100 dB).

MEDICAL NOISE

Compression chambers may be an aural hazard from both pressure equalization and noise near the air valves. There is at least one reported case of deafness inside such a chamber.[60] Drills used for ear surgery have raised concern; although it appears that airborne sound levels from drills are high, Kylen et al concluded that for the patient, the mixture of air- and bone-conducted sound may be damaging, particularly if large burs are used.[61] The smaller the bur is, and the greater the number of teeth, the less the sound level. Man and Winerman, however, studied bone-conduction thresholds in operated and normal unoperated ears and found that the normal unoperated ear was unaffected by drilling.[62] Because bone-conducted sound is transmitted across the skull without loss, this is strong evidence that sensorineural loss in chronic otitis media is caused by factors other than bur noise. Concern has also been raised about the noise of suction units that, at the ear, is extremely intense.

Hospitals contain many noise hazards. Apart from normal commercial industrial workshops, dishwashers, and air-conditioning units, sound levels in intensive care units and incubators are high. The noise in intensive care units can certainly be disturbing. The background signals from many life support devices working at once make hearing and localizing warning signals difficult. Magnetic resonance machines are very loud at the patient's head

position, and use of rubber or foam earplugs is advocated.

NONAUDITORY EFFECTS OF NOISE

There is great controversy about noise as a cause of hypertension. This is a difficult subject to untangle. Certainly, sound may precipitate acute cardiovascular changes, but these are almost certainly normal physiologic responses to warning signals. Whether this leads in turn to chronic change remains debatable. Large population studies suggested that living in areas of high noise is more likely to produce hypertension, although a well-controlled study in the US Air Force by Von Gierke and Harris failed to demonstrate such an effect.[63] If it exists, it is extremely small. Sleep deprivation produced by loud sounds is also a matter of concern. These and other factors are summarized by Kryter[6] and in the WHO guidelines for community noise.[54]

There have been persistent reports of an association between excessive community noise exposure and mental disease. These have been extensively reviewed by Kryter.[6] There probably is no direct association. This has been confirmed by a careful study by Stansfeld et al.[64] Children's education may be hindered by chronic noise exposure because of interference with speech perception and because of the development of strategies to block out sound. The behavior of teachers may also be affected.[65]

SOCIAL ANTHROPOLOGY OF HEARING LOSS

The interaction between the individual with hearing impairment and the general population is currently much studied and is known as the social anthropology of hearing loss. One aspect is the impact on the lives of those with occupational hearing loss and their families. The middle-aged male with NIHL is isolated from his family; his children may become alienated as they feel that their father is disinterested in their activities, and he, in turn, feels that he is being excluded from family conversation. There may be considerable stress between husband and wife; communication becomes more difficult, producing anger, fatigue, and feelings of guilt on both sides.[66,67]

CLINICAL FEATURES OF NOISE-INDUCED HEARING LOSS

There are no clear-cut clinical features that distinguish NIHL from several other causes of sensorineural hearing impairment. The diagnosis is based on history, physical examination, and appropriate laboratory investigation including full hearing tests. Although there are audiometric configurations suggestive of the diagnosis, they may have many variations, and no one configuration fits all cases or excludes a case. To compound matters, NIHL, like any other chronic disorder, may coexist with other lesions. The diagnosis, which is circumstantial, is largely based on a careful history and is frequently made by exclusion, that is, if other causes of hearing loss have been excluded, noise exposure has been adequate, and there is an appropriate hearing loss, it is customary to attribute the loss to that cause. The diagnosis must be made individually, and one should avoid the epidemiologic error of believing that noise exposure and hearing loss are necessarily causally related. To believe this is to believe that working in noise protects one from all other forms of ear disease. In a large consecutive series of claims for NIHL taken at random from a working population in Ontario, at least 5% had other ear disease as the major cause of their hearing loss.[68]

Noise-induced hearing loss may be superimposed on other diseases such as chronic middle ear disease, Meniere's disease, familial hearing loss, otosclerosis, and the like. Unfortunately, people's memory about their own health is short, and little current information is available in the literature about the prevalence of ear disease in industrial populations.

The history includes questions about familial hearing loss, length of symptoms, childhood problems, school screening (whether school hearing screening has ever been failed), ear discharge, other diseases related to hearing loss such as renal problems, potential use of ototoxic drugs, and head injuries. The audiogram quantifies the hearing loss. Its pattern may give a clue toward the diagnosis because it may help to distinguish among various sites of lesion, conductive, cochlear, or retrocochlear. Where indicated, further investigation should be initiated, including hematologic, serologic, and imaging with computed tomography and magnetic resonance. When noise exposure is believed to be the cause of the hearing loss, various venues of loss must

be sought (eg, recreational, military, and occupational). Even within the realm of occupational hearing loss, it is often difficult to attribute the cause to a specific employer and ultimately may become a matter of clinical skill and exquisite judgment to establish both the cause of the hearing loss and the proportion of it that should be attributed to any given employer.

The responsibility of the assessing physician is to the patient: to quantify the hearing loss accurately, establish a diagnosis, and make a recommendation about treatment or rehabilitation. If physical signs or symptoms suggest other ear disease, this should be followed through, and, in particular, causes of asymmetric hearing loss should be investigated. It is possible to have an asymmetric hearing loss from noise exposure, such as, for example, individuals who fire guns off one shoulder, concert violinists, workers who use certain industrial equipment such as rock drills, and people who drive farm tractors with one ear to the engine and the head constantly turned over the shoulder, protecting the other ear. A dosimeter study showed a 9 dB difference in sound level between the ears when using a heavy handheld drill to make holes in a concrete form. Unequal hearing thresholds were present in about 10% of a large series of claimants for occupational hearing loss,[69] much of which was attributable to other ear disease, as well as asymmetric industria1 exposure from machinery and unequal damage to ears from explosions at work. The significance of asymmetric hearing loss in military personnel has recently been evaluated by Caldera and Pearson[70] There are recent thoughtful discussions about the diagnosis of NIHL from a North American standpoint by Dobie[50] and from a British view by Coles et al.[71]

HEARING TESTS

Hearing tests used in occupational hearing loss are of two types, screening and diagnostic. Screening tests will be dealt with under "Hearing Conservation Programs." This section deals with diagnostic tests. The degree of accuracy required here is greater than in routine clinical diagnostic audiology; it cannot be overemphasized that the tester must be experienced in this type of work. The basis for most pension and compensation awards is the pure-tone audiogram, which must be accurate to be fair to both employer and employee. At the time of testing, there should be no TTS. Authorities have legislated different time periods out of noise for this, ranging from 6 months to 48 hours. It is the author's view that if the claimant has been subjected to chronic noise exposure, a period of 48 hours free of noise is adequate for compensation assessment, and, in fact, anything much longer is impractical, particularly if the employee is continuing to work. The situation is different after a major blast accident or head injury, in which recovery may continue for a period of many weeks, and a compensation assessment should not be undertaken for a minimum of 3 months after the episode.

The types of tests to be undertaken vary by local use rather than any intrinsic factors related to the tests. They must be undertaken with properly calibrated equipment and in appropriately calibrated soundproof enclosures. Although screening tests make much use of the automatic Békésy audiometry, a test to quantify hearing loss accurately requires standard behavioral audiometry: air-conduction, bone-conduction, and speech reception thresholds. Further tests are undertaken as indicated. If there is an asymmetric loss, site-of-lesion tests are performed; if there is doubt about the accuracy of the behavioral test, the appropriate special tests are performed.

Many audiometric tests are used to identify exaggerated hearing loss, but there are few that quantify it. Quantification for compensation purposes must be frequency-specific and parallel tonal audiometry. In unilateral hearing loss, the Stenger test is effective. In bilateral hearing loss, electric response testing is necessary, and in this group of patients, slow vertex response (SVR) audiometry is the most effective. Auditory brainstem responses, currently much in vogue, usually provide only a click threshold, which gives a general idea of the level of hearing but not a simulacrum of an audiogram. To obtain this by means of brainstem audiometry requires expensive filtering techniques and a much lengthier test. The relative merits of various evoked response tests have been well described by Hyde et al[72] and the SVR audiometry in this type of patient by Alberti et al.[73] They use certain guidelines to identify the patient who should have SVR audiometry and place considerable emphasis on the discrepancy between speech reception threshold and pure-tone average and on the 500 Hz air-conduction threshold. Klockhoff et al[74] and Alberti et al[75] both demon-

strated that the probability of a hearing loss of 40 dB or greater at 500 Hz being caused by something other than noise is high. These patients produce a high yield of other ear disease and of nonorganic hearing loss.

THERAPY AND REHABILITATION

Experimental work suggests that symptoms of an acute traumatic hearing loss can be helped by the use of hyperbaric oxygen[76] or carbogen,[77] but this treatment does not have widespread acceptance. The use of antioxidants is also still experimental in that there is no easy delivery system to the cochlea.

The clinical onset of occupational hearing loss is usually insidious, but ultimately the complaints are those of any sensorineural hearing loss: difficulty in distinguishing one voice from many, difficulty in hearing in crowds, family complaints about the loudness of the television set, and inattentiveness. The psychological impact on the worker and his or her family has recently received considerable attention. Denial of problems, withdrawal, and lack of family interaction may have a severe effect on interpersonal family relationships.[66,67]

Much can be done to ameliorate the impairment with both specific and general devices. We have found a telephone handset amplifier (obtainable from most telephone companies), a louder bell, and a specific television amplifier to be the most useful. Many patients use a hearing aid to amplify television, but this also amplifies all other sounds in the environment, such as paper rustling, children crying, dishes clattering, and other people talking. A specific amplifier or television such as a headphone or extension loudspeaker serves both to make the sound louder and to improve the signal-to-noise ratio. Hearing aids are useful for and well accepted by those with NIHL.[78] They should not be withheld.

HEARING CONSERVATION PROGRAMS

Noise-induced hearing loss is incurable but largely preventable. The technique of prevention is known as hearing conservation. A hearing conservation program has four main features: sound measurement, engineering and administrative controls, personal hearing protection, and audiometric monitoring of the population at risk, with the following elements: (1) noise hazard identification, (2) engineering controls, (3) personal hearing protection, (4) monitoring, (5) record keeping, (6) health education, (7) enforcement, and (8) program evaluation.

A good hearing conservation program is multidisciplinary, involving the industrial hygienist, engineer, nurse, audiometric technician, and frequently supervisory audiologist and otologist. Good reviews of such programs are found in specific texts on hearing conservation.[79] Sound measurement is a skilled task devised to answer two questions: (1) What are the absolute sound levels in the work place? and (2) What is the sound exposure of the individual? Carefully performed sound surveys measuring sound levels throughout the plant and throughout the working day indicate the absolute environmental levels and help identify those who may be exposed to potentially hazardous noise. The exposure of individual workers, if the job is static, may be extrapolated from these measurements, but it is far better to undertake individual noise dosimetry studies. Dosimeters are small but expensive devices that sample and store the total sound received. The measurement of impact and impulsive noise is particularly difficult and may require special laboratory equipment.

A great effort must be made to have total collaboration of the workforce; to be successful, it requires bottom up planning and implementation and top down guidance.

SOUND ABATEMENT AND ADMINISTRATIVE CONTROL

The most effective way of reducing the risk of NIHL is to engineer out noise at the source by either silencing machinery or enclosing it. It is sometimes possible to retrofit machinery and provide sound-silencing barriers. Vicarious noise exposure should be eliminated; for example, there is no need to have a workbench beside a drop forge and to expose the mechanic, as well as the forge operator, to the noise of the forge. This happens all too often. In many industrial plants, it is possible to put the equipment operator inside a sound-enclosed booth and separate him from noisy machinery, for example, in ships' engine rooms, hydroturbine halls, and paper-making plants.

PERSONAL HEARING PROTECTION

Hearing protectors are widely used where noise levels remain hazardous even after appropriate engineering and administrative controls have been put in place. It is, however, not enough to provide a protector; there must be a program to encourage their use and instruct in their proper fitting and maintenance. The plant physician has a major role to play in this area and should be familiar with the various types of protectors, their advantages and disadvantages, and their application. Earplugs are by far the most commonly used; they constitute about 85% of personal protector devices in use in the United States.[80] Earplugs should fit firmly and may require a few days of use before they are comfortable. The preformed plugs of the V51R or fir cone Comfit type must be sized appropriately for the individual user. Workers usually choose one size too small because it is more comfortable. Universal plugs, which are compressed before being placed in the ear canal and expand to close it, are widely used and range from waxed wool to the commonly used foam polyurethane EAR or Deci-Damp. They work well for most people but are difficult to fit into those with small ear canals, particularly some women. Soft, individually molded plugs, which were widely used, have not proven effective on a large scale. Theoretically, the most effective type of hearing protector is the earmuff, consisting of a solid cup containing a sound-absorbing material such as foam, sealed to the side of the head with a soft malleable gasket and applied with sufficient pressure to produce an airtight seal. Muffs are, however, less comfortable to wear than plugs, and hard hats and safety glasses interfere with their performance. They tend to be reserved for heavy industry and industries in which intermittent use of a protector is appropriate. Figure 15–2 shows the average attenuation of a wide range of hearing protectors under average-use conditions —a far cry from manufacturers' claims. Because the majority of industrial noise lies in the mid–90 dB range, however, 10 dB of attenuation are, in fact, effective. One must guard against inappropriate protection as well as underprotection.

One of the problems with hearing protectors is that they distort external sounds by attenuating high-frequency more than low-frequency sound; another is that they exclude ambient signals as well as unwanted noise. Increasing use is therefore being made of active hearing protectors consisting of good-quality muffs with built-in microphone and amplifier, which, however, cut out above certain predetermined ambient noise levels. They act much like a good-quality compression hearing aid with less and less amplification as the ambient noise rises, and, above a predetermined figure, usually 85 dB, the amplifier automatically cuts out when the device becomes a passive hearing protector. They are particularly useful in intermittent and impact noise and have been widely adopted by hunters who wish to hear the sounds of the countryside but to protect their ears from the sound of gunshots. These muffs have widening application.[80] Modifications have also been made to passive muffs by the incorporation of filters that flatten the attenuation curve and provide a more natural environment for their user.[81]

HEARING SCREENING

A crucial feature of any hearing conservation program is the regular monitoring of workers' hearing. New employees must be screened and the effectiveness of the program monitored by routine, usually annual, hearing tests. The annual hearing test has many functions; important among them is the opportunity it provides to reinforce the hearing conservation program and to monitor proper use of hearing protectors. It detects those with hearing loss and identifies individuals with changes and whether a total program is effective. It indicates whether a program is working, whether an individual worker needs counseling, or if the whole plant needs further instruction. There are practical problems related to hearing screening. When should one do the test? Is TTS a problem? It used to be taught that screening should not be performed unless the worker was out of noise for 16 and preferably 48 hours before the test. This is impractical, and now testing is done throughout the workday, for in an effective program, there should be no TTS. Those who fail can be retested after a period in quiet.

What is a failure? Many suggestions have been made in the literature. The current belief is that a change of 15 dB or more in two or more frequencies should be an indication for a further referral. Ten dB changes are usually assumed to be within the realm of test–retest variability of screening audiometry. Proposed new federal regulations have been delayed.

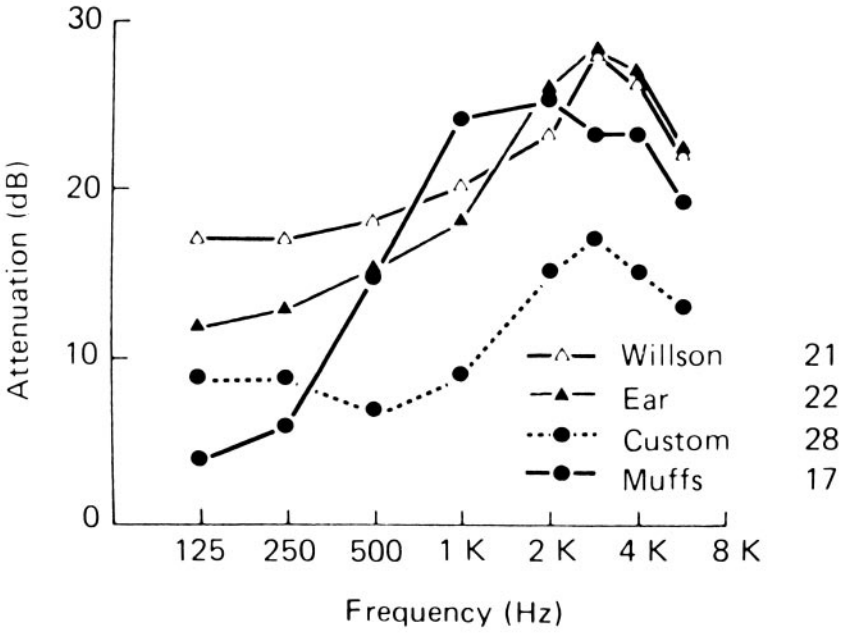


FIGURE 15–2. Mean attenuation characteristics of three types of earplugs and a group of earmuffs, as issued to industrial workers and fitted by them. Reproduced with permission from Alberti PW, Riko K, Abel SM, Kristensen R. The effectiveness of hearing protectors in practice. J Otolaryngol 1979; 8:354–9.

A further question is how, and for how long, to store records. This issue has been significantly reduced by the introduction of computer programs designed to deal with the mass of figures generated, which are becoming widely used. The ideal system of hearing screening includes direct computer storage of the audiometric results from the audiometer and a printout of the current hearing test results, the most immediate prior one, and the first one on record.

MEDICAL, LEGAL, AND SOCIAL IMPLICATIONS

There is no clear-cut method of determining the impairment produced by hearing loss, and definitions vary. In turn, there is little congruity about levels at which compensation should be paid, for what, and for how long. One may compensate for loss of earnings or loss of enjoyment of life. Compensation may be paid only while retraining occurs or for life.

Differences in philosophy have given rise to widely differing practices. The American Academy of Otolaryngology-Head and Neck Surgery suggests that hearing loss should be compensated if the average loss of the frequencies 500, 1,000, 2,000, and 3,000 exceeds 25 dB. States vary widely in their interpretation of these recommendations.

REFERENCES

1. Barr T. Transactions of the Philosophical Society of Glasgow 1886;17:223–39.
2. Habermann J. Uber die Schwerhortighkelt des Kesselschmiede. Arch Ohrenhelik 1890;30:1–25.
3. Fowler EP. Limited lesions of the basilar membrane. Arch Otolaryngol 1928;8:151.
4. Bunch CC. Traumatic deafness. In: Fowler EP, editor. Nelson's loose leaf medicine of the ear. New York: Thomas Nelson; 1939. p. 349–67.
5. Mills JH. Effects of noise on auditory sensitivity, psychophysical tuning curves and suppression. In: Hamernik RP, Henderson D, Salvi R, editors. New perspectives on noise-induced hearing loss. New York: Raven Press; 1982. p. 249–63.
6. Kryter KD. The effects of noise on man. 2nd ed. New York: Academic Press; 1985.
7. Melnick W. Human temporary threshold shift (TTS) and damage risk. J Acoust Soc Am 1991;90:147–64.
8. Puel JC, d'Aldium C, Ruel J, et al. Perspectives in inner ear pharmacology and clinical applications. In: Prasher D, Canlon B, editors. Cochlear pharmacology and noise trauma. London: NRN Publications; 1999. p. 1–12.
9. Saunders JC, Cohen YE, Szymko YM. The structural and functional consequences of acoustic injury in the cochlea and peripheral auditory system: a five-year update. J Acoust Soc Am 1991;90:136–46.

10. Liberman MC. Chronic ultrastructural changes in acoustic trauma. Serial section reconstruction of stereocilia and cuticular plates. Hear Res 1987;16: 65–88.
11. Liberman MC, Dodds LW. Acute ultrastructural changes in acoustic trauma. Serial section reconstruction of stereocilia and cuticular plates. Hear Res 1987;26:45–64.
12. Kiang NYS, Liberman MC, Sewell WF, et al. Single unit clues to cochlear mechanisms. Hear Res 1986;22: 171–82.
13. Schacht J, Canlon B. Biochemistry of the inner ear. In: Alberti PW, Ruben RJ, editors. Otologic medicine and surgery. New York: Churchill Livingstone; 1988. p. 151–78.
14. Slepecky N. Overview of mechanical damage to the inner ear—noise as a tool to probe cochlear function. Hear Res 1986;22:307–21.
15. Gao WY, Ding DL, Zheng XY, et al. A comparison of changes in the stereocilia between temporary and permanent hearing losses in acoustic trauma. Hear Res 1992;62:27–41.
16. Henderson D, McFadden SL, Zheng XY, et al. Intervention possibilities for noise induced hearing loss. In: Prasher D, Canlon B, editors. Cochlear pharmacology and noise trauma. London: NRN Publications; 1999. p. 85–95.
17. Miller JM, Altschuler R. Prevention of noise-induced hearing loss. Noise and Health 2000;3:83.
18. Ohinata Y, Miller JM, Altschuler R, Schacht J. Intense noise induces formation of vasoactive lipid peroxidation products in the cochlea. Brain Res 2000; 878:163–73.
19. Ohinataab Y, Yamasobac T, Schacht J, Miller JM. Glutathione limits noise-induced hearing loss. Hear Res 2000;146:28–34.
20. Henderson D, McFadden SL, Liu CC, et al. The role of antioxidants in protection from impulse noise. Ann N Y Acad Sci 1999;884:368–80.
21. Duan M, Agerman K, Ernfors P, Canlon B. Complementary roles of neurotrophin 3 and N-methyl-D-aspartate antagonist in the protection of noise and aminoglycoside-induced ototoxicity. Proc Natl Acad Sci U S A 2000;97:597–602.
22. Canlon B, Borg E, Flock A. Protection against noise trauma by pre-exposure to low level acoustic stimuli. Hear Res 1988;34:197–200.
23. Liberman MC, Mulroy MJ. Acute and chronic effects of acoustic trauma: cochlear pathology and auditory nerve pathophysiology. In: Hamernik RP, Henderson D, Salvi R, editors. New perspectives on noise induced hearing loss. New York: Raven Press; 1980. p. 105–35.
24. Marshall L, Miller JL, Heller LM. DPOAES as screening tool for NIHL. Noise and Health 2000;3:76–7.
25. Glorig A. Noise and the ear. New York: Grune & Stratton; 1958.
26. Hamernik RP, Ahroon WA. Susceptibility of the noise-toughened auditory system to noise-induced trauma. Hear Res 1999;132:140–8.
27. Yoshida N, Liberman MC. Sound conditioning reduces noise-induced permanent threshold shift. Hear Res 2000;148:213–9.
28. Mills JH. Threshold shifts as produced by a 90-day exposure to noise. In: Henderson D, Hamernik RP, Dosanjh D, Mills JH, editors. Effects of noise on hearing. New York: Raven Press; 1976. p. 265–75.
29. Taylor W, Pearson JCC, Mair A, et al. Study of noise and hearing in jute weaving. J Acoust Soc Am 1965; 33:13–120.
30. Dobie RA. The relative contributions of occupational noise and aging in individual loss of hearing. Ear Hear 1992;13:19–27.
31. Li HS, Borg E. Auditory degeneration after acoustic trauma in two genotype mice. Hear Res 1993;68: 19–27.
32. Davis RR, Newlander JK, Ling X, et al. Genetic basis for susceptibility to noise-induced hearing loss in mice. Hear Res 2001;155:82–90.
33. Burns W, Robinson DW. Hearing and noise in industry. London: HM Stationery Office; 1970.
34. Ward WD. The role of intermittence in PTS. J Acoust Soc Am 1991;90:164–9.
35. Hamernik RP, Hsueh KD. Impulse noise—some definitions, physical acoustics and other considerations. J Acoust Soc Am 1991;90:189–96.
36. ISO 1999. 2nd ed. New York: International Standards Organization; 1990.
37. Lutman ME. What is the risk of noise-induced hearing loss at 80, 85, 90 dB(A) and above? Occup Med (Lond) 2000;50:274–5.
38. Passchier-Vermeer W. Noise induced hearing loss from exposure to intermittent and varying noise. In: Proceedings of International Congress on Noise as a Public Health Problem. Washington (DC): US Environmental Protection Agency; 1983. p. 169–200.
39. Axelsson A, Barreras ML. Tinnitus in noise induced hearing loss. In: Dancer AL, Henderson D, Salvi RJ, Hamernik RP, editors. Noise induced hearing loss. St. Louis: Mosby Year Book; 1992. p. 269–76.

40. McShane D, Hyde ML, Alberti PW. Tinnitus prevalence in occupational hearing loss claimants. Clin Otolaryngol 1988;13:323–30.
41. Tyler RS, editor. Tinnitus handbook. San Diego: Singular; 2000.
42. Tschopp K, Probst R. Acute acoustic trauma. A retrospective study of influencing factors in different therapies in 268 patients. Acta Otolaryngol (Stockh) 1989;108:378–84.
43. Rybak LP. Hearing—the effects of chemicals. Otolaryngol Head Neck Surg 1992;106:677–86.
44. Morata TC, Dunn DE, Kretschmer LW, et al. Effects of occupational exposure to organic solvents and noise on hearing. Scand J Work Environ Health 1993;19:245–54.
45. Morata TC, Fiotini AC, Fischer FM, et al. Toluene induced hearing loss among rotogravure printing workers. Scand J Work Environ Health 1997;23: 289–98.
46. Iki M, Kurumatami N, Hirata K, et al. Association between vibration-induced white finger and hearing loss in forestry workers. Scand J Work Environ Health 1986;12:365–70.
47. Macrea JM. Presbyacusis and noise induced permanent threshold shift. J Acoust Soc Am 1991;90:2513–6.
48. Robinson DW. Relation between hearing threshold level and its component parts. Br J Audiol 1991; 25:93–103.
49. Corso JF. Age correction factor in noise-induced hearing loss: a quantitative model. Audiology 1980; 19:221–32.
50. Dobie RA. Medical-legal evaluation of hearing loss. 2nd ed. San Diego: Singular; 2001.
51. Singleton GT, Whitaker DL, Keim RJ, Kemker FJ. Cordless telephones: a threat to hearing. Ann Otol Rhinol Laryngol 1984;93:565–8.
52. Cunningham CD, Weber PC, Cure J. Neurotologic complications associated with deployment of airbags. Otolaryngol Head Neck Surg 2000;123:637–9.
53. Phillips YY, Zajtchuk JT. Blast injuries of the ear in military operations. Ann Otol Rhinol Laryngol Suppl 1989;140:3–4.
54. Berglund B Lindvall T, Schwelas DH, editors. Guidelines for community noise. WHO 1999. www.who.int/peh/noise.guidelines2.html. 2001.
55. Roy B, Das P. Status of road traffic noise in Calcutta metropolis, India. J Acoust Soc Am 1997;101:943–9.
56. Gupta D, Vishwarama SK. Toy weapons and firecrackers—a source of hearing loss. Laryngoscope 1989;99:330–4.
57. Axelsson A, Hellstrom PA, Altschuler R, Miller JM. Inner ear damage from toy cap pistols and fire-crackers. Int J Pediatr Otorhinolaryngol 1991;21:143–8.
58. Clark WW. Noise exposure from leisure activities—a review. J Acoust Soc Am 1991;91:175–81.
59. Kramer MB, Wood D. Noise induced hearing loss in rural school children. Scand Audiol 1982;11:279–80.
60. Hughes KB. Sensorineural deafness due to compression chamber noise. J Laryngol Otol 1976;90:803–7.
61. Kylen P, Stjernvall JE, Arlinger S. Variables affecting the drill generated noise levels in ear surgery. Acta Otolaryngol (Stockh) 1977;84:252–9.
62. Man A, Winerman I. Does drill noise during mastoid surgery affect the contralateral ear. Am J Otol 1985; 6:334–5.
63. Von Gierke HE, Harris CS. On the potential association between noise exposure and cardiovascular disease. In: Berglund V, Karlsson J, Lindvall T, editors. Noise as a public health problem. Part 1, ed. v. Stockholm: Swedish Council for Building Research; 1990.
64. Stansfeld SA, Gallacher J, Babisch W, et al. Road traffic noise, noise sensitivity and psychiatric disorder —preliminary prospective findings from the Caerphilly study. In: Vallet M, editor. Noise as a public health hazard. Vol 3. Bron: Institute Nationale de Recherche sur les Transports et leur Sécurité (INRETS); 1993. p. 268–73.
65. Bronzaft AL. The effects of noise on learning, cognitive development and social behaviour. In: Fay TH, editor. Noise and health. New York: New York Academy of Medicine; 1991. p. 87–92.
66. Hetu R, Lalonde M, Getty L. Psychosocial disadvantages associated with occupational hearing loss as experienced in the family. Audiology 1987;26:141–52.
67. Hallberg LR-M, Barreras ML. Living with a male with noise induced hearing loss—experiences from the perspective of spouses. Br J Audiol 1993;27:255–61.
68. Alberti PW, Blair RL. Occupational hearing loss— an Ontario perspective. Laryngoscope 1987;92:535–9.
69. Alberti PW, Symons F, Hyde ML. Occupational hearing loss—the significance of asymmetrical hearing thresholds. Acta Otolaryngol (Stockh) 1979;87: 255–63.
70. Caldera SR, Pearson CR. Risk management of asymmetrical hearing impairment in an armed forces population. J Laryngol Otol 2000;114:345–9.
71. Coles RR, Lutman ME, Buffin JT. Guidelines on the diagnosis of noise-induced hearing loss for medicolegal purpose. Clin Otolaryngol 2000;25: 264–73.

72. Hyde ML, Alberti PW, Matsumoto N, et al. Auditory evoked potentials in audiometric assessment of compensation and medical-legal patients. Ann Otol Rhinol Laryngol 1986;95:514–9.
73. Alberti PW, Hyde ML, Riko K. Exaggerated hearing loss in compensation claimants. J Otolaryngol 1987; 16:362–6.
74. Klockhoff I, Drettner B, Svedberg A. Computerized classification of the results of screening audiometry in groups of persons exposed to noise. Audiology 1974;13:326–34.
75. Alberti PW, Morgan PP, Czuba I. Speech and pure tone audiometry as a screen for exaggerated hearing loss in industrial claims. Acta Otolaryngol (Stockh) 1978;85:328–31.
76. Demaertelaere L, Van Opstal M. Treatment of acoustic trauma with hyperbaric oxygen. Acta Otorhinolaryngol Belg 1981;35:303–14.
77. Witter HL, Deko RC, Lipscomb DM, et al. Effects of pre-stimulatory carbogen inhalation on noise-induced temporary threshold shifts in humans and chinchilla. Am J Otol 1980;1:227–32.
78. Riko K, Hyde MI, McShane D, Alberti PW. Hearing aid usage in occupational hearing loss claimants. J Otolaryngol 1990;19:25–30.
79. Berger EH, Royster LH, Royster JD, et al. The noise manual. 5th ed. Akron (OH): American Industrial Hygiene Association; 2000.
80. Alberti PW. Active hearing protectors. In: Berglund V, Karlsson J, Lindvall T, editors. Noise as a public health problem. Part 1, ed. V. Stockholm: Swedish Council for Building Research; 1990. p. 79–87.
81. Allen CH, Berger E. Development of a unique passive level-dependent hearing protector. J Acoust Soc Am 1987;Suppl 81:55.

SUGGESTED READING

Alberti PW, editor. Personal hearing protection in industry. New York: Raven Press; 1980.

Axelesson A, Borchgrevink HM, Hamernik RP, et al, editors. Scientific basis of noise-induced hearing loss. New York: Thieme; 1996.

Burns W. Noise and man. 2nd ed. London: Murray; 1973.

King PF, Coles RR, Lutman ME, Robinson DW. Assessment of hearing disability, guidelines for medicolegal practice. London: Whurr; 1992.

Morata TC, Dunn DE, editors. Occupational hearing loss. Occupational medicine state of the art reviews. Vol 10. Philadelphia: Hanley Begus; 1995.

Nilsson P, Arlinger S, editors. International symposium on effects of impulse noise on hearing. Scand Audiol 1971;Suppl 12.

Prasher D, Luxon L, Pyykko I, editors. Advances in noise research. Vol 2. Protection against noise. London: Whurr; 1998.

Rossi G, editor. Proceedings of the 4th International Congress: noise as a public health hazard. Torino, Italy: Amplifon; 1983.

Salvi RJ, Henderson D, Hamernik RP, Colletti V, editors. Basic and applied aspects of noise-induced hearing loss. New York: Plenum Press; 1985.

Schafer RM. The tuning of the world. Toronto: McClelland and Stewart; 1977.

Tobias JV, Jansen G, Ward WD, editors. Noise as a public health problem. Rockville (MD): American Speech-Language-Hearing Association; 1980. ASHA Report 10.

Ward WD. Proceedings of the International Congress: noise as a public health problem. Washington (DC): US Environmental Protection Agency; 1973.

CHAPTER 16

Ototoxicity

Leonard P. Rybak, MD, PhD, John Touliatos, MD

Ototoxicity can be defined as the capacity of a drug or chemical to damage inner ear structure and/or function. The damage may occur in the auditory or vestibular portion or in both parts of the inner ear. A large number of agents with the propensity for ototoxicity exist. This chapter deals with several of the more common drugs and chemicals likely to cause hearing loss and/or cochlear injury. These agents include aminoglycoside antibiotics, the antineoplastic drugs cisplatin and carboplatin, loop diuretics, and salicylates. The risk factors, pharmacokinetics, and monitoring of blood levels are discussed.

AMINOGLYCOSIDE ANTIBIOTICS

Although they were not the first drugs with adverse effects on the inner ear, it was the widespread use of aminoglycoside antibiotics in the late 1940s and 1950s as chemotherapy against tuberculosis that drew the attention of the medical community to the problem of ototoxicity.

The aminoglycosides provide bactericidal activity against gram-negative aerobic bacteria and include streptomycin, dihydrostreptomycin, tobramycin, gentamicin, neomycin, netilmicin, sisomicin, amikacin, and kanamycin. Netilmicin appears to be the least ototoxic drug in the group.

Amikacin and kanamycin have more of an effect on the cochlea than on the vestibular system, whereas streptomycin and gentamicin appear to exert greater effects on the vestibular system than on the cochlea, although auditory damage does occur with their use. However, with any of these drugs, both auditory and vestibular toxicity can occur simultaneously.

Because only about 3% of an orally administered dose is absorbed from the gastrointestinal tract, these drugs usually are given parenterally. Tissue concentrations of aminoglycosides are usually about one-third of corresponding serum concentrations. Penetration of aminoglycosides across the blood-brain barrier is minimal, except in neonates.

Aminoglycosides are excreted in the urine by glomerular infiltration; therefore, concentrations in the urine may exceed those in the serum. Impaired renal function reduces the rate of excretion and may lead to accumulation in the body. Doses are frequently adjusted downward in patients with decreased renal function.

In ototoxic doses, the cochleotoxic agents (neomycin, kanamycin, amikacin, sisomycin, lividomycin) appear to elicit patterns of injury to the cochlea in experimental animals that are similar to the patterns seen in human temporal bones. These compounds tend to cause selective destruction of the outer hair cells in the basal turn of the cochlea, with progression toward the apex as the dose and duration of treatment are increased.

Several prospective studies of aminoglycoside ototoxicity have been carried out. Fee[1] and Smith et al[2] compared gentamicin ototoxicity with tobramycin ototoxicity and found cochlear damage in 10 to 16% of patients and vestibular toxicity in 5 to 15%.

The ototoxic potential of aminoglycoside administration is determined by total serum concentration over time. In practice, however, monitoring peak and trough levels is more practical and provides reasonable protection against toxicity, as well as ensures that the therapy is adequate. Lerner and associates found an association of ototoxicity with elevated mean trough levels.[3]

Moore et al reported that bacteremia, elevated temperatures, liver dysfunction, and elevated serum urea nitrogen-to-creatinine ratios were risk factors associated with ototoxicity in prospective double-blind clinical trials of gentamicin, tobramycin, and amikacin.[4] In another study, univariate and multi-

variate analysis of risk factors of ototoxicity showed only age as a predisposing factor for toxicity.

One of the newly identified risk factors may be genetic susceptibility. Hypersensitivity to aminoglycoside ototoxicity can be transmitted by mothers as a genetic trait. A mitochondrial deoxyribonucleic acid (DNA) polymorphism, 1555[G], is associated with aminoglycoside-induced deafness.[5] The basis for this hypersensitivity to ototoxicity may be based on the three-dimensional structure of the ribosome, favoring the binding of aminoglycoside, leading to disruption of protein synthesis and cell death in the cochlea.[6]

Acute damage to the cochlea is often preceded by tinnitus, although ototoxic effects can occur in the absence of tinnitus. The loss of hearing that results from the use of these drugs initially affects the high frequencies. Since patients do not complain about or have noticeable hearing losses until they have 30 dB losses that include frequencies as low as 3,000 to 4,000 Hz, the detection of ototoxicity is difficult unless audiometric monitoring is performed. Monitoring is particularly important in high-risk patients, such as those with renal insufficiency, those on higher-than-usual drug doses, those with persistently elevated serum levels despite dosage adjustments, or those with some degree of a baseline hearing or balance disorder. When audiometric measurements are obtained, the usual definition of ototoxicity is a hearing loss of 10 dB bilaterally at any frequency; some authors have used 20 dB as the cutoff. Again, since the usual threshold for perception of hearing loss by patients is a 30 dB drop, there will be at least that much of a loss by the time evaluation of the patient's complaints is begun.

The formation of free radicals (reactive oxygen species, reactive oxygen metabolites) by aminoglycosides has come to the forefront as the etiologic mechanism responsible for ototoxicity. The observation that gentamicin may act as an iron chelator was key to establishing that the formation of free radicals by aminoglycosides is an important mechanism in the development of ototoxicity.[7]

Several protective models are currently being investigated to prevent free radical–induced apoptotic cell death in cochlear hair cells. Those models that have been experimentally successful include administration of appropriate neurotrophins or growth factors, administration of caspase-specific inhibitors that remove the essential catalyst caspase from the apoptotic cascade, and gene therapy to enhance expression of genes that will synthesize neurotrophins or other protective substances.

ANTINEOPLASTIC DRUGS: CISPLATIN AND CARBOPLATIN

Cisplatin

After a 1-hour intravenous infusion of 70 mg/m^2 of cisplatin, the plasma platinum concentrations were found to have a biphasic clearance with half-life values of 23 minutes and 67 hours. Seventeen percent of the administered dose was excreted in the first 24 hours. Renal excretion is primarily by glomerular filtration. Ninety percent of cisplatin is bound to serum proteins,[8] and this cisplatin-protein complex is inactive against tumor cells. The serum levels of non–protein-bound platinum display different kinetics than those found for total platinum levels in serum. It is not known whether toxic metabolites of cisplatin are formed either in the inner ear or in the other parts of the body. However, the liver has been shown to convert cisplatin into nontoxic metabolites within 1 hour.

Early after its introduction, standard cisplatin doses were 50 mg/m^2. Subsequent cisplatin treatment regimens have been developed using higher doses of cisplatin (100 to 120 mg/m^2 = high dose and 150 to 225 mg/m^2 = very high-dose regimens). These increased-dosage regimens have resulted in a higher incidence of hearing loss than that observed with the lower dosage of 50 mg/m^2, as well as some different perspectives on cisplatin-induced hearing loss. In a study of 54 patients receiving high-dose cisplatin, Laurell and Jungnelius found that 81% of the patients had significant threshold elevations (15 dB or more at one frequency and 10 dB or more in three frequencies).[9] After therapy, which ranged from one to seven courses, 41% of the patients had significant deterioration of hearing in the speech frequency range of 500 to 2,000 Hz. Twenty-five percent of patients lost 25% of their remaining high-frequency hearing after each course. Preexisting hearing loss did not seem to predispose to ototoxicity, but advanced age was slightly associated with risk of hearing loss. The audiogram after the first course did not predict future deterioration of hearing during treatment with high-dose protocols. In this study, the ototoxic risk was determined more by the amount of the sin-

gle dose than by the cumulative dose levels. No ototoxic effects were seen at a peak plasma concentration of less than 1 μg/L. Based on their findings, the authors recommended that patients undergoing high-dose cisplatin treatment undergo audiometric testing before the initiation of therapy and before each of the subsequent courses. Less frequent testing was deemed necessary for patients receiving low- and moderate-dose treatment.[9]

Kopelman et al reported that all of their patients complained of decreased hearing after very high-dose cisplatin administration (150 to 225 mg/m^2). Ototoxic reactions appear to be more likely in patients with low serum albumin and in those with anemia.[10]

The exact incidence and severity of cisplatin-induced auditory effects is difficult to assess because of inconsistencies including variable dosing in previous studies and the lack of complete data from patients too ill to cooperate for pretreatment and post-treatment audiograms. Symptoms that strongly suggest cisplatin ototoxicity include otalgia, tinnitus, and subjective hearing loss. Tinnitus has been reported in 2 to 36% of patients receiving cisplatin. Often the tinnitus is transient, lasting from a few hours up to a week after cisplatin therapy. The incidence of hearing loss among patients treated with cisplatin has been as low as 9% and as high as 91%.[11] In patients with head and neck cancer treated with cisplatin, about half develop hearing loss.[12] The hearing loss is usually bilateral and appears first at high frequencies (6,000 and 8,000 Hz). Progression to lower frequencies (2,000 and 4,000 Hz) may occur with continued therapy. The hearing loss may be asymmetric and may not appear until several days after treatment. Patients may experience some degree of reversibility, but when the hearing loss is profound, it appears to be permanent. Because the hearing loss tends to occur at the higher frequencies, it may escape detection without audiometry. Cochlear toxicity may be detected earlier with high-frequency audiometry (up to 20 kHz) than with conventional audiologic testing.[13] Speech discrimination scores may be markedly reduced when cisplatin ototoxicity occurs. The hearing loss may be gradual, progressive, and cumulative or may present suddenly. Several patterns of hearing loss have been described.[14]

The critical cumulative dose of cisplatin has been reported as 3 to 4 mg/kg body weight. Since ototoxicity may be more severe after bolus injection, the ototoxic effects can be reduced by using slow infusion and dividing the doses over several months.[11] Bokemeyer et al evaluated a group of adult patients treated with cisplatin for testicular cancer.[15] Symptoms of ototoxicity persisted in 20% of patients (tinnitus and/or hearing loss). Ten percent had experienced completely reversible ototoxic symptoms that lasted from 1 to 18 months after treatment. Statistically significant risk factors for ototoxicity were (1) a high cumulative dose of cisplatin, (2) a history of noise exposure, and (3) a high dose of vincristine, which seemed to cause reversible ototoxic symptoms. Persistent ototoxicity may occur in about 20% of patients with testicular cancer who are treated at a standard dose but may affect more than 50% of patients receiving cumulative doses of cisplatin exceeding 400 mg/m^2. Previous noise exposure may also result in a threefold increased risk for cisplatin ototoxicity. Future studies should examine these risk factors as important stratification criteria for trials to prevent or, at least, minimize cisplatin ototoxicity.

A new method of treatment for superficially accessible tumors of the head and neck, esophagus, and trunk has been reported. Cisplatin is injected as a gel, which also contains epinephrine. The antitumor agent was injected weekly for 4 weeks in 45 patients. The initial dose of cisplatin was 1 mg/cm^3 tumor volume and escalated up to 6 mg/cm^3 as needed, depending on observed toxicities. The overall objective tumor response was 50%, with 40% complete response with a median-response duration of 160 days. No dose-limiting cisplatin-related toxicities, such as nephrotoxicity, neurotoxicity, or ototoxicity, were observed with this method.[16]

Children receiving high cumulative doses of cisplatin (above 540 mg/m^2) have a high incidence of hearing loss, which is cumulative and dose dependent.[17] However, a plateau effect has been reported with no further deterioration of hearing at doses greater than 600 mg/m^2. Still, a number of patients will develop more severe hearing losses in the 2,000 to 8,000 Hz range even after one or two courses of therapy, showing exceptions to the plateau effect.[18] Adults with a preexisting history of otologic problems are said to experience a higher incidence of ototoxicity of cisplatin affecting both lower (1,000 to 8,000 Hz) and higher (10,000 to 18,000 Hz) frequencies.[19] Ototoxicity of cisplatin appears to be enhanced by cranial irradiation.[20]

A study of otoacoustic emissions was carried out in pediatric patients treated with cisplatin chemotherapy. Transient evoked otoacoustic emissions could be measured in 11 of 12 patients studied. When the middle ear status was normal, a significant correlation was found between the transient evoked otoacoustic emissions and the pure-tone threshold, and 90.5% of patients had a significant sensorineural hearing loss at 8,000 Hz. Increased hearing loss was associated with young age at first dose of cisplatin, high number of chemotherapy cycles, and high cumulative dose of cisplatin.[21]

Temporal bones removed from patients with cisplatin-induced hearing loss have demonstrated the following: large, fused stereocilia; damage to the cuticular plate of the outer hair cells; and extensive loss of sensory cells in the vestibular labyrinth in specimens studied with scanning electron microscopy.[22] In addition to noting outer hair cell degeneration in the basal turn of the cochlea, Strauss et al reported degeneration of spiral ganglion cells and cochlear neurons in patients with documented cisplatin ototoxicity.[23] However, the vestibular neurons appeared normal. In a study of human temporal bones removed from patients treated with cisplatin, radiation, or a combination of both treatments, the spiral ganglion cells and inner and outer hair cells were decreased in number, and the stria vascularis was found to be atrophic. The otopathology identified in these temporal bones also included vascular changes, serum effusion, or fibrosis that accompanied sensorineural hearing losses.[24]

Carboplatin

Clinical Studies Carboplatin is a newer analog of cisplatin that was introduced into clinical trials in 1981. This drug was found to have less nephrotoxicity than does cisplatin. Initial reports suggested that carboplatin was also less ototoxic than is cisplatin. The primary dose-limiting toxicity of carboplatin has been bone marrow suppression.[25] This toxic effect has been overcome by the use of autologous stem cell rescue and/or the use of hematopoietic growth factors. This has allowed the use of higher doses of carboplatin and increased antitumor effectiveness.

A study of children with neuroblastoma who were treated with higher-dose carboplatin (2 g/m^2 total dose) showed that 9 of the 11 children (82%) had hearing loss in the speech frequencies that was severe enough that hearing aids were recommended. These children had all previously been treated with cisplatin, and several patients had previously received aminoglycoside antibiotics. The investigators concluded that high-dose carboplatin is ototoxic, especially in children who have been exposed to previous platinum therapy or other ototoxic agents. They felt that in children in whom hearing losses are inevitable, owing to cumulative ototoxic exposure, parents need to be properly informed of the balance of risks and benefits.[26]

Studies by Van Der Hulst et al[19] and Kennedy et al[25] have shown a significant incidence of hearing loss in carboplatin-treated patients. Seventy-five percent of patients treated with carboplatin had measurable hearing loss in the former study, whereas 20% of patients treated with carboplatin in the latter study had hearing loss. The authors in the latter study stated that no clinically significant hearing loss occurred.[26] On the other hand, cumulative sensorineural hearing loss has been shown in 11 of 22 children who received carboplatin chemotherapy.[27]

Carboplatin has been successfully used to treat malignant brain tumors when administered in conjunction with osmotic blood-brain barrier disruption with mannitol. Unfortunately, a larger percentage of patients developed high-frequency hearing loss when they were treated with this regimen. Twenty-nine patients received various doses of the protective agent sodium thiosulfate, given intravenously 2 hours after carboplatin. Doses of sodium thiosulfate were escalated from 4 g/m^2 to 8, 12, 16, and 20 g/m^2 in consecutive months. Audiograms were performed before treatment and monthly after treatment with carboplatin. These audiograms were compared with those of 19 historical control patients who were treated similarly but who did not receive sodium thiosulfate. The incidence of hearing loss in the control group was 79% (15/19). The group of patients protected with sodium thiosulfate lost only 3.7 ± 2 dB at 8,000 Hz after one carboplatin treatment, compared with the historical group, which had an average loss of 20.8 ± 5.9 dB. Patients in the sodium thiosulfate–treated group with excellent baseline hearing had little change in hearing thresholds at 8 kHz after the second treatment (8.0 ± 8.3 dB) compared with the historical control patients with excellent baseline hearing (40.5 ± 8.6 dB). These findings support the concept that sodium

thiosulfate in doses of 16 or 20 g/m^2 decreases carboplatin-induced hearing loss. Because the thiosulfate does not seem to cross the blood-brain barrier, there would appear to be no interference with the antitumor action of carboplatin in these patients.[28]

An intriguing study predicted carboplatin ototoxicity in patients with ovarian cancer, who were treated with a combination of cisplatin and carboplatin. A total of 105 patients with ovarian cancer who received up to six courses of carboplatin (300 mg/m^2) were evaluated, and values for first-course carboplatin area under the curve (AUC = area under the plasma drug concentration versus time curve, a measure of drug exposure) were determined retrospectively. Thrombocytopenia of grade 3 to 4 was found in 10% of patients with low AUC (< 4 mg/mL × min). No patients in the low AUC group had ototoxicity, but 12% of patients in the high AUC group were found to have ototoxicity and 44.6% had thrombocytopenia. These data could prove valuable for prevention of thrombocytopenia and ototoxicity for the first cycle of combined treatment with carboplatin and cisplatin.[29]

LOOP DIURETICS

Loop diuretics are a group of potent synthetic drugs that act on the loop of Henle to inhibit the reabsorption of sodium, potassium, and chloride ions by the Na^+-K^+-$2Cl^-$ cotransporter system. The ototoxicity of these agents has been reviewed.[30]

Most administered furosemide is excreted in the urine. Furosemide essentially follows a three-compartment pharmacokinetic model with an average half-life for renal elimination of 29½ minutes. Furosemide elimination decreases greatly in patients with advanced renal failure, in whom the half-life can be prolonged to 10 to 20 hours. Long half-lives also have been observed in patients with congestive heart failure, especially those undergoing long-term furosemide treatment. The plasma half-life of furosemide is approximately 45 to 92 minutes in healthy subjects but is prolonged to approximately 3 hours in patients with renal failure. The gastrointestinal uptake of furosemide is high; most authors have concluded that approximately 65% of an oral dose of furosemide is absorbed. Plasma levels exceeding 50 mg/L are frequently associated with hearing disturbances.[30]

Although the use of ethacrynic acid is very low currently, occasionally, it is used in patients who are allergic to furosemide or refractory to its diuretic effect.

A study by Arnold et al revealed some morphologic effects in the temporal bones from a patient who received a total of 5,000 mg of furosemide plus 250 mg of ethacrynic acid over 5 days before dying of renal failure.[31] The inner and outer hair cells were normal at both the light and electron microscopic levels, but the stria vascularis had marked cystic changes, as has been reported in animal studies. The dark cell areas of the vestibular system were found to have marked cystic changes, with dilation of the intracellular fluid spaces, suggesting effects on fluid transport within the inner ear.

Tuzel summarized the reported incidence of hearing loss, as documented by pure-tone audiometry in 179 patients treated with bumetanide and 62 patients receiving furosemide.[32] Among patients receiving bumetanide, only 2 (1.1%) had audiometric changes of 15 dB or greater. Patients treated with furosemide had a 6.4% incidence of hearing loss. Tuzel did not report whether these hearing losses were temporary or permanent. Although most cases of loop diuretic ototoxicity are temporary and fully reversible, some appear to be permanent. New loop diuretics are being developed and tested in animals, and some may prove to be even less ototoxic than bumetanide when clinical trials are performed.

Permanent, profound mid- and high-frequency sensorineural hearing loss in a transplant patient receiving ethacrynic acid illustrates that deafness from loop diuretics continues to be a problem that is clinically significant.[33] Permanent hearing loss has been reported in high-risk premature neonates treated with furosemide.[34]

SALICYLATES

Orally administered salicylates are absorbed very rapidly in the gastrointestinal tract, with an apparent half-life of 6 to 15 minutes. Absorption of salicylates is influenced significantly by gastric emptying time and by the presence of food in the stomach; food may more than double the absorption half-life of aspirin. Aspirin is broken down in the body to salicylic acid, with a biologic half-life of 15 to 20 minutes.[35] Salicylates are distributed mainly to the extracellular water compartments.[36] Concentrations of the salicylates are higher in the liver and kidney than in the serum, whereas brain concentrations are

usually about 10% of those in the serum. Salicylates are excreted mainly in the urine.

Salicylate concentrations of 20 to 50 mg/dL in serum have been associated with hearing losses of up to 30 dB.[37] Human volunteers receiving aspirin at various doses experienced hearing loss and tinnitus at relatively low concentrations of total plasma salicylate. A linear relationship between hearing loss and unbound salicylate concentrations has been reported.[38] Hearing loss can be observed at salicylate concentrations below 20 mg/dL, and there seems to be no threshold for salicylate-induced hearing loss. The severity of tinnitus increases as plasma salicylate concentrations exceed 40 mg/L. Work by Lue and Brownell links the reversible salicylate ototoxicity to reduction of lateral membrane stiffness in outer hair cells.[39]

REFERENCES

1. Fee WE. Aminoglycoside ototoxicity in the human. Laryngoscope 1980;90 Suppl 24:1–19.
2. Smith CR, Lipsky JJ, Laskin OL, et al. Double blind comparison of the nephrotoxicity and auditory toxicity of gentamicin and tobramycin. N Engl J Med 1980;302:1106–9.
3. Lerner SA, Seligsohn R, Bhattacharya I, et al. Pharmacokinetics of gentamicin in the inner ear perilymph of man. In: Lerner SA, Matz GJ, Hawkins JE, editors. Aminoglycoside ototoxicity. Boston: Little Brown & Co.; 1981. p. 357–69.
4. Moore RD, Smith CR, Leitman PS. Risk factors for the development of auditory toxicity in patients receiving aminoglycosides. J Infect Dis 1984;149:23–30.
5. Prezant TR, Agapian JV, Bohlman C, et al. Mitochondrial ribosomal RNA mutation associated with both antibiotic-induced and nonsyndromic deafness. Nat Genet 1993;4:289–94.
6. Hutchin T, Haworth I, Higashi K, et al. A molecular basis for human hypersensitivity to aminoglycoside antibiotics. Nucleic Acids Res 1993;21:4174–9.
7. Forge A, Schacht J. Aminoglycoside antibiotics. Audiol Neurootol 2000;5:3–21.
8. Gormley PE, Bull JM, LeRoy AF, Cysyk R. Kinetics of cis-dichlorodiammineplatinum. Clin Pharmacol Ther 1971;25:351–7.
9. Laurell G, Jungnelius U. High-dose cisplatin treatment: hearing loss and plasma concentrations. Laryngoscope 1990;100:724–34.
10. Kopelman J, Budnick AS, Sessions RB, et al. Ototoxicity of high-dose cisplatin by bolus administration in patients with advanced cancer and normal hearing. Laryngoscope 1988;98:858–64.
11. Moroso MJ, Blair RL. A review of cisplatinum ototoxicity. J Otolaryngol 1983;12:365–9.
12. Blakley BW, Gupta AK, Myers SF, Schwan S. Risk factors for ototoxicity due to cisplatin. Arch Otolaryngol Head Neck Surg 1994;120:541–6.
13. Fausti SA, Henry JA, Schaffer HI, et al. High-frequency monitoring for early detection of cisplatin ototoxicity. Arch Otolaryngol Head Neck Surg 1993;119:661–6.
14. Blakley BW, Myers SF. Pattern of hearing loss resulting from cisplatinum therapy. Otolaryngol Head Neck Surg 1993;109:385–91.
15. Bokemeyer C, Berger CC, Hartmann JT, et al. Analysis of risk factors for cisplatin-induced ototoxicity in patients with testicular cancer. Br J Cancer 1998; 77:1355–62.
16. Burris HA, Vogul CL, Castro D, et al. Intratumoral cisplatin/epinephrine-injectable gel as a palliative treatment for accessible solid tumors: a multi-center pilot study. Otolaryngol Head Neck Surg 1998; 118:496–503.
17. Kluba J, Kluba U, von Sprecht H, Mittler U. Zur Überwachung des Gehörs wahrend der ototoxischen Cisplatin-Therapie [Monitoring hearing during ototoxic cisplatin therapy]. Padiatr Grenzgeb 1990;29: 19–23.
18. Myers SF, Blakley BW, Schwan S, et al. The "plateau effect" of cisplatinum induced hearing loss. Otolaryngol Head Neck Surg 1991;1104:122–7.
19. Van Der Hulst RJ, Dreschler WA, Urbanus NA. High frequency audiometry in prospective clinical research of ototoxicity due to platinum derivatives. Ann Otol Rhinol Laryngol 1988;97:133–7.
20. Granowetter L, Rosenstock JG, Packer RJ. Enhanced cisplatinum neurotoxicity in pediatric patients with brain tumors. J Neurooncol 1983;1:293–7.
21. Allen GC, Tiu C, Koike K, et al. Transient-evoked otoacoustic emissions in children after cisplatin chemotherapy. Otolaryngol Head Neck Surg 1998; 118:584–8.
22. Wright CG, Schaefer SD. Inner ear histopathology in patients treated with cis-platinum. Laryngoscope 1982;92:1408–13.
23. Strauss M, Towfighi J, Lord S, et al. Cisplatinum ototoxicity: clinical experience and temporal bone histopathology. Laryngoscope 1983;93:1554–9.

24. Hoistad DL, Ondrey FG, Mutlu C, et al. Histopathology of human temporal bone after cis-platinum, radiation or both. Otolaryngol Head Neck Surg 1998;118:825–32.
25. Kennedy IC, Fitzharris BM, Colls BM, Atkinson CH. Carboplatin is ototoxic. Cancer Chemother Pharmacol 1990;26:232–4.
26. Parsons SK, Neault MW, Lehmann LE, et al. Severe ototoxicity following carboplatin-containing conditioning regimen for autologous marrow transplantation for neuroblastoma. Bone Marrow Transplant 1998;22:669–74.
27. Macdonald MR, Harrison RV, Wake M, et al. Ototoxicity of the carboplatin: comparing animal and clinical models at the Hospital for Sick Children. J Otolaryngol 1994;23:151–9.
28. Neuwelt EA, Brummett RE, Doolittle ND, et al. First evidence of otoprotection against carboplatin-induced hearing loss with a two-compartment system in patients with central nervous system malignancy using sodium thiosulfate. J Pharmacol Exp Ther 1998;286:77–84.
29. Obermair A, Speiser P, Thoma M, et al. Prediction of toxicity but not of clinical course by determining carboplatin exposure in patients with epithelial ovarian cancer treated with a combination of carboplatin and cisplatin. Int J Oncol 1998;13:1023–30.
30. Rybak LP. Ototoxicity of loop diuretics. Otolaryngol Clin North Am 1993;26:829–44.
31. Arnold W, Nadol JB, Weidauer H. Ultrastructural histopathology in a case of human ototoxicity due to loop diuretics. Acta Otolaryngol (Stockh) 1981;91: 399–414.
32. Tuzel IH. Comparison of adverse reactions to bumetanide and furosemide. J Clin Pharmacol 1981;21:615–9.
33. Rybak LP. Ototoxicity of ethacrynic acid (a persistent clinical problem). J Laryngol Otol 1988;102: 518–20.
34. Brown DR, Watchko JF, Sabo D. Neonatal sensorineural hearing loss associated with furosemide: a case control study. Dev Med Child Neurol 1991; 33:816–23.
35. Rowland M, Riegelman S. Pharmacokinetics of acetylsalicylic acid and salicylic acid after intravenous administration in man. J Pharm Sci 1968; 57:1313–9.
36. Hollister L, Levy G. Some aspects of salicylate distribution and metabolism in man. J Pharm Sci 1965; 54:1126–9.
37. Myers EN, Bernstein JM. Salicylate ototoxicity: a clinical and experimental study. Arch Otolaryngol 1965;82:483–93.
38. Day RO, Graham GG, Bieri D, et al. Concentration-response relationships for salicylate-induced ototoxicity in normal volunteers. Br J Clin Pharmacol 1989;28:695–702.
39. Lue AT, Brownell WE. Salicylate induced changes in outer hair cell lateral wall stiffness. Hear Res 1999;135:163–8.

CHAPTER 17

Idiopathic Sudden Sensorineural Hearing Loss

Robert A. Dobie, MD

DEFINITION, CLINICAL PRESENTATION, AND PROGNOSIS

Idiopathic sudden sensorineural hearing loss (ISSNHL) is a term that requires little elaboration. Most patients describe waking up with hearing loss that was not present the night before, or they can identify the time of day when hearing loss began. Nevertheless, the literature reflects a consensus to consider onset over a period of less than 3 days as "sudden."

Hearing loss, whether sudden or more gradual in onset, can obviously range from barely detectable to profound, but, once again, a convention has evolved: ISSNHL is generally defined as demonstrating a threshold change for the worse of at least 30 dB in at least three contiguous audiometric frequencies. Occasionally, this change can be assessed by comparing current to prior audiograms. More often, the affected ear is compared with the unaffected ear, assuming that the two ears were equal before. Fortunately, ISSNHL is almost always unilateral.

According to Byl, ISSNHL affects about 1 person per 10,000 per year[1]; this would correspond to over 25,000 cases per year in the United States. In contrast, Hughes and colleagues estimate only 4,000 cases per year in the United States.[2] All ages and both sexes are affected; the peak ages are 30 to 60 years.[3]

Tinnitus is usually present. Vertigo is frequent, either spontaneous (as in acute vestibular neuronitis) or as isolated positional vertigo.[4] Sixty-five percent of patients recover completely with or without treatment, most within 2 weeks.[5] Low-frequency losses have a better prognosis than high-frequency losses; of course, this is true of all types of rapidly progressive or fluctuating sensorineural hearing loss (SNHL), not just ISSNHL.[6]

ETIOLOGY

Since ISSNHL is, by definition, idiopathic, the cause of any individual case is not known. Nevertheless, there has been considerable speculation about viral and vascular causes, with rather more evidence for the former than the latter.

Viral

Many otologists believe that viral infection (or reactivation of latent viruses) is the most important cause of ISSNHL. The circumstantial evidence for that belief can be summarized as follows:

1. Many viral diseases (eg, mumps) can cause congenital or sudden SNHL.
2. Cochlear histopathology after these diseases resembles that seen after ISSNHL.
3. Patients with ISSNHL often demonstrate immunologic evidence for viral infection.
4. Viruses such as herpes simplex may lie dormant in neural tissue for years and have been identified in human spiral ganglia.

Intrauterine viral infection may damage many organ systems. The rubella (German measles) virus caused thousands of cases of congenital deafness during a 1964–1965 epidemic[7]; most were associated with a syndrome that included eye, heart, and brain abnormalities. Since the introduction of the mumps-measles-rubella (MMR) vaccine, the congenital rubella syndrome has virtually disappeared in the United States.

At present, the most important viral cause of congenital deafness is cytomegalovirus (CMV), a member of the herpes virus family. Cytomegalovirus has been cultured from the inner ear[8] and causes cochlear lesions in experimentally infected animals.[9]

Congenital CMV syndrome includes deafness and lesions of the eye, brain, liver, and spleen.[10] It is important to note, however, that most CMV-infected neonates do not demonstrate the full-blown CMV syndrome. Indeed, "asymptomatic" CMV infection causes more hearing loss than "symptomatic" (syndromic) CMV infection. Neonatal CMV screening has shown an infection rate (virus cultured from urine) of 1.3%[11]; 10% of these cases had either congenital or progressive hearing loss. Most of the babies who had (or developed) hearing loss were otherwise healthy and would have been missed by conventional high-risk criteria for congenital hearing loss. No vaccine is yet available to reduce or eliminate hearing loss from CMV.

Neonatal herpes simplex virus (HSV) infection is associated with hearing loss in about 10% of cases.[12]

Mumps (viral parotitis) is the prototypical example of sudden unilateral hearing loss in childhood. Rare since the widespread use of the MMR vaccine, mumps deafness occurred in about 5 of every 10,000 mumps cases[13] and was usually unilateral. Like CMV, the mumps virus has been cultured from the inner ear[14] and causes cochlear lesions in experimental animals.[15] Measles (rubeola virus) accounted for 3 to 10% of bilateral deafness in children prior to the MMR vaccine.[16]

The varicella-zoster (or herpes zoster) virus causes chickenpox in children and shingles in adults. When the geniculate ganglion is involved, with facial paralysis plus auricular bullae, it is called herpes zoster oticus (HZO) or Ramsay Hunt syndrome. About 6% of HZO cases exhibit SNHL,[17] making HZO probably the most common type of hearing loss in adults in which a viral cause is reasonably certain. Herpes zoster viral deoxyribonucleic acid (DNA) has been identified in temporal bone tissue taken from a patient with HZO and sudden hearing loss.[18]

Human immunodeficiency virus (HIV) may cause hearing loss directly but more commonly makes the host more susceptible to other viruses, many of which (CMV, HSV, adenovirus) have been cultured from the HIV-infected inner ear,[19] as well as to other infectious agents.

Schuknecht and Donovan first showed that ears from patients who had suffered ISSNHL demonstrated atrophy of the organ of Corti and tectorial membrane and spiral ganglion cell losses; these findings were similar to those seen in cases (mumps, measles) known to be caused by viral infections and different from the cochlear pathology (fibrous and osseous proliferation) seen after ischemic deafness.[20]

Given the otopathologic similarity between ISSNHL and hearing loss of known viral origin, investigators looked for immunologic evidence of viral infection in ISSNHL. Veltri and colleagues showed that patients with ISSNHL frequently had seroconversion (increasing antibody levels) for several viruses, including mumps, influenza (coincidental?), measles, HSV, rubella, and CMV.[21] Wilson focused specifically on HSV, noting that 16% of ISSNHL patients showed seroconversion compared with only 4% of controls.[22] The identification of DNA from HSV in human spiral ganglion[23] and the development of an animal model for HSV neurolabyrinthitis[24] have added to suspicion that this virus may be an important cause of ISSNHL. Pitkaranta and colleagues noted that there are many negative serologic studies, failing to implicate HSV or other viruses, but also noted that long-dormant neurotropic viruses can reactivate and cause localized disease without triggering changes in systemic immunoglobulin levels.[25] This is true for herpetic cold sores, for example, and may also be true for diseases such as ISSNHL, Bell's palsy, and vestibular neuronitis. Unfortunately, the difficulty of inner ear culture or biopsy makes proving this hypothesis in individual cases impossible.

Vascular

Diseases of blood vessels affect many organ systems and constitute some of the most important causes of death and disability in developed countries. Risk factors for vascular disease include diabetes, hypertension, obesity, and hyperlipidemia; controlling these risk factors can reduce the incidence of heart attacks, strokes, blindness, kidney failure, and premature death. Equally important (but uncontrollable) risk factors include age, sex, and family history. Since no organ can survive without adequate blood supply and since vascular disease affects so many different organ systems, it is natural that investigators have looked for evidence that SNHL is linked to vascular disease (or to its risk factors). Given the intensity of their efforts, the evidence is extremely weak.

In other organ systems, vascular disease may cause either gradual or sudden/stepwise deteriora-

tion of function. Congestive heart failure may develop insidiously in some patients, whereas others have sudden-onset myocardial infarctions. Even when the disease process (eg, atherosclerotic narrowing of arterial lumen) is gradual, the symptoms and signs of dysfunction are often sudden, when a critical point is reached or a new event such as thrombosis or embolism totally shuts off blood flow. If vascular disease is important for a particular organ (eg, the cochlea), one should expect it to cause both sudden and gradual changes in function.

Unfortunately, otologists cannot yet measure cochlear blood flow in clinical practice, so whether a person (or an ear) has cochlear ischemia cannot be determined. Reliance must be placed on correlations between hearing loss and vascular disease risk factors (or established vascular disease elsewhere in the body) and to a lesser extent on histopathologic evidence.

Consideration can begin with age-related hearing loss (ARHL). Age is not really a cause of SNHL but rather an association; still, after excluding common causes of SNHL (noise, head injury, etc), the vast majority of people with adult-onset SNHL cannot be labeled any more precisely than to call them age related. If vascular disease was an important component of ARHL, it might be expected that patients (and longitudinal studies of ARHL) would describe stepwise progression, with substantial asymmetry as one ear suffers ischemic events whereas the other escapes, at least for a while. The rarity of such reports in ARHL should already cast doubt on the vascular hypothesis, but those doubts can momentarily be put aside and one might ask whether vascular disease and its risk factors are correlated with ARHL. For example, do diabetics have more hearing loss than nondiabetics of the same age and sex? The extensive literature addressing this issue has been reviewed and can best be described as inconclusive.[26] If vascular disease is responsible for some proportion of ARHL, its contribution is small.

Focusing on patients presenting with ISSNHL, are some of these cases attributable to "vascular accidents" (thrombosis, embolism, or hemorrhage) affecting the cochlea? If vascular disease caused a substantial fraction of ISSNHL, the incidence of ISSNHL would be expected to be much higher in men than in women and to rise sharply with age; neither is true. Frequent recurrences and (over time) bilaterality would also be expected, yet most people who suffer ISSNHL have only one event, affecting only one ear. Studies of vascular disease and risk factors have shown only that diabetics with ISSNHL have a poorer prognosis,[27] not that any of these factors predict the likelihood of suffering ISSNHL.

None of this proves that vascular disease is totally irrelevant to ISSNHL. Inner ear hemorrhages with sudden deafness do occur rarely in patients with leukemia, sickle cell disease, and thalassemia,[28–30] although the temporal bone histopathologic findings after hemorrhage or ischemia are very different from those seen in patients who have suffered ISSNHL, as previously discussed. Sudden SNHL may[31] or may not[32] be an occasional complication of cardiopulmonary bypass surgery (presumably owing to emboli). Until otologists are able to assess cochlear blood flow reliably and noninvasively in their patients, we will probably never know much about the role of vascular disease in ISSNHL (or SNHL in general); at this point in time, it appears to be somewhere between small and negligible.

Other

Simmons postulated that double inner ear membrane breaks could be responsible for some cases of ISSNHL,[33] and it is hard to prove that he was wrong, but there is little evidence that he was right either. Because Simmons hypothesized that one of these membrane breaks would involve the oval or round window, producing a perilymph fistula (PLF), the 1980s saw a wave of enthusiasm for middle ear exploration, looking for PLF, in patients with ISSNHL. Today, most otologists believe that PLFs do occur, causing SNHL and dizziness, but almost exclusively in the context of identifiable barotrauma: after scuba diving, violent nose blowing, or extreme exertion during breath holding (eg, weight lifting with improper technique). In the absence of such a history, few otologists would diagnose membrane breaks and/or PLF or recommend therapy appropriate to such a diagnosis (bed rest, middle ear exploration).

DIFFERENTIAL DIAGNOSIS

Table 17–1 shows that dozens of otologic and systemic disorders have been associated with sudden SNHL. Many of these disorders (eg, Meniere's disease, autoimmune inner ear disease) typically cause

rapidly progressive or fluctuating hearing loss rather than truly sudden loss. The more important identifiable causes of sudden SNHL (by definition, these are not "idiopathic") are clinically easy to diagnose, for example, meningitis, acoustic trauma, head injury. Others, such as multiple sclerosis, are usually missed when sudden SNHL is the first manifestation. In areas where Lyme disease is endemic, it is probably wise to ask patients with sudden SNHL about recent rash or arthralgia.

TABLE 17–1. Some Causes of Sudden Sensorineural Hearing Loss

Infectious
Meningitis (bacterial, fungal)
Labyrinthitis (bacterial, fungal, viral, parasitic, spirochetal)
Traumatic
Head injury (with/without fracture)
Barotrauma (with/without perilymph fistula)
Acoustic trauma
Iatrogenic injury
Neoplastic
Acoustic tumor
Metastases (to meninges or temporal bone)
Hematologic malignancies (leukemia, myeloma)
Immunologic
Autoimmune inner ear disease
Systemic immune diseases (Cogan's, Wegener's granulomatosis, polyarteritis, temporal arteritis)
Ototoxic
Vascular (see text)
Neurologic
Multiple sclerosis
Focal pontine ischemia
Metabolic
Disturbance of iron metabolism
Renal failure/dialysis
Miscellaneous
Meniere's disease
Functional hearing loss

Adapted from Hughes GB et al.[2]

INVESTIGATIONS

Patients presenting with sudden unilateral hearing loss without obvious antecedent cause should receive a careful otologic history and examination, as well as an audiogram (including the Stenger test to detect functional hearing loss). Otoacoustic emissions (OAEs) are sometimes present (implying sparing of outer hair cells) in ISSNHL,[34] but it is unclear whether OAE testing is clinically useful, that is, whether the results can assist in selecting therapy.

Routine blood testing in ISSNHL is of dubious value. Children with ISSNHL (rare!) should receive a complete blood count as a screening test for leukemia. If there are clinical clues pointing toward Lyme disease, confirmatory testing is indicated. Patients with symptoms of systemic disease should probably be referred to an internist rather than receive blood tests selected by an otolaryngologist. If corticosteroid therapy is being considered in a patient who has not been tested for diabetes in recent years, a blood glucose test is probably in order.

About 1% of patients presenting with sudden SNHL have acoustic tumors,[35] and even complete recovery does not completely rule out this diagnosis. When hearing in the affected ear is quite poor (> 60 dB at 2 kHz), magnetic resonance imaging (MRI) with gadolinium contrast injection is the most appropriate study to detect or rule out a tumor. For patients with better hearing, either MRI or auditory brainstem response (ABR) testing (with MRI reserved for patients showing abnormal ABR) may be appropriate.

The increasingly frequent use of MRI in cases of sudden SNHL has shed additional light on the etiology. Precontrast films can document inner ear hemorrhage in the rare cases associated with hematologic disorders.[29] Soon after the onset of hearing loss, gadolinium enhancement of the cochlea is seen in typical cases of ISSNHL, probably owing to a transiently leaky blood-brain barrier.[36] Six months or more after onset, no fibrotic or osseous prolifera-

tion (such as would be expected after ischemic damage) could be identified.[37]

INTERVENTIONS

Wilson et al showed, in a randomized clinical trial, that patients receiving oral corticosteroids achieved substantial recovery (defined as more than 50% recovery vis-à-vis the uninvolved ear) more frequently than patients receiving placebo.[38] This benefit was apparent only for "moderate" degrees of loss (Figure 17–1); 78% of patients in this category who received corticosteroids recovered at least half of their loss, compared with 38% of those who received placebo. Patients with mid-frequency losses did well with or without steroids; those with profound losses did poorly regardless of treatment.

Unfortunately, Wilson et al's is the only randomized trial of corticosteroid therapy to date. A nonrandomized study by Moskowitz and colleagues showed benefits similar to those seen in the Wilson et al study (89% recovery with corticosteroids, 44% recovery without treatment).[39] Many, if not most, otologists offer oral corticosteroids to patients with ISSNHL who have no contraindications (eg, diabetes, active duodenal ulcer, tuberculosis). A typical regimen is prednisone, 1 mg/kg/day, for 7 to 10 days, with or without a taper. Useful oral antiviral drugs have been available in recent years but have yet to be found helpful for patients with ISSNHL.[40]

The vascular hypothesis is extremely popular in Europe, where treatments such as oral pentoxifylline and intravenous dextran (intended to reduce blood viscosity), apheresis (intended to remove low-density lipoprotein cholesterol from the blood), carbogen (a mixture of 10% carbon dioxide and 90% oxygen), and papaverine (intended to dilate blood vessels) are often used. None of these has been shown to be superior to placebo.[41–46]

When a cause for sudden SNHL is found (see Table 17–1), the case is not idiopathic, and the treatment should depend on the diagnosis. The management of persistent tinnitus and hearing loss, when recovery is incomplete, is nonspecific. Some patients choose to try hearing aids, but, for most, sympathetic counseling is the best approach.

REFERENCES

1. Byl F. Sudden hearing loss: eight years' experience and suggested prognostic table. Laryngoscope 1984; 94:647–66.

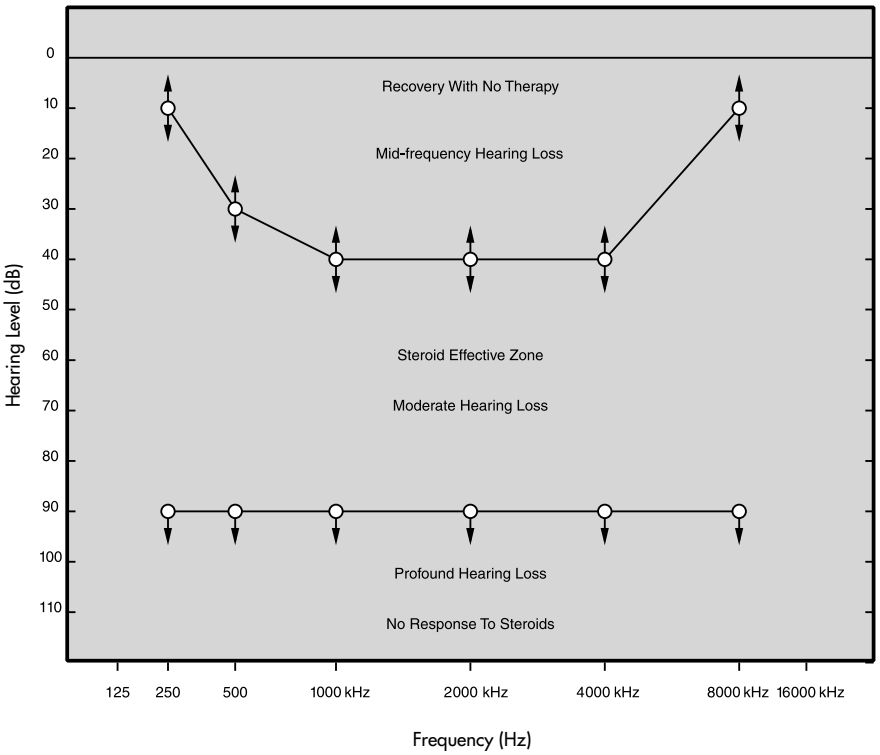


FIGURE 17–1. Audiometric categories predicting corticosteroid responsiveness (after Wilson et al[38]). Patients with mid-frequency losses, most of which are more severe than shown in this figure but always with better hearing in the low and high frequencies than in the mid-frequencies, always recovered without treatment (n = 14). Patients with profound losses (> 90 dB, all frequencies) did poorly with or without corticosteroids (n = 34). The remaining patients (n = 74) were in the "steroid-effective" zone.

2. Hughes GB, Freedman MA, Haberkamp TJ, Guay ME. Sudden sensorineural hearing loss. Otolaryngol Clin North Am 1996;29:393–405.
3. Megighian D, Bolzan M, Barion U, Nicolai P. Epidemiological considerations in sudden hearing loss: a study of 183 cases. Arch Otorhinolaryngol 1986; 243:250–3.
4. Karlberg M, Halmagyi GM, Buttner U, Yavor RA. Sudden unilateral hearing loss with simultaneous ipsilateral posterior semicircular canal benign paroxysmal positional vertigo. Arch Otolaryngol Head Neck Surg 2000;126:1024–9.
5. Mattox DE, Simmons FB. Natural history of sudden sensorineural hearing loss. Ann Otol Rhinol Laryngol 1977;86:463–80.
6. Dobie RA. Drug treatments for sensorineural hearing loss and tinnitus. In: Berlin CI, editor. Neurotransmission and hearing loss: basic science, diagnosis and management. San Diego: Singular; 1997. p. 147–59.
7. Trybus RJ, Karchmer MA, Kerstetter PP, Hicks W. The demographics of deafness resulting from maternal rubella. Am Ann Deaf 1980;125:977–84.
8. Davis LE, James CG, Fiber F, McLaren LC. Cytomegalovirus isolation from a human inner ear. Ann Otol Rhinol Laryngol 1979;88:424–6.
9. Keithley EM, Woolf NK, Harris JP. Development of morphological and physiological changes in the cochlea induced by cytomegalovirus. Laryngoscope 1989;99:409–14.
10. Stagno S, Reynolds DW, Amos CS, et al. Auditory and visual defects resulting from symptomatic and subclinical congenital cytomegaloviral and toxoplasma infections. Pediatrics 1977;59:669–78.
11. Hicks T, Fowler K, Richardson M, et al. Congenital cytomegalovirus infection and neonatal auditory screening. J Pediatr 1993;123:779–82.
12. Dahle AJ, McCollister FP. Audiological findings in children with neonatal herpes. Ear Hear 1988; 9:256–8.
13. Everberg G. Deafness following mumps. Acta Otolaryngol (Stockh) 1957;48:397–403.
14. Westmore GA, Pickard BH, Stern H. Isolation of mumps virus from the inner ear after sudden deafness. BMJ 1979;1:14–5.
15. Tanaka K, Fukuda S, Terayama Y, et al. Experimental mumps labyrinthitis in monkeys (*Macaca irus*)—immunohistochemical and ultrastructural studies. Auris Nasus Larynx 1988;15:89–96.
16. Tieri L, Masi R, Marsella P, Pinelli V. Sudden deafness in children. Int J Pediatr Otorhinolaryngol 1984;7:257–64.
17. Harbert F, Young IM. Audiologic findings in Ramsay Hunt syndrome. Arch Otolaryngol 1967;85:632–9.
18. Wackym PA. Molecular temporal bone pathology: II. Ramsay Hunt syndrome (herpes zoster oticus). Laryngoscope 1997;107:1165–75.
19. Davis LE, Rarey KE, McLaren LC. Clinical viral infections and temporal bone histologic studies of patients with AIDS. Otolaryngol Head Neck Surg 1995;113:695–701.
20. Schuknecht HF, Donovan ED. The pathology of idiopathic sudden sensorineural hearing loss. Arch Otorhinolaryngol 1986;243:1–15.
21. Veltri RW, Wilson WR, Sprinkle PM, et al. The implication of viruses in idiopathic sudden hearing loss: primary infection or reactivation of latent viruses? Otolaryngol Head Neck Surg 1981;89: 137–41.
22. Wilson WR. The relationship of the herpes virus family to sudden hearing loss: a prospective clinical study and literature review. Laryngoscope 1986; 96:870–7.
23. Schulz P, Arbusow V, Strupp M, et al. Highly variable distribution of HSV-1-specific DNA in human geniculate, vestibular and spiral ganglia. Neurosci Lett 1998;252:139–42.
24. Stokroos RJ, Albers FW, Schirm J. The etiology of idiopathic sudden sensorineural hearing loss. Experimental herpes simplex virus infection of the inner ear. Am J Otol 1998;19:447–52.
25. Pitkaranta A, Vasama J-P, Julkunen I. Sudden deafness and viral infections. Otorhinolaryngol Nova 1999;9:190–7.
26. Dobie RA. Medical-legal evaluation of hearing loss. 2nd ed. San Diego: Singular; 2001.
27. Wilson WR, Laird N, Moo-Young G, et al. The relationship of idiopathic sudden hearing loss to diabetes mellitus. Laryngoscope 1982;92:155–160.
28. Schuknecht HF, Igarashi M, Chasin WD. Inner ear hemorrhage in leukemia. A case report. Laryngoscope 1965;75:662–8.
29. Vakkalanka S, Ey E, Goldenberg RA. Inner ear hemorrhage and sudden sensorineural hearing loss. Am J Otol 2000;21:764–5.
30. Tavin ME, Rubin JS, Camacho FJ. Sudden sensorineural hearing loss in haemoglobin SC disease. J Laryngol Otol 1993;107:831–3.

31. Plasse HM, Mittleman M, Frost JO. Unilateral sudden hearing loss after open heart surgery: a detailed study of seven cases. Laryngoscope 1981;91:101–9.
32. Ness JA, Stakilwicz JA, Kaniff T, et al. Sensorineural hearing loss associated with aortocoronary bypass surgery: a prospective analysis. Laryngoscope 1993; 103:589–93.
33. Simmons FB. The double-membrane break syndrome in sudden hearing loss. Laryngoscope 1979; 89:59–66.
34. Sakashita T, Minowa Y, Hachikawa K, et al. Evoked otoacoustic emissions from ears with idiopathic sudden deafness. Acta Otolaryngol Suppl (Stockh) 1991;486:66–72.
35. Shaia FT, Sheehy JL. Sudden sensori-neural hearing impairment: a report of 1,220 cases. Laryngoscope 1976;86:389–98.
36. Mark AS, Seltzer S. Nelson-Drake J, et al. Labyrinthine enhancement on gadolinium-enhanced magnetic resonance imaging in sudden deafness and vertigo: correlation with audiologic and electronystagmographic studies. Ann Otol Rhinol Laryngol 1992;101:459–64.
37. Albers FWJ, Demerynck KMNP, Casselman JW. Three-dimensional magnetic resonance imaging of the inner ear in idiopathic sudden sensorineural hearing loss. ORL J Otorhinolaryngol Relat Spec 1994;56:1–4.
38. Wilson WR, Byl FM, Laird N. The efficacy of steroids in the treatment of idiopathic sudden hearing loss. A double-blind clinical study. Arch Otolaryngol 1980; 106:772–6.
39. Moskowitz D, Lee KJ, Smith HW. Steroid use in idiopathic sudden sensorineural hearing loss. Laryngoscope 1984;94:664–6.
40. Stokroos RJ, Albers FW, Tenvergert EM. Antiviral treatment of idiopathic sudden sensorineural hearing loss: a prospective, randomized, double-blind clinical trial. Acta Otolaryngol (Stockh) 1998;118: 488–95.
41. Fisch U. Management of sudden deafness. Otolaryngol Head Neck Surg 1983;91:3–8.
42. Edamatsu H, Hasegawa M, Oku T, et al. Treatment of sudden deafness: carbon dioxide and oxygen inhalation and steroids. Clin Otolaryngol 1985;10:69–72.
43. Kallinen J, Laurikainen E, Laippala P, Grenman R. Sudden deafness: a comparison of anticoagulant therapy and carbogen inhalation therapy. Ann Otol Rhinol Laryngol 1997;106:22–6.
44. Kronenberg J, Almagor M, Bendet E, Kushnir D. Vasoactive therapy versus placebo in the treatment of sudden hearing loss: a double-blind clinical study. Laryngoscope 1992;102:65–8.
45. Probst R, Tschopp K, Ludin E, et al. A randomized, double-blind, placebo-controlled study of dextran/ pentoxifylline medication in acute acoustic trauma and sudden hearing loss. Acta Otolaryngol (Stockh) 1992;112:435–43.
46. Suckfull M, Thiery J, Schorn K, et al. Clinical utility of LDL-apheresis in the treatment of sudden hearing loss: a prospective, randomized study. Acta Otolaryngol (Stockh) 1999;119:763–6.

CHAPTER 18

Perilymphatic Fistulae

Robert I. Kohut, MD

A perilymphatic fistula (PLF) is an abnormal passage through the bony labyrinthine capsule that is permeable to perilymph, causing abnormal, sometimes reversible, cochlear or vestibular function.

The author's attention to this disorder occurred with a patient encounter. This patient's severe vestibular symptoms had previously been incorrectly attributed to a central neurovascular event and later to psychological problems. The presence of a positive unilateral Hennebert's sign/symptom in this patient who had no other signs of syphilis (and was proven to be free of syphilis) prompted the conclusion that the findings were owing to a pathologic process in which the positive and negative air pressure applied to the external ear canal was transmitted to the inner ear. On exploration of the middle ear, fluid reaccumulation just anterior to the stapes was observed, and this site was grafted. Dramatic rapid improvement occurred in this patient.[1]

The purpose of this chapter is to make available to the reader a set of diagnostic criteria to be used during clinical or research activities concerning PLF that cause balance disorders, sensory hearing disorders, or both. To achieve this purpose, a summary of historical facets and research activities extending over 20 years related to this disorder will be reviewed. Following this review, discussion will concern a predictive masked experimental paradigm, which was used to test the diagnostic criteria. In these experiments, the temporal bone histologic diagnostic criteria were developed whereby clinical events were predicted based on the presence of specific histologic findings. Then, using these histologic diagnostic criteria, another experiment was designed in which the presence of the histologic diagnostic criteria for each donor temporal bone pair was predicted on the basis of clinical chart data recorded during life.

The treatment methods that will be described are the products of the experience of many otologists, and these are still in a dynamic phase of development.

The reader is encouraged to consider the thoughts shared in this chapter, to criticize them, and help develop, through his/her own clinical experience and investigative efforts, further knowledge concerning this disorder. Many in the recent past have considered PLF as a nonentity. However, the preponderance of clinical and investigative data has given most of these clinicians reason to pause and recognize the pathogenic process called PLF or perilymphic leaks. A very thorough review article was published by Friedland and Wackym.[2] Their conclusion that PLF is an uncommon disorder appears to be based on clinical reports, particularly the findings on endoscopic examination. The potential weakness of this type of examination is discussed elsewhere in this chapter and relates to the heat produced by the endoscope with possible evaporation of the minute quantity of perilymph present, thereby precluding fluid detection.

HISTORY OF PERILYMPHATIC FISTULAE

In 1883, Gellé observed, during manipulation (without dislodgment) of a patient's exposed stapes, that the patient suffered vertigo with each movement of the ossicle. He attributed this to hypermobility of the stapes.[3]

Hennebert, some 22 years later, observed that compression and rarefaction of the air in the external auditory canal in some patients, having intact tympanic membranes, caused small-amplitude, brief, nystagmic movement of the eyes (Hennebert's sign).[4] Some patients having this provoked nystagmus also experienced a sense of displacement in space: vertigo (Hennebert's symptom).[5] Many, but

not all, of Hennebert's patients had syphilis. Since syphilis was a popular medical topic at that time, a generalization was developed wherein a patient with an intact tympanic membrane and with the findings of a nystagmic response to a fistula test (compression and rarefaction of the air of the external canal) was considered to have syphilis.

A positive fistula test, before Hennebert, had been observed in patients with bone-eroding ear disease such as cholesteatoma, and these were documented surgically. Yet there was no evidence in Hennebert's patients of an opening in the bony labyrinthine capsule as seen in these bone-eroding processes.

Discovering no bone-eroding process and reflecting on Gellé's observation, it seemed reasonable to conclude that the symptoms were attributable to a hypermobile stapes. However, there were some who thought that the symptoms and/or signs caused by a positive fistula test in patients having an intact tympanic membrane were most like those of patients having erosive otic disorders in which there was an opening through the bony labyrinthine capsule joining the middle ear and the inner ear (labyrinthine fistula). Some of these individuals were prominent otologists such as Alexander, Leidler, and McKenzie.[1]

When no fistula was found at surgery in patients having an intact tympanic membrane and a positive fistula test, the surgeons were criticized as having performed needless surgery. Note that the surgery was performed without the availability of the operating microscope.

The Nobel Laureate Robert Bárány, in 1910, concluded that because of the negative surgical findings and the previous proposals of stapes hypermobility, the hypermobile stapes hypothesis seemed reasonable.[6] This seemed to be confirmed later in 1921, when Ruttin reported the histologic findings of the temporal bones from a patient having a Hennebert's sign: no abnormal labyrinthine fistulae were histologically identified.[7]

This thought, that a positive Hennebert's sign was diagnostic of syphilis, was not re-examined until the advent of stapes surgery. In 1967, PLF were reported as a postoperative complication of stapedectomy for otosclerosis. The main concern here was hearing loss.[8,9] One year later, Fee reported a patient having a PLF caused by nonsurgical trauma.[10] Over the next decade, this was followed by similar reports, which indicated that sudden or rapidly progressive hearing loss[11–15] in patients with or without vertigo, or even with vertigo alone, some having a positive fistula test (Hennebert's sign or symptom), could be caused by a PLF.[12–17]

Okano and associates demonstrated a crack-like formation about the oval window or connecting the round window niche and the posterior semicircular canal ampulla.[18,19] Although not clinically correlated, these openings were suggested as a form of opening between the middle ear and the inner ear, perhaps related to inner ear disease. This type of histologic opening had not been considered by Ruttin.[7]

In 1986, Kohut et al reported a relationship between patencies of the fissula ante fenestram (FAF) and/or patencies of the fissure connecting the round window niche/posterior semicircular canal ampulla (round window fissure [RWF]) and specific fistula symptoms and signs.[16] These findings were supported by subsequent studies wherein these patencies could be predicted as being present or absent histologically based on archival medical record data.[17]

It appeared that histologic changes (noninterrupted formations of the FAF and/or RWF) could be predicted on strict clinical data related to PLF. Clinical prediction of pathologic findings, of course, is the "acid test" for disorders such as appendicitis, Laënnec's cirrhosis, malignancies, and other disorders related to histologically defined criteria. Cracks of the labyrinthine capsule in other areas existed in some of the temporal bone specimens examined; however, these never seemed to be related to clinical signs or symptoms.[17]

Since specific clinical criteria for PLF were related to vertigo and/or hearing loss, challenging temporal bone histologic investigations included one by El Shazly and Linthicum in which patencies of the labyrinthine capsule were not found to be related to sudden sensory hearing loss.[20] In this study, the authors found that the presence of labyrinthine capsule patencies did not predict the presence of sudden sensory hearing loss.[20] This was not in disagreement, although, at first glance, it appears so, with the previous studies by Kohut et al and Hinojosa et al.[21,22] The presence of sudden sensory hearing loss in Kohut et al's study predicted patencies of the labyrinthine capsule, not vice versa.[21]

Again, El Shazly and Linthicum's study tested the hypothesis that patencies of the labyrinthine

capsule predicted sudden sensory hearing loss,[20] not sudden hearing losses predicting the presence of patencies. As such, the studies are not in conflict.

However, one small series of temporal bones from patients having had sudden sensory hearing loss during life lacked the use of the necessary diagnostic exclusionary criteria for PLF described below. In this series, no patencies of the FAF or RWF were found. Yet in two specimen pairs (the only pairs that met the diagnostic exclusionary criteria), there was histologic evidence of patencies of the central portions of the stapes footplates. This condition has been seen by the author during surgery for PLF and has been repaired with improvement of symptoms.[23] Current studies suggest that this histologic change is not a normal variant.[24,25]

In another temporal bone histopathologic study by the author and other colleagues, the prevalence of patencies of the FAF or RWF using 200 temporal bone pairs from consecutive autopsies, therefore representing a cross-section of society, was 24%, which is surprisingly high. Logic compels one to conclude that the prevalence of patients having PLF symptoms and signs are magnitudes smaller. Therefore, the mere histologic presence of these specific patencies seems not to predict the presence of symptoms of PLF during life. On the other hand, when hearing or balance symptoms had been present during life, the presence of these labyrinthine capsule patencies predicts the specific set of symptoms and signs related to PLF. One explanation to this seeming dilemma seems plausible, that is, there is another variable, perhaps distinguishable only with special histologic techniques. The author and his colleagues are engaged in an investigation testing the hypothesis that "some patencies are permeable to fluid, some are not, and only those permeable to fluid are related to clinical manifestations." So far, in preliminary studies, differences in proteoglycans (extracellular mucopolysaccharides bound to protein chains in covalent complexes) appear to be so related.[26] Definitive studies await completion.

COMMENTS REGARDING OVERALL CLINICAL PERSPECTIVES

The diagnostic criteria for PLF are simple but do require specific exclusionary criteria: a PLF cannot be accurately diagnosed if any of these exclusionary disorder types are present: inflammation, granuloma, neoplasia, or anatomically related neurologic disorders with symptoms and signs similar to PLF. These disorders can mimic symptoms and signs otherwise attributable to PFL. Symptoms and signs of PLF are related to the vestibular system and cochlear function separately, even though they can be present in combination.

As has been reviewed, the temporal bone histopathologic/clinical studies tested specific clinical criteria. The studies indicate that the probability and accuracy of these criteria are $p < .001$, sensitivity 59%, and specificity 91%. It appears reasonable to present the criteria used in these studies to the reader. A sensitivity of 59% suggests underdiagnosis; therefore, a heightened alertness to the possibility of a PLF appears reasonable. Additional diagnostic refinements are required to identify clinically the remaining 41% (100 – 59 = 41) of patients having PLF who are now presumably misdiagnosed as having another disorder. These refining new diagnostic criteria remain the responsibility of the readers or others to identify.

DIAGNOSTIC EXCLUSIONARY CRITERIA

For all clinical considerations regarding PLF, initially, there must be the exclusion of inflammation, granuloma, neoplasia, or anatomically related neurologic disorders. These are easily accomplished with a sedimentation rate, a microhemoagglutination assay for antibodies to *Treponema pallidum* for syphilis, magnetic resonance imaging for tumor, and a neurologic evaluation pertinent to this disorder as regularly accomplished by otologists. All of the above occur after arrival at the clinical suspicion of PLF, which is developed following a detailed clinical history and otoneurologic evaluation. The clinical elements that allow a heightened suspicion of PLF are very specifically related to either hearing or balance or both, as described below.

Hearing

For hearing, each patient with sudden sensory hearing loss (hours to days) or rapidly progressive sensory hearing loss (weeks to months) is suspected of having a PLF.

After the hearing loss occurs, rapid evaluation of these patients is required (no more than 3 days and preferably within several hours).

Balance

For disorders of balance or concomitant physical findings in patients with hearing loss but perhaps unrecognized vestibular symptoms, the aspects of the clinical examination are very deliberate. Without recognition of the subtleties, the diagnosis may be missed. The three diagnostic criteria include (1) constant dysequilibrium, however mild; (2) a positive Hennebert's sign or symptom; and (3) positional nystagmus or postural vertigo. The common tests for these elements are inadequate and will lead the clinician to false conclusions. Therefore, sensitive tests are warranted and easily accomplished. Be aware that all three types of findings can sometimes not be present during a single examination. This requires that in some patients, repeat testing is necessary to determine the presence or absence of any of the sought finding(s).

For dysequilibrium, observe the patient walking with eyes closed. Often the patient will be surprised that there is dysequilibrium, which was never before recognized; eyes are usually open during ambulation, adding vision to aid stability. On rare occasion, the patient must "walk on tiptoes" (the balls of the feet) with eyes closed to allow the physician to observe and the patient to recognize the presence of dysequilibrium.

For Hennebert's sign or symptom, the patient should be standing so that the vestibulospinal reflex appears most sensitive. The patient should be standing with eyes closed and feet together (as close to this as possible if unstable). The fistula (Hennebert's) test is then performed using a Bruening otoscope, usually with a flexible tympanogram earpiece fitted over the speculum earpiece to allow a convenient ear canal seal. Other clinical otoscopes do not provide an adequate air seal. There should be enough pressure to cause the ear canal skin to move or blanch with pressure and rarefaction. Be prepared to catch or support the patient quickly should he/she lose balance. A positive vestibulospinal response is the observation of body sway (sign) or the patient's sensation of sway or loss of balance (symptom). Interestingly, some of these patients also exhibit a Tullio phenomenon (vertigo or nystagmus in response to high-intensity sound).

For eye movement observation with fistula tests, naked eye observation will most often be nonrevealing. Frenzel lenses may allow for the observation of one, two, or three nystagmic beats but often will give a false-negative result. Electrode eye movement recordings result in false positives. The most sensitive test for a Hennebert's sign (a few small-amplitude nystagmic beats) requires the use of infrared nystagmography with video recording and with time-locked signals for positive and negative ear canal air pressure. In this way, with infrared observation in darkness, visual fixation is avoided, which can quell the diagnostic nystagmic beats. Be aware that both the vestibulospinal and vestibulo-ocular responses fatigue quickly, so serial repeat testing is usually not confirmatory and can give one a false impression of a negative result.

For positional nystagmus or vertigo, the usual Hallpike maneuvers are employed. However, the usual methods of no lenses or Frenzel lenses are often inadequate. Therefore, infrared eye movement recordings are often necessary to give objective evidence to the patient's report of postural vertigo. While seeking the above information, the results of a posturography test are often revealing. Any evidence of dysequilibrium in patients complaining of postural vertigo heightens one's suspicion of the presence of a PLF.

Again, be aware that repeat testing at the initial or a subsequent clinical visit is often required to find the positive results that allow accurate diagnosis.

The Possibility of a Perilymphatic Fistula: Elements That Heighten One's Awareness; Absence Not an Excluding Factor For either hearing loss (sudden sensory or rapidly progressive) or vestibular symptoms, there is often a precipitating factor related to barometric pressure change, sometimes severe, as with scuba diving, or as mild as a travel venture with ascent or descent of a mountain or even an airplane flight. The patient says, "When I got off the plane, I noticed I had lost some of my sense of balance" or "The hearing loss occurred after diving." The absence of historical events of trauma does not exclude the presence of PLF in that the event can be as mild as a vigorous sneeze. It appears that only the absence of the array of clinical findings on repeat testing is exclusionary.

CLINICAL MANAGEMENT

Diagnostic Aids

The use of fluorescin or other perilymphatic indicators, thus far, appears to be not regularly reliable.

When a positive result is present, it is supportive of the diagnosis, but a negative result means nothing. Therefore, specific surgical observations remain our best form of confirmation of a PLF diagnosis. Diagnostic variables based on outcome analysis of treatment should be judged and evaluated with caution, considering the need for further refining of diagnostic criteria related to this disorder, which are not yet defined but are anticipated. There are also possible unrecognized treatment variables that may occur coincidentally with the intended treatment changes for other disorders when the correct diagnosis of PLF has been missed.

There is a question regarding the value of endoscopic middle ear examination for evidence of perilymph.[27] It appears that the endoscope becomes hot. Not only is this important to patient comfort, but perhaps even more significant in terms of diagnostic accuracy. The volume of perilymph present at the specific sites of accumulation (FAF or RWF) is often just a few microliters, easily evaporated by modest amounts of heat. In fact, even at surgery, a touch of the minute amount of fluid with a sharp-pointed instrument may be necessary to confirm its presence by the resulting light-reflection change.

Medical Management

The initial treatment for PLF usually should be medical management except for sudden sensory hearing loss. Sudden hearing loss requires more intense and immediate restrictions. The general elements of patient instructions are quite simple but specific and and relate to hydrodynamics:

1. Keep the head above the level of the heart.
2. Avoid lifting anything over 10 pounds.
3. Have someone elevate the head of the bed 4 inches by putting blocks under the headboard bedposts.

Equivalents of these should be recognized and avoided (eg, straining at stool, vigorous sneezing [should open mouth], and hard nose blowing).

These elements of management allow vestibular symptom improvement in about 40% of patients; for hearing loss, it appears to be significantly less. These modifications should be followed for 6 weeks. If there is no significant improvement, surgery should be considered. Sudden hearing loss is a special consideration. For sudden sensory hearing loss, 5 days of inpatient bed rest with the head of the bed elevated and bathroom privileges with help are required for patients who have been seen shortly after the onset (weeks). Audiograms (bedside) are performed daily. If adequate improvement occurs, the patient is given the above restriction directions and is discharged from the hospital. If there is no improvement, surgery is performed on the sixth day. If there is worsening of hearing, surgery is performed at the time of worsening.

Surgery

Whenever possible, surgery is performed under local anesthesia with minimal or no sedation to avoid postanesthetic straining and masking of surgically related symptoms. Exposure of the middle ear is accomplished using a standard tympanomeatal flap designed to give generous exposure of the oval window and round window niche. Care should be taken to provide visualization of the area of the FAF just anterior to the stapes. Also, although rarely involved, the area of the fossula post fenestram (FPF) posterior to the stapes should be considered. The RWF is located on the floor of the round window niche extending anteriorly, often initially obscured by mucosal bands (not to be confused with the round window membrane).

The surgical manipulations should be done in such a way as to avoid blood entering the middle ear. If such should occur, careful suctioning should be used to remove it. Use as small a suction tip as possible to avoid the "hair dryer" evaporation qualities of air movement, which have the potential of evaporating the small quantities of perilymph present.

After exposure, carefully examine the area just anterior to the stapes footplate for crystal clear fluid collecting in minute quantities. Often there is a roughened appearance or a minute cleft-like formation of the bony surface. With a 26-gauge suction tip, aspirate any fluid present in this area. Repeated reaccumulation of fluid often but not always takes several minutes and is diagnostic of the presence of perilymph.[28] Not infrequently, the presence of fluid is not immediately apparent because its total volume is only a few microliters (total inner ear perilymph volume is 76 μL). In these instances, touching the area(s) in question with a sharp needlepoint instrument will cause the fluid to flash a sparkle as its reflective surface is altered.

Not infrequently, the perilymph is present only under the mucosa. With careful surgical (nonbloody) mucosal reflection and subsequent suction drying of the area in question, diagnostic reaccumulation of fluid is observed unrelated to the distant reflected mucosal edges. If laser is used for mucosal obliteration, removal of any char with a fine pick is necessary prior to observation for fluid as this can inhibit fluid flow.

After diagnostic reaccumulation perilymph has been observed, sampling of the fluid can be done. The author prefers to use glass micropipettes with a tip size of about 10 μm. Often preuse tip breakage is necessary in that the usual micropipette has a tip size of 3 to 5 μm, too small to allow the fluid to rise in the pipette by capillary action. Often the fluid volume is so small that only the frustum of the pipette will fill, sometimes just a portion of the frustum. Other times, a whole pipette full can be obtained, about 10 μL. Surprisingly, there does not seem to be a relationship between fluid volume and symptom or sign intensity.

After the surgical diagnosis (reaccumulation of fluid) has been made, graft site preparation is necessary and is meticulously performed. The mucosa is reflected or desiccated by laser away from the site of fluid origin (FAF or RWF). Care must be taken at the round window niche in that the singular nerve may be herniated and exposed at the middle ear surface.[19] Any tissue debris, even the slightest amount, must be removed and the mucosal edges reflected far enough from the site to allow graft pieces to be placed without a risk of overlaying any mucosal element. The mucosal edges must be identified as being free from the graft elements after graft replacement.

Graft Materials The most effective graft material appears to be the loose areolar fibrous tissue overlying the temporalis muscle fascia.

The graft material is easily obtained through a 1 cm postauricular incision. Rarely will any preoperative postauricular shaving be necessary. A piece of graft 1 cm in diameter is more than enough. Once obtained, the graft is minced in very fine pieces, almost pinpoint in size. These pieces are cut on a hard surface, usually metal.

The graft application to the recipient site is performed using slightly blunted needle-tip ear picks. These picks are loaded with individual pieces of graft by the person assisting the otologic surgeon. The first is placed in the area of leak source, with subsequent pieces placed over and around the initial piece to help hold it in position. The opening producing the fluid passage is only 50 μm in maximal diameter; therefore, the initial graft piece must be very small and specifically placed. (It appears histologically that there is fibrous tissue ingrowth of this graft material and/or fibrous elements stimulated by this graft material into the FAF. Interestingly, the RWF appears to seal differently, with a dense collagen formation.[23])

Alternative Surgical Method There are some patients whose general condition may not allow the surgery described above. Although the author has no experience with the method, it appears simple and has been reported as effective for poststapedectomy PLFs (JB Causse, personal communication, March 2001). Simply, autologous blood is injected into the middle ear through the tympanic membrane. Presumably, a fibrous seal results owing to the sterile inflammatory response to the blood.

Additional Surgical Considerations, Including Recurrence For vestibular disorders caused by PLF, an operation can be considered successful if after 6 weeks the patient is significantly improved or symptom free. Symptom fluctuation before 6 weeks is not unusual. Any evidence of the return or worsening of symptoms after a 6-week period of symptom improvement should be considered a recurrence. Recurrences will occur in about 25% of patients. In this instance, medical management should be followed as described above before considering reoperation, assuming hearing stability. If surgery is warranted, the procedures are as described above with care to re-examine all areas of risk (FAF, FPF, and RWF), even those seemingly not involved at the time of the previous operation.

For hearing, although rapid improvement has been seen, it is not the rule. Stability is sought. Any sudden worsening prompts measures considered for sudden hearing loss (above).

Special Consideration Concerning Endolymphatic Hydrops In histologic studies, endolymphatic hydrops has been seen at times to be present in specimens from patients meeting the clinical diagnostic criteria and diagnosed as having PLF dur-

ing life and having the predicted FAF or RWF patencies histologically.[17]

Similarly, clinical studies without histologic confirmation of patients with PLFs have indicated that some patients exhibit clinical findings consistent with the presence of endolymphatic hydrops. In these patients, an endolymphatic shunt appeared to be helpful to those who used an outcome analysis experimental study design.[29] The only difference between the histologic studies and the clinical study is the number or percentage of patients so afflicted. There appeared to be a greater frequency in the clinical study. This raises the additional question of possible long-term tissue changes in cases of endolymphatic hydrops—a possible time-related resolution of the otherwise histologically identifiable tissue distention of endolymphatic hydrops. The answer to this question will await appropriate clinical/histologic studies.

Postoperative Care After surgery for a PLF, the author has had patients remain in the hospital for 5 days, with the head of the bed elevated 30 degrees, and has only allowed bathroom privileges with help. The personal help is required because of the risk of falls. Immediate postoperative exacerbation of vestibular symptoms is not unusual. The 5 days of bed rest are ordered to avoid, as much as possible, graft dislodgment. The author has the experience of reoperating on the patients of other surgeons who have experienced recurrence. The only apparent change in care was the 5-day hospitalization followed by 6 weeks of house confinement, and there were the usual successes. All patients are reinstructed regarding the physical restrictions required for medical management (above).

Unusual Locations of Perilymphatic Leak On a few occasions, the author has observed the source of perilymph leakage in the central portion of the stapes footplate. In these instances, the surgical repair was performed as described above. Ballottement of the mucosa of the stapes footplate may be necessary to give evidence of the trapped mucosal-covered fluid leaking from the stapes footplate defect. Again, the defect may be minute, and recurrent fluid emission may be the only evidence of a PLF.

Anticipated Postoperative Course Not infrequently, patients become dramatically improved or even free of their preoperative vestibular symptoms. Unfortunately, this is not the rule for sensory hearing loss. Usually, in this case, improvement is slow. However, there are rare cases in which hearing is improved dramatically in but a few days or even as short as 12 hours. Just as infrequently, the author has observed cases in which socially inadequate hearing has improved to socially functional hearing over the course of several years. The physiologic mechanisms and dynamics of the inner ear causing hearing loss or vestibular symptoms remain unexplained. However, the dramatic sudden improvement in some suggests a bioelectric or biochemical process and not a biomechanical process.

Complications Few complications are experienced by patients undergoing medical or surgical PLF repair. Infection has been experienced by two of the author's patients, both of whom had allowed their ears to get wet before healing occurred. They did not lose hearing or vestibular function. The infection remained restricted to the external canal/middle ear.

The hoped for improvement in hearing sometimes does not occur, and progression of the hearing loss continues. When the preoperative symptoms were vestibular only, a mild sensory hearing loss has occurred in some patients postoperatively.

In all instances, the author has advised the patients preoperatively that any complication is possible (including death), particularly the risk to any structure in the surgical field, specifying these structures and their function.

THOUGHTS REGARDING FUTURE NEEDED STUDIES

Knowledge regarding the pathogenic mechanisms of PLFs remains conjecture based on an extension of pilot studies concerning intracellular proteoglycans related to fluid. Definitive studies are needed, or the disorder will remain a medical mystery.

In that the disease mechanism related to PLF appears to be bioelectric, there should be a search for measurement methods that are applicable in vivo by which these phenomena can be identified, perhaps substituting for the sought perilymph indicators. The implications of this type of measurement for other ear disease may indeed be more significant than its singular use for PLF.

The current methods of treatment of PLF, developed clinically using the best techniques avail-

able, remain crude: patching a minute hole and/or altering endolymph dynamics by manipulation of the endolymphatic sac. Even in these two instances, the physiologic consequences remain an enigma. It appears, however, that they help patients improve in their health status. Is this "dumb luck?"

Regarding PLF, as with all other disease, "what" remains the junior question, and "how" and "why" the cascade of events are triggered are of paramount importance. With about 25% of the population having these temporal bone pathways, patencies, but apparently only a fraction of these having fluid permeable patencies, the questions "how" and "why" loom as all important.[26] Proteoglycan dynamics may provide insight.

REFERENCES

1. Kohut RI, Waldorf RA, Haenel JL, Thompson JN. Minute perilymph fistulas: vertigo and Hennebert's sign without hearing loss. Ann Otol Rhinol Laryngol 1979;88:153–9.
2. Friedland DR, Wackym PA. A critical appraisal of spontaneous perilymphatic fistulas of the inner. Am J Otol 1999;20:261–76.
3. Gellé R. Etude clinique du vertige de Menier dans ses rapports avec les lesions des fenetres ovales et rondes. Arch Neurol (Paris) 1883;4:273.
4. Hennebert C. Labyrinthite double, reflexe moteur oto-oculaire. Clin (Brux) 1905;19:214–5.
5. Ramadier J. L'epreuve pneumatique du vestibul dans l'heredo-syphilis auriculaire, le sign de las fistule sans fistule (signe d'Hennerbert). Presse Med 1921; 19:624.
6. Bárány R. Fall von labyrinthlues. Monatssch Ohrenheik 1910;11:40.
7. Ruttin E. Uber Fiselsymptom ohne Fisel. Verh Ges Deutsch Hais-Hasen-U Ohrenarzte (Leipz) 1921; 8:286–97.
8. Harrison WH, Shambaugh GE, Derlacki EL, Clemis JD. Perilymph fistula in stapes surgery. Laryngoscope 1967;77:836–49.
9. House HP. The fistula problem in otosclerotic surgery. Laryngoscope 1967;77:1410–26.
10. Fee FA. Traumatic perilymphatic fistulas. Arch Otolaryngol 1968;88:477–80.
11. Stroud MH, Calcaterra TC. Spontaneous perilymph fistulas. Laryngoscope 1970;80:479–87.
12. Goodhill V, Brockman SJ, Harris I, Hantz O. Sudden deafness and labyrinthine window ruptures. Audiovestibular observations. Ann Otol Rhinol Laryngol 1973;82:2–12.
13. Facer GW, Farrell KH, Cody DT. Spontaneous perilymph fistula. A medical emergency. Mayo Clin Proc 1973;48:203–6.
14. Healy GB, Strong MS, Sampogna D. Ataxia, vertigo, and hearing loss. A result of rupture of the inner ear window. Arch Otolaryngol 1974;100:130–5.
15. Healy GB, Friedman JM, Strong MS. Vestibular and auditory findings of perilymphatic fistula: a review of 40 cases. Trans Am Acad Ophthalmol Otolaryngol 1976;82:44–9.
16. Kohut RI, Hinojosa R, Budetti JA. Perilymphatic fistula: a histopathologic study. Ann Otol Rhinol Laryngol 1986;95:466–71.
17. Kohut RI, Hinojosa R, Ryu JH. Perilymphatic fistulae: a single-blind clinical histopathologic study. Adv Otorhinolaryngol 1988;42:148–52.
18. Okano Y, Myers EN, Dickson DR. Microfissure between the round window niche and posterior canal ampulla. Ann Otol Rhinol Laryngol 1977;86:49–57.
19. Okano Y, Myers EN. Herniation of the singular nerve into the round window niche. Arch Otolaryngol 1976;102:478–81.
20. El Shazly MA, Linthicum FH. Microfissures of the temporal bones: do they have any clinical significance? Am J Otol 1991;12:169–71.
21. Kohut RI, Hinojosa R, Ryu JH. Sudden-onset hearing loss in 11 consecutive cases: a temporal bone histopathologic study with identification of perilymphatic fistulae. Trans Am Otol Soc 1989:81–8.
22. Hinojosa R, Kohut RI, Lee JT, Ryu JH. Sudden hearing loss due to perilymphatic fistulae II. A quantitative temporal bone histopathologic study. Trans Am Otol Soc 1990:121–7.
23. Kohut RI, Hinojosa R, Ryu JH. [In preparation]
24. Kohut RI, Ryu JH. [In preparation]
25. Kohut RI, Hinojosa R, Ryu JH. The histological characteristics of the core of the fistula ante fenestram. Acta Otolaryngol Suppl (Stockh) 1991;48:158–62.
26. Kohut RI, Hinojosa R, Ryu JH. [In preparation]
27. Bottrill I, Perrault DR, Poe D. In vitro and in vivo determination of the thermal effect of middle ear endoscopy. Laryngoscope 1996;106:213–6.
28. Selzer S, McCabe BF. Perilymph fistula: the Iowa experience. Laryngoscope 1986;96:37–49.
29. Fitzgerald DC. Perilymphatic fistulas and Meniere's disease. Clinical series and literature review. Ann Otol Rhinol Laryngol 2001;110:430–6.

CHAPTER 19

Autoimmune Inner Ear Disease

Jeffrey P. Harris, MD, PhD, Elizabeth M. Keithley, PhD

BASIC SCIENCE

The idea that immunity might participate in the etiology of inner ear disease has only recently become accepted. Using the brain as a model, the inner ear was initially viewed as an immunologically "privileged" site, separated from cellular and humoral immunity by a blood-labyrinthine barrier. Apart from inner ear inflammation associated with viral or bacterial labyrinthitis, it seemed inconceivable that the immune system could operate within the bony labyrinth. However, experiments beginning in the 1980s demonstrated that antigen introduced into the inner ear of naive or systemically primed animals resulted in a brisk systemic immune response as well as a local one.[1–4] Moreover, antigen introduced into the inner ear of a systemically immunized animal resulted in hearing loss and even profound deafness owing to a vigorous secondary immune response[2,5,6] (Figure 19–1). The endolymphatic sac also becomes inflamed (Figure 19–2). The magnitude of these responses was dependent on an intact endolymphatic sac since ablation of the sac or even blockage of the endolymphatic duct reduced inner ear immune responses.[7,8] Experiments then demonstrated that the inner ear, although protected by systemic immunity, was damaged by bystander injury when cell-mediated immunity became involved. For example, systemic immunity can protect the inner ear from viral infection through circulating antiviral antibodies.[9,10] If, however, there is a cellular response to viral inoculation of a naive animal, there is hearing loss and cochlear damage. This damage can be reduced by systemic immunosuppression.[11] The route of entry of the inflammatory cells into the inner ear following antigen or viral challenge of the inner ear appears predominantly to be via the spiral modiolar vein (Figure 19–3). During the inflammatory response, this vein takes on characteristics of an activated venule and expresses intercellular adhesion molecule-1 (ICAM-1) on the endothelial cell surface that facilitates the passage of circulating immunocompetent cells into the scala tympani.[12–15] Once within the inner ear, these cells divide,[15] release inflammatory mediators, and set in motion events that lead to cellular proliferation and eventually osteoneogenesis[6] (Figure 19–4). Shortly after entering the cochlea, hearing loss occurs.[16] Endolymphatic hydrops often accompanies these end-stage reactions (see Figure 19–4). Of note, after the single inoculation of antigen, cells may continue to be stimulated, undergoing cell division for up to 6 weeks.[17]

CLINICAL STUDIES

In 1979, McCabe first brought attention to a possible discrete clinical entity when he presented a series of patients with bilateral, progressive hearing loss showing improvement following treatment with corticosteroids.[18] Although others had over the years described patients with ear-related illnesses associated with systemic immune disorders,[19–21] none had collected such a large series or speculated that these patients might have an organ-specific illness. Since that time, autoimmunity has been proposed as a cause for other inner ear disorders, including Meniere's disease,[22] sudden sensorineural hearing loss (SNHL),[23] and acute vertigo.[24] A relationship between inner ear disorders and systemic autoimmune disease has also been documented. A number of systemic autoimmune disorders such as polyarteritis nodosa, systemic lupus erythematosus (SLE), relapsing polychondritis, ulcerative colitis, and Wegener's granulomatosis include auditory and vestibular symptoms.[25] Moreover, some patients with suspected autoimmune inner ear disease (AIED) have either presented with or later developed systemic autoimmune diseases.[26,27]

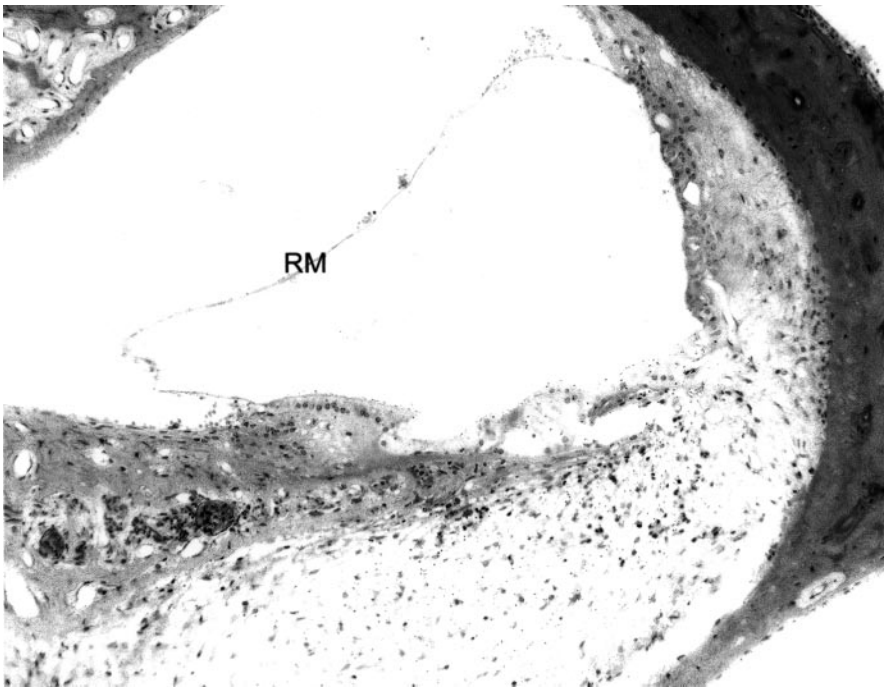


FIGURE 19–1. Photomicrograph of the middle turn of an inflamed guinea pig cochlea (× 20 original magnification). The scala tympani contains inflammatory cells and fibrotic tissue, and the position of Reissner's membrane (RM) suggests endolymphatic hydrops. Toluidine blue staining of plastic embedded tissue.

Unlike other organs and tissues, the inner ear is not amenable for biopsy for the express intent of investigating the underlying immunopathogenesis of purported autoimmune disorders affecting it. What little histopathology has been published from patients with suspected AIED shows fibrosis and/or bone deposition in the labyrinth, consistent with the late sequelae of inflammation.[28–31] Specific immune reactivity against inner ear antigens is often detected in patients with suspected AIED, but the results vary. Lymphocyte migration assays using inner ear tissue as a target have been disappointing, providing at best low stimulation indexes.[32] More promising results have been obtained with Western blotting. Significantly more patients with suspected AIED show reactivity against a 68 kD antigen (Figure 19–5) than do matched normal-hearing or rheumatic controls.[27,33] Autoimmune inner ear disease has also been associated with reactivity against antigens of different molecular weights, especially 45 to 50 kD, 30 kD, and 20 kD.[34] Although reactivity of patient sera against tissue sections of the inner ear has produced reproducible inner ear labeling,[35] it does not appear to be as practical or as useful as immunoblotting.

A number of animal studies have supported an autoimmune cause for some types of SNHL. Immunization of animals with crude extracts of inner ear tissues results in hearing loss in approximately one-third of subjects.[36] Circulating monoclonal antibodies against inner ear tissues also produce hearing loss.[37] Animal models of systemic autoimmune disease, such as the MRL-Fas^{lpr} mouse model of SLE, display hearing loss.[38]

ETIOLOGY

Because of the relative rarity of this condition and its recent recognition, as noted above, very few temporal bones with a diagnosis of AIED have been evaluated. More studies have assessed the inner ears of patients with systemic autoimmune disorders. Recently, Sone et al assessed 14 temporal bones from 7 individuals with SLE.[39] The duration of disease and ages varied widely. The most consistent findings were hair cell and spiral ganglion cell loss. However, unusual accretions were observed in the stria vascularis of 6 of 14 temporal bones.

Animal models have been valuable adjuncts in the study of AIED since the antigen and immunization history can be rigorously controlled, and histopathology is routinely available. Initial studies using immunization of guinea pigs with bovine inner ear extracts resulted in the development of

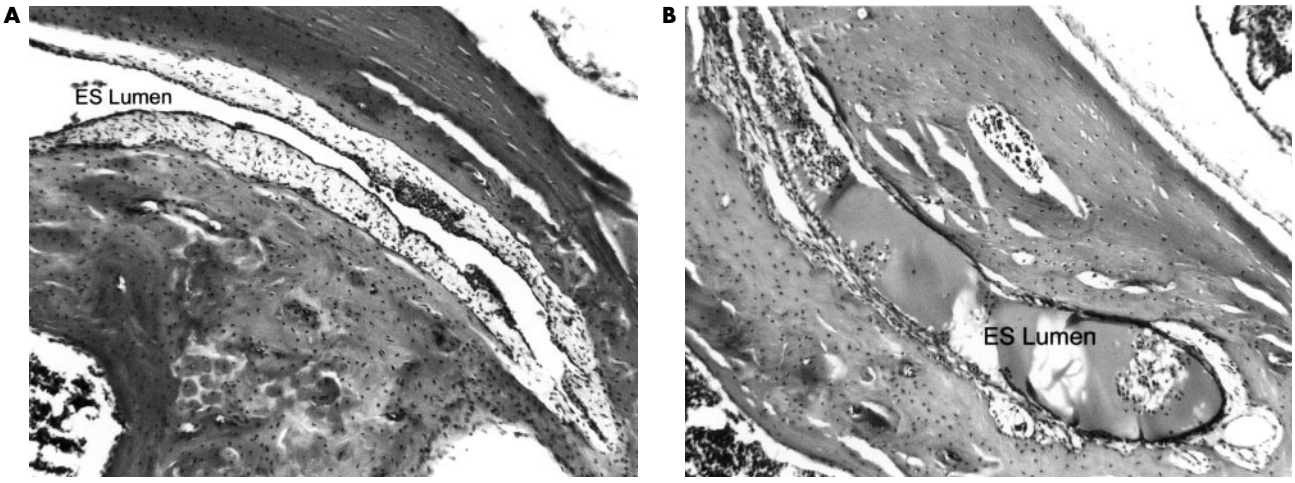


FIGURE 19–2. A, Photomicrograph of a normal endolymphatic sac (ES) in the guinea pig inner ear. Hematoxylin and eosin–stained paraffin section (× 20 original magnification). *B,* Photomicrograph of an ES from a guinea pig with cochlear inflammation following cochlear challenge with a foreign protein to which the animal was sensitized (× 20 original magnification).

hearing loss and mild inflammatory changes in the inner ears of a subset of animals.[36] More recent work has confirmed these findings and indicated that the hearing losses induced by this procedure tend to be modest. Bouman et al found that immunization of animals with swine inner ear extracts produced modest declines in compound action potentials recorded from the guinea pigs 2 and 6 weeks after immunization but no changes in the cochlear microphonic.[34] This suggests that the events responsible for hearing loss occurred at the level of the inner hair cell and/or spiral ganglion neuron rather than at the level of the outer hair cell. Hearing losses were associated with increased Western blot reactivity to 68 kD and other antigens.[34]

Immunization with specific proteins has also resulted in hearing loss. Based on the observation that myelin protein P0 was associated with immunoreactivity against a 30 kD inner ear protein in patients with AIED,[40] Matsuoka et al immunized mice with purified bovine P0.[41] They observed an approximately 10 dB hearing loss and a monocellular infiltrate in the eighth nerve within the cochlear modiolus. Experimental autoimmune encephalomyelitis can be induced by immunization with the neuronal S-100β calcium-binding protein and by passive transfer of T cells sensitized to this antigen. Based on this earlier finding, Gloddek et al found that passive transfer of S-100β-reactive T cells produced a 10 dB hearing loss in rats as well as a cellular infiltrate into the perilymph.[42,43] A monoclonal antibody raised against cochlear tissues of approximately 68 kD that specifically reacts with supporting cells in the organ of Corti has been shown to produce high-frequency hearing loss in mice carrying the hybridoma.[37] More recently, Nair et al infused this antibody into the cochlear perilymph using an osmotic minipump.[44] After 13 days, an approximately 20 dB hearing loss developed, associated with minor losses of hair cells.

To explore the origin of lymphocytes in the region of the endolymphatic sac and the reaction of T cells to self-antigens in the inner ear, Iwai et al used a model of graft-versus-host disease.[45] T cells from C57BL/6 mice injected into the systemic circulation of BALB/c mice infiltrated and proliferated in the perisaccular region surrounding the endolymphatic sac but not into other regions of the inner ear. These findings confirm the role of the endolymphatic sac region in mediating immunity in the inner ear, as well as the communication of the normal sac with circulating lymphocytes. This provides an additional foundation for autoimmunity as an etiology in disorders involving the sac, such as Meniere's disease.

Animal models have also been used to study the relationship between systemic autoimmune disease and the inner ear. The MRL-Fas^{lpr} mouse is used as a model of SLE owing to the accumulation of

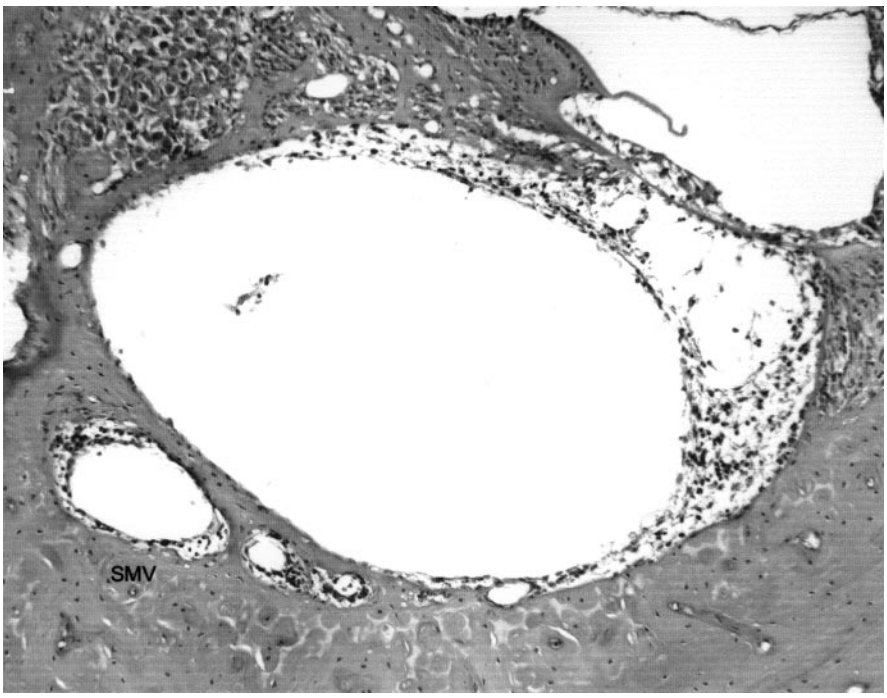


FIGURE 19–3. Photomicrograph of the middle turn of an inflamed guinea pig cochlea. Extravasated immunocompetent cells can be seen within the bony channel surrounding the spiral modiolar vein (SMV) and its tributary on their way to the scala tympani.

autoreactive T cells normally eliminated by Fas-mediated apoptosis. This model also displays progressive hearing loss. Ruckenstein et al found that the most striking inner ear pathology in this model was observed in the stria vascularis,[38] with progressive, hydropic degeneration of intermediate cells, consistent with the strial pathology observed in human SLE temporal bones as described above.[39] In addition, Ruckenstein and Hu observed the deposition of both complement-fixing and non–complement-fixing antibodies in the stria vascularis and, to a lesser extent, in other structures.[46] All antibodies were bound to the capillary walls and were not associated with signs of inflammation. The same group found that systemic treatment with dexamethasone suppressed antibody deposition within the stria and other structures of the inner ear.[38] However, the treatment failed to suppress strial degeneration and hearing loss, suggesting that perhaps the hearing loss seen in these animals had a genetic basis. In contrast to this result, Wobig et al found that systemic prednisolone treatment protected hearing in MRL-Fas^{lpr} mice.[47] The Palmerston-North mouse is also employed as a model of SLE with hearing loss. These animals develop abnormal mineralization of connective tissue in the region of the eighth nerve root within the modiolus. However, there is no deposition of antibody in or cellular infiltration of the cochlea in this strain.[48]

As better imaging becomes available, the localization of inflammatory processes to specific regions of the inner ear should become possible. Having this information will certainly improve our diagnostic capability as well as our knowledge of the basic pathogenesis of this disorder.

DIAGNOSIS

Diagnosis of AIED is still problematic. There is no universally accepted set of diagnostic criteria or diagnostic test for a condition that appears to have several independent etiologies. In general, in all cases of idiopathic, rapidly progressive, bilateral hearing loss, AIED should be suspected. However, there is no doubt that involvement of the second ear may occur months or even years after presentation of symptoms in the first ear. Hearing loss may be manifested as either diminished hearing acuity, decreased discrimination, or both and may involve significant fluctuations over time. In bilateral Meniere's disease, with its triad of vestibular dysfunction, low-tone, fluctuant hearing loss, and tinnitus, AIED should also be suspected, especially when the second ear becomes involved within a short period of time of

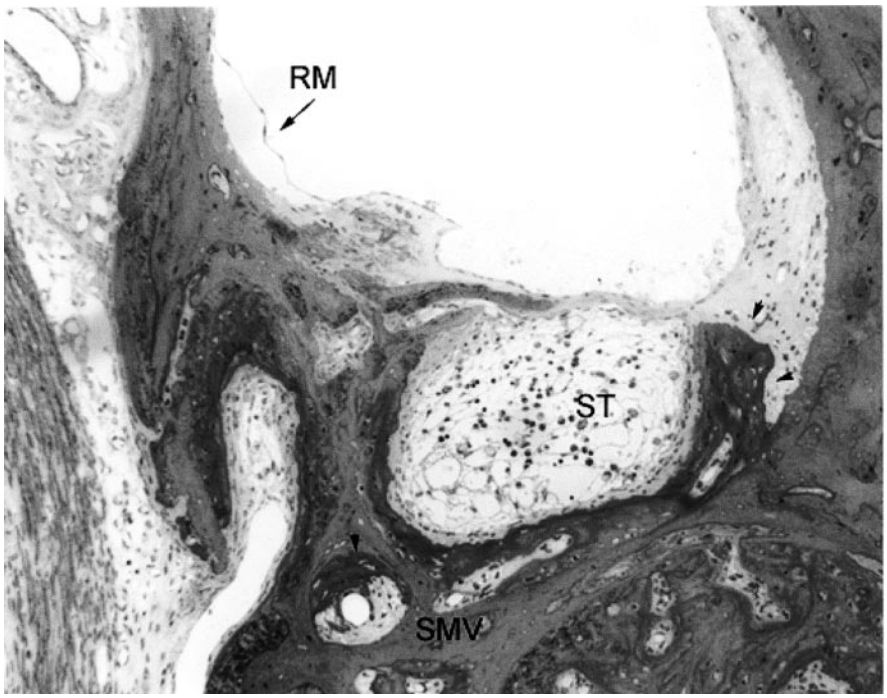


FIGURE 19–4. Photomicrograph of the middle turn of an inflamed guinea pig cochlea 4 weeks following challenge with a foreign protein to which the animal was sensitized. The scala tympani still contains infiltrated inflammatory cells, but the fibrotic matrix has begun to ossify both within the scala tympani (ST) and around the spiral modiolar vein (SMV) (*arrowheads*). The location of Reissner's membrane (RM) indicates severe hydrops.

the first. Aside from an empiric trial with high-dose corticosteroids showing improved inner ear function, Western blot assays are currently the most widely used category of diagnostic test for AIED. The initial assays used for this purpose were based on proteins extracted from bovine inner ear tissue, and inner ear extracts are still used. Reactivity to an approximately 68 kD antigen was detected in a significant proportion of patients with AIED and Meniere's disease.[27,33,49–51] This or an antigen with shared epitopes was later shown to be present in kidney and to be a member of the heat shock protein (HSP) family.[52,53] Reactivity against inner ear antigens of other molecular weights, especially in the 45, 30, and 20 kD ranges, has frequently been reported.[34,40] Few of these other antigens have been characterized or found to be present in statistically significant proportions of hearing loss versus control populations. For example, studies examining the antigenic profile of inner ear tissues consistently demonstrate a multitude of antigens against which human sera reacts; however, these bands seen on Western blot, failed to reach significance when carefully matched against controls.[33,49,54] Immunoreactivity to a 30 kD antigen has been associated with myelin protein P0.[41] The 68 kD antigen has been associated with HSP 70, and immunoreactivity against the bovine HSP 70 (bHSP 70) has been found to be correlated with AIED.[52] However, preabsorption with bHSP 70 does not remove all reactivity to the 68 kD inner ear antigen,[52] and immunization of animals with HSP 70 does not appear to produce hearing loss.[55] Antibodies against other antigens have also been implicated. Several investigators have reported that immunoreactivity to inner ear proteins with molecular weights in the 42 to 45 kD range is also positively correlated with AIED, although with a lower level of specificity than the 68 kD protein or HSP 70. For example, Atlas et al found that positive Western blots against 68 and/or 42 to 45 kD inner ear proteins were present in significantly more patients with Meniere's disease than in normal controls, whereas reactivity against 35 to 36 and 20 kD proteins was not different in the two groups.[51] Moreover, reactivity appeared to be related to disease state in that patients with active disease (at least one episode of vertigo within 1 month) were significantly more likely to be positive than those with inactive disease. Other specific antigenic targets have also been studied. Modugno et al observed antithyroid antibodies in 27% of patients with benign paroxysmal positional vertigo, significantly more than were observed in a group of normal controls.[56] Using immunoblotting and

enzyme-linked immunosorbent assay, Yamawaki et al observed antibodies against the sulfoglucuronosyl glycolipid SGLPG, but not SGPG, in 37 of 74 patients with AIED, as compared with only 3 of 56 pathologic and 2 of 28 healthy controls.[57] In a negative result, Lopez-Gonzalez et al found that anti–type II collagen antibodies and stimulation indexes were unrelated to idiopathic or sudden sensorineural hearing loss or Meniere's disease.[58]

In one of the few prospective studies attempting to correlate the significance of Western blotting for the 68 kD antigen with response to treatment, Moscicki et al found that a positive result for 68 kD was associated with a 75% rate of hearing improvement with corticosteroids, compared with 18% of patients who were Western blot negative.[50] In this study, disease activity was an important predictor of a positive antibody response and response to treatment. Eighty-nine percent of patients with active bilateral progressive hearing loss had a positive 68 kD antibody, whereas patients with inactive disease were uniformly negative.

A number of other recent studies have focused on the diagnostic utility of antibodies against HSP 70. In a retrospective case series, Hirose et al evaluated a variety of assays for systemic autoimmune disease, as well as a Western blot assay against bHSP 70, for their utility in predicting the responsiveness of rapidly progressive SNHL to corticosteroids.[59] Again in this study, positivity in the HSP 70 blot assay was the best predictor of corticosteroid responsiveness in AIED. Although the sensitivity was low (42%), the specificity was 90% and the positive predictive value was 91%. Bloch et al tested the serum of 52 patients with bilateral, progressive, idiopathic hearing loss or Meniere's disease in Western blot assays to recombinant bHSP 70 and recombinant human HSP 70.[60] Reactivity against recombinant bHSP 70 was observed in 40 of 52 patients. Only 12 also reacted to recombinant human HSP 70. They also tested the positive sera against a panel of recombinant peptide fragments of bHSP 70. Reactivity was observed to widely separate epitopes. However, most positive patients reacted preferentially or only to an amino acid segment from the carboxy terminus of rbHSP (recombinant bovine), aa 427-461. Within this dominant epitope, the bovine peptide differs from the corresponding human peptide by only one amino acid.

Western blotting has also been used to explore the possibility that a subset of patients has Meniere's disease with an immunologic basis. Gottschlich et al demonstrated that 32% of patients with Meniere's disease were anti–68 kD positive.[33] Rauch et al reported that anti–HSP 70 antibodies were found in 47% of patients with Meniere's disease and that level increased to 58% when the disease was bilateral.[61] Recently, however, Rauch et al had a similar level of sensitivity in 134 patients with Meniere's disease but a much lower level of specificity owing to a relatively

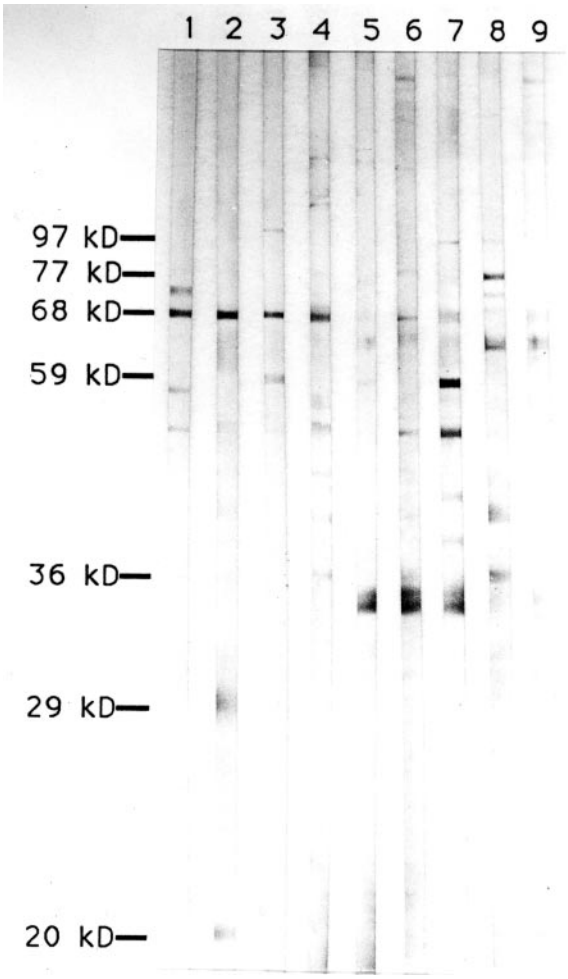


FIGURE 19–5. Western blot indicating reactivity of patient sera against antigens extracted from bovine inner ear tissue. Lanes 1 to 3 and 5 to 8 are sera from patients with sensorineural hearing loss. Sera have reactivity against a 68 kD protein or 33 to 35 kD proteins. There is also reactivity with 77, 52, and 58 kD antigens. Lane 4 is serum from a patient with giant cell arteritis and lane 9 is from a normal-hearing individual. Reproduced with permission from Gottschlich S et al.[33]

high level of reactivity in blood donor control sera.[62] This high rate of positivity in the control serum was unexplained and contradicts the previously low level of positives in control sera as well as those high rates of positivity reported by others.[27,53,61] Serial serum samples revealed no correlation between antibody level and the clinical course of the disease. These observations led them to question the clinical utility of the HSP 70 assay in Meniere's disease. The answer to whether a subset of patients with classic Meniere's disease is immune mediated is currently unclear.

A number of investigators have reacted patient sera against tissue sections to detect immunoreactivity against inner ear antigens. This technique has been used on a research basis since the 1980s, when Arnold et al reported a high degree of labeling within the cochlea with patient serum.[63] Bachor et al detected immunoreactivity to rat cochlear sections in 14 of 15 patients showing progressive or sudden hearing loss in the cochlea opposite an ear deafened by trauma or inflammation.[64] Using rat cryosections of the inner ear, Arbusow et al detected antibodies against vestibular sensory epithelia in 8 of 12 patients with idiopathic bilateral vestibular pathology as compared with 1 of 22 healthy controls and 0 of 6 patients with systemic autoimmune disease.[65] Ottaviani et al detected immunoreactivity against endothelial cells using sections of rat kidney tissue in 8 of 15 patients with sudden hearing loss, as compared with 2 of 14 normal controls.[66] Helmchen et al observed positive but low levels of immunoreactivity against inner ear sections using serum from patients with Cogan's syndrome.[67] However, unlike anticorneal antibodies, the anticochlear immunoreactivity levels were not correlated with disease stage.

CLASSIFICATION OF AUTOIMMUNE INNER EAR DISEASE

Over the past two decades since McCabe's published article on AIED,[18] many patients have been diagnosed and treated for rapidly progressive SNHL and many have had their hearing maintained or even improved with treatment. As a result of the growing experience with patients with corticosteroid sensitive hearing loss, a pattern has begun to emerge that warrants a classification scheme to sort out patients better as they present with such a broad category of inner ear dysfunction. Although the following classification scheme is intended specifically for that purpose, it is likely that over the next few years it will be further refined:

Type 1: Organ (Ear) Specific

- Rapidly progressive bilateral SNHL
- All age ranges, although middle age is most common
- No other clinical evidence of systemic autoimmune disease
- Positive Otoblot (Western blot 68 kD or HSP 70)
- Negative serologic studies (antinuclear antibody [ANA], erythrocyte sedimentation rate, rheumatoid factor (RF), C1q binding assay, etc)
- Greater than 50% response rate to high-dose corticosteroids

Type 2: Rapidly Progressive Bilateral Sensorineural Hearing Loss with Systemic Autoimmune Disease

- Rapidly progressive bilateral SNHL
- Hearing loss often worst with flare of autoimmune condition
- Other autoimmune condition is present (SLE, ulcerative colitis, polyarteritis nodosa, vasculitis, rheumatoid arthritis, Sjögren's syndrome)
- Otoblot may be positive or negative
- Serologic studies will be positive in accordance with the illness (ie, ANA-high titers, RF positive, circulating immune complexes)
- Corticosteroid responsive and may be managed with targeted therapies for underlying illness

Type 3: Immune-Mediated Meniere's Disease

- Bilateral, fluctuating SNHL with vestibular symptoms that may predominate
- Subset of patients with delayed contralateral endolymphatic hydrops or recent instability of better-hearing ear in a patient with burned out Meniere's disease
- Otoblot positive 37 to 58%; may show presence of circulating immune complexes
- Corticosteroid responsive; may require long-term immunosuppression owing to relapses

Type 4: Rapidly Progressive Bilateral Sensorineural Hearing Loss with Associated Inflammatory Disease (Chronic Otitis Media, Lyme Disease, Otosyphilis, Serum Sickness)

- Evidence of profound drop in hearing with long-standing chronic otitis media

- May show inflammation of the tympanic membrane and perforations
- Hearing loss progresses despite treatment of the infectious agent (treponemal or rickettsial)
- Otoblot negative; serologic tests for the underlying disease may be positive; patient should be evaluated for granulomatous disease and vasculitis by biopsy if tissue is available
- Corticosteroid responsive and may require long-term immunosuppression
- Serum sickness has been reported after vaccinations, although anecdotal

Type 5: Cogan's Syndrome

- Sudden onset of interstitial keratitis and severe vestibuloauditory dysfunction
- Otoblot negative for 68 kD but positive for 55 kD antigen
- Responds to high-dose corticosteroids, although becomes resistant over long term

Type 6: Autoimmune Inner Ear Disease-Like

- Young patients with idiopathic rapidly progressive bilateral SNHL leading to deafness
- Severe ear pain, pressure, and tinnitus
- Otoblot and all serology negative
- May have an unrelated, nonspecific inflammatory event that initiates ear disease
- Not responsive to immunosuppressive drugs, although they are tried

TREATMENT

Once a diagnosis of AIED is established or considered highly presumptive, high-dose prednisone is the mainstay of treatment for this condition. Early institution of 60 mg of prednisone daily for a month is now widely used as short-term or lower-dose long-term therapy and has either been ineffective or fraught with the risk of relapse. Prednisone is then tapered slowly if a positive response to therapy is obtained. If during the taper hearing suddenly falls, reinstitution of high-dose prednisone is indicated. One sensitive predictor of imminent relapse can be the appearance of loud tinnitus in one or both ears. If patients show corticosteroid responsiveness but attempts at taper result in relapse, the addition of a cytotoxic drug should be considered. The most widely used of these agents are methotrexate (MTX) and cyclophosphamide (Cytoxan). The former has the advantage of being less toxic and has fewer long-term hematopoietic risks, such as the development of neoplasia.[68] If MTX is used, it should be given as an oral dose 7.5 to 20 mg weekly with folic acid. The patient should be monitored closely for toxicity with complete blood count, platelets, blood urea nitrogen, creatinine, liver function tests, and urinalysis. It should be noted that the prednisone-sparing effects of MTX may take 1 to 2 months to achieve; therefore, prednisone should be maintained until such effects are obtained. Also, if high-dose prednisone has not been effective in restoring hearing, it is unlikely that MTX will offer additional efficacy. Therefore, for patients with severe hearing losses, positive 68 kD Western blots, and nonresponsiveness to prednisone or MTX therapy, consideration should be given to a trial of cyclophosphamide.[26] At oral doses of 1 to 2 mg per day taken each morning with liberal amounts of fluid, the risk of hemorrhagic cystitis or drug effects on the bladder can be minimized. Again, appropriate monitoring of peripheral blood counts is required. Cyclophosphamide should not be administered to children, and the risk of permanent sterility should be outlined. If, on the other hand, no response to high-dose prednisone is achieved, and the patient is 68 kD Western blot negative, it may be futile to continue potentially toxic drugs, with little evidence for AIED as the cause. As this field continues to evolve, there are, however, no hard and fast rules, and a practitioner may be justified in trying cytotoxic drugs on an empiric basis because unrelenting progressive deafness is a serious handicap for a previously normal-hearing person. Luetje recommended plasmapheresis for difficult to manage patients,[69] and this can be a useful adjunct to the above-mentioned immunosuppressive drugs. At the time of writing, a multi-institutional clinical trial is under way to compare the efficacy of MTX and prednisone versus prednisone alone for the management of AIED. The results of this trial should help to delineate appropriate therapy for suspected AIED.

Parnes et al noted that local corticosteroids appear to be more effective in the treatment of other autoimmune disorders, such as corneal inflammation owing to Cogan's syndrome.[70] They therefore investigated the pharmacokinetics of hydrocortisone, methylprednisone, and dexamethasone in perilymph and endolymph after oral, intravenous, or

intratympanic administration. Dexamethasone was found to be largely excluded from the cochlea by the blood-labyrinthine barrier. Both methylprednisone and hydrocortisone reached inner ear fluid after systemic administration, attenuated presumably by the blood-labyrinthine barrier. Much higher levels of all three drugs were observed in cochlear fluid after intratympanic administration, with rapid declines over a 6- to 24-hour period. Similar results were noted by Chandrasekhar et al.[71] Parnes et al also reported that repeated intratympanic administration of corticosteroids in a small series of patients with hearing loss of diverse origins was followed by improvement in some patients, but no control group was included.[70] In contrast, Yang et al found that local immunosuppression had no effect on experimental immune-mediated SNHL in an animal model.[72] It should be noted that local effects are not, of course, the only basis for the therapeutic efficacy of immunosuppressants. By decreasing peripheral blood leukocytes, these agents reduce the population of cells that can be recruited to the inner ear to participate in immune and inflammatory damage. An experiment designed to prevent entry of cells into the cochlea using antibodies to ICAM-1 did show a reduced number of infiltrated inflammatory cells in the cochlea following antigen challenge.[73] Although the inflammation was not entirely prevented, such a strategy may be worth pursuing. An analogous situation exists for ocular immune disorders such as uveitis. Despite the greater accessibility of the eye to topical drugs than the inner ear, ophthalmologists would never consider local therapy in lieu of high-dose corticosteroids for these disorders. Perhaps a lesson taken from their experience might lessen the enthusiasm that currently exists for treatment solely by local middle ear corticosteroid instillation.

DISCUSSION

Debate continues as to whether AIED exists as a separate entity. Some authors prefer to refer to this condition as immune-mediated inner ear disease. Clearly, the evidence for specific autoimmunity is indirect. Hearing and vestibular problems that are diagnosed as autoimmune in origin are often responsive to corticosteroids. Although this suggests that the condition involves inflammation, one cannot infer the involvement of specific immunity. The fact that inner ear disease is often present in systemic autoimmune disorders provides strong evidence that autoimmune processes can damage the labyrinth but does not speak to the issue of organ-specific disease. Animal models of hearing loss and/or vestibular dysfunction secondary to immunization with inner ear antigens provide stronger evidence of specific autoimmunity.

Autoimmune inner ear disease is difficult to diagnose. Although it is generally agreed that the condition should be bilateral and rapidly progressive, the involvement of the second ear may take months or even years to manifest. Although rapidly progressing conditions are more readily held to be autoimmune, AIED is increasingly considered as a potential cause of Meniere's disease and as a less likely cause of sudden hearing loss.

Improved diagnostic tests are clearly required. No one test appears to be positive in more than 30 to 40% of patients who otherwise fit the criteria for autoimmune disease. One possible explanation for this is that rapidly progressive SNHL has a number of different causes, including autoimmune, viral, genetic, developmental, vascular, and perhaps metabolic. Many of these cannot be separated by their presentation; therefore, it would not be unusual or unexpected for many of these patients to have negative antibody testing, and some who were not autoimmune might even improve with corticosteroids (eg, viral). Another possibility is that autoimmunity exists to a variety of inner ear antigens. Given the variety of autoimmune disorders that can affect the inner ear, the variety of antigens with which sera from patients with AIED will react, and the fact that immunization with a variety of proteins can lead to hearing loss in animal models, this would appear to be a strong possibility.

The usefulness of Western blotting for antibodies directed against the 68 kD or HSP 70 antigen as diagnostic assay seems clear, although there is little evidence to support an etiologic role for HSP 70. It is possible that HSP 70 shares one or more epitopes with an inner ear antigen, although reactivity to widely variable epitopes of HSP 70[60] argues against this. Alternatively, HSP 70 immunoreactivity may all be a well-correlated epiphenomenon, perhaps produced by immunization of self-proteins during inflammatory responses arising from other causes. Lastly, initial studies with serum tested by Western blotting were with the use of 68 kD inner ear tissue as the target antigen. After the recognition

that HSP 70 showed results similar to 68 kD by several investigators, a number of groups have adopted HSP 70 as the target for immunologic testing. In fact, this may be the wrong approach if HSP 70 merely shares epitopes with but is not the actual antigen in 68 kD inner ear immunoreactivity. Future studies will certainly improve our knowledge of the actual antigenic target(s) involved in AIED.

Despite uncertainty over etiology and difficulties in diagnosis, this condition is frequently responsive to treatment with immunosuppressive drugs. Since there are few forms of SNHL that can be treated other than symptomatically, AIED represents a unique opportunity to reverse SNHL and vestibular disorders. For this reason alone, the diagnosis should be considered when symptoms are appropriate, and both clinical and basic research on this condition is warranted.

REFERENCES

1. Harris JP. Immunology of the inner ear: response of the inner ear to antigen challenge. Otolaryngol Head Neck Surg 1983;91:18–32.
2. Harris JP. Immunology of the inner ear: evidence of local antibody production. Ann Otol Rhinol Laryngol 1984;93:157–62.
3. Harris JP, Ryan AF. Immunobiology of the inner ear. Am J Otolaryngol 1984;5:418–25.
4. Harris JP, Woolf NK, Ryan AF. Elaboration of systemic immunity following inner ear immunization. Am J Otolaryngol 1985;6:148–52.
5. Woolf NK, Harris JP. Cochlear pathophysiology associated with inner ear immune responses. Acta Otolaryngol (Stockh) 1986;102:353–64.
6. Ma C, Billings P, Harris JP, Keithley EM. Characterization of an experimentally induced inner ear immune response. Laryngoscope 2000;110:451–6.
7. Tomiyama S, Harris JP. The endolymphatic sac: its importance in inner ear immune responses. Laryngoscope 1986;96:685–91.
8. Tomiyama S, Harris JP. The role of the endolymphatic sac in inner ear immunity. Acta Otolaryngol (Stockh) 1987;103:182–8.
9. Harris JP, Woolf NK, Ryan AF, et al. Immunologic and electrophysiological response to cytomegaloviral inner ear infection in the guinea pig. J Infect Dis 1984;150:523–30.
10. Woolf NK, Harris JP, Ryan AF, et al. Hearing loss in experimental cytomegalovirus infection of the guinea pig inner ear: prevention by systemic immunity. Ann Otol Rhinol Laryngol 1985;94:350–6.
11. Darmstadt GL, Keithley EM, Harris JP. Effects of cyclophosphamide on the pathogenesis of cytomegalovirus-induced labyrinthitis. Ann Otol Rhinol Laryngol 1990;99:960–8.
12. Harris JP, Fukuda S, Keithley EM. Spiral modiolar vein: its importance in inner ear inflammation. Acta Otolaryngol (Stockh) 1990;110:357–65.
13. Stearns GS, Keithley EM, Harris JP. Development of high endothelial venule-like characteristics in the spiral modiolar vein induced by viral labyrinthitis. Laryngoscope 1993;103:890–8.
14. Takahashi M, Harris JP. Analysis of immunocompetent cells following inner ear immunostimulation. Laryngoscope 1988;98:1133–8.
15. Suzuki M, Harris JP. Expression of intercellular adhesion molecule-1 during inner ear inflammation. Ann Otol Rhinol Laryngol 1995;104:69–75.
16. Keithley EM, Woolf NK, Harris JP. Development of morphological and physiological changes in the cochlea induced by cytomegalovirus. Laryngoscope 1989;99:409–14.
17. Chen MC, Harris JP, Keithley EM. Immunohistochemical analysis of proliferating cells in a sterile labyrinthitis animal model. Laryngoscope 1998; 108:651–6.
18. McCabe BF. Autoimmune sensorineural hearing loss. Ann Otol Rhinol Laryngol 1979;88:585–9.
19. Schiff M, Brown M. Hormones and sudden deafness. Laryngoscope 1974;84:1959–81.
20. Clemis JD, Mastricola PG, Schuler-Vogler M. Sudden hearing loss in the contralateral ear in postoperative acoustic tumor: three case reports. Laryngoscope 1982;92:76–9.
21. Teryama Y, Saski Y. Studies on experimental allergic (isoimmune) labyrinthitis in guinea pigs. Acta Otolaryngol (Stockh) 1963;58:49–61.
22. Hughes GB, Barna BP, Kinney SE, et al. Autoimmune endolymphatic hydrops: five-year review. Otolaryngol Head Neck Surg 1988;98:221–5.
23. Moskowitz D, Lee KJ, Smith HW. Steroid use in idiopathic sudden sensorineural hearing loss. Laryngoscope 1984;94:664–6.
24. Ariyasu L, Byl FM, Sprague MS, Adour KK. The beneficial effect of methylprednisone in acute vestibular vertigo. Arch Otolaryngol Head Neck Surg 1990; 166:700–3.

25. Harris JP, Ryan AF. Fundamental immune mechanisms of the brain and inner ear. Otolaryngol Head Neck Surg 1995;112:639–53.
26. McCabe BF. Autoimmune inner ear disease. In: Bernstein J, Ogra P, editors. Immunology of the ear. New York: Raven Press; 1987. p. 427–35.
27. Harris JP, Sharp PA. Inner ear autoantibodies in patients with rapidly progressive sensorineural hearing loss. Laryngoscope 1990;100:516–24.
28. Schuknecht HF. Ear pathology in autoimmune disease. Adv Otorhinolaryngol 1991;46:50–70.
29. Hoistad DL, Schachern PA, Paparella MM. Autoimmune sensorineural hearing loss: a human temporal bone study. Am J Otolaryngol 1998;19:33–9.
30. Jenkins HA, Pollak AM, Fisch U. Polyarteritis nodosa as a cause of sudden deafness: a human temporal bone study. Am J Otolaryngol 1981;2:99–107.
31. Keithley EM, Chen MC, Linthicum F. Clinical diagnoses associated with histologic findings of fibrotic tissue and new bone in the inner ear. Laryngoscope 1998;108:87–91.
32. Hughes GB, Barna BP, Kinney SE, et al. Predictive value of laboratory tests in "autoimmune" inner ear disease: preliminary report. Laryngoscope 1986; 96:502–5.
33. Gottschlich S, Billings PB, Keithley EM, et al. Assessment of serum antibodies in patients with rapidly progressive sensorineural hearing loss and Meniere's disease. Laryngoscope 1995;105:1347–52.
34. Bouman H, Klis SF, Meeuwsen F, et al. Experimental autoimmune inner ear disease: an electrocochleographic and histophysiologic study. Ann Otol Rhinol Laryngol 2000;109:457–66.
35. Soliman AM, Zanetti F. Improvements of a method for testing autoantibodies in sensorineural hearing loss. Adv Otorhinolaryngol 1988;39:13–7.
36. Harris JP. Experimental autoimmune sensorineural hearing loss. Laryngoscope 1987;97:63–76.
37. Nair TS, Raphael Y, Dolan DF, et al. Monoclonal antibody induced hearing loss. Hear Res 1995; 83:101–13.
38. Ruckenstein MJ, Sarwar A, Hu L, et al. Effects of immunosuppression on the development of cochlear disease in the MRL-Fas(lpr) mouse. Laryngoscope 1999;109:626–30.
39. Sone M, Schachern PA, Paparella MM, Morizono N. Study of systemic lupus erythematosus in temporal bones. Ann Otol Rhinol Laryngol 1999; 108:338–44.
40. Cao MY, Gersdorff M, Deggouj N, et al. Detection of inner ear disease autoantibodies by immunoblotting. Mol Cell Biochem 1995;146:157–63.
41. Matsuoka H, Cheng KC, Krug MS, et al. Murine model of autoimmune hearing loss induced by myelin protein P0. Ann Otol Rhinol Laryngol 1999;108:255–64.
42. Gloddek B, Lassmann S, Gloddek J, Arnold W. Role of S-100beta as potential autoantigen in an autoimmune disease of the inner ear. J Neuroimmunol 1999;101:39–46.
43. Gloddek B, Gloddek J, Arnold W. A rat T-cell line that mediates autoimmune disease of the inner ear in the Lewis rat. ORL J Otorhinolaryngol Relat Spec 1999;61:181–7.
44. Nair TS, Prieskorn DM, Miller JM, et al. KHRI-3 monoclonal antibody-induced damage to the inner ear: antibody staining of nascent scars. Hear Res 1999;129:50–60.
45. Iwai H, Tomoda K, Sugiura K, et al. T cells infiltrating from the systemic circulation proliferate in the endolymphatic sac. Ann Otol Rhinol Laryngol 1999;108:1146–50.
46. Ruckenstein MJ, Hu L. Antibody deposition in the stria vascularis of the MRL-Fas(lpr) mouse. Hear Res 1999;127:137–42.
47. Wobig RJ, Kempton JB, Trune DR. Steroid-responsive cochlear dysfunction in the MRL/lpr autoimmune mouse. Otolaryngol Head Neck Surg 1999; 121:344–7.
48. Khan DC, DeGagne JM, Trune DR. Abnormal cochlear connective tissue mineralization in the Palmerston North autoimmune mouse. Hear Res 2000;142:12–22.
49. Shin SO, Billings PB, Keithley EM, Harris JP. Comparison of anti-heat shock protein 70 (anti-hsp70) and anti-68-kDa inner ear protein in the sera of patients with Ménière's disease. Laryngoscope 1997;107:222–7.
50. Moscicki RA, San Martin JE, Quintero CH, et al. Serum antibody to inner ear proteins in patients with progressive hearing loss. JAMA 1994;272: 611–6.
51. Atlas MD, Chai F, Boscato L. Ménière's disease: evidence of an immune process. Am J Otol 1998; 19:628–31.
52. Billings PB, Keithley EM, Harris JP. Evidence linking the 68 kilodalton antigen identified in progressive sensorineural hearing loss patient sera with heat

shock protein 70. Ann Otol Rhinol Laryngol 1995; 104:181–8.
53. Bloch DB, San Martin JE, Rauch SD, et al. Serum antibodies to heat shock protein 70 in sensorineural hearing loss. Arch Otolaryngol Head Neck Surg 1995;121:1167–71.
54. Yamanobe S, Harris JP. Inner ear specific autoantibodies. Laryngoscope 1993;103:319–25.
55. Trune DR, Kempton JB, Mitchel CR, Heferneider SH. Failure of elevated heat shock protein 70 antibodies to alter cochlear function in mice. Hear Res 1998;116:65–70.
56. Modugno GC, Pirodda A, Ferri GG, et al. A relationship between autoimmune thyroiditis and benign paroxysmal positional vertigo? Med Hypoth 2000;54:614–5.
57. Yamawaki M, Ariga T, Gao Y, et al. Sulfoglucuronosyl glycolipids as putative antigens for autoimmune inner ear disease. J Neuroimmunol 1998;84:111–6.
58. Lopez-Gonzalez MA, Lucas M, Sanchez B, et al. Autoimmune deafness is not related to hyperreactivity to type II collagen. Acta Otolaryngol (Stockh) 1999;119:690–4.
59. Hirose K, Wener MH, Duckert LG. Utility of laboratory testing in autoimmune inner ear disease. Laryngoscope 1999;109:1749–54.
60. Bloch DB, Gutierrez JA Jr, Guerriero V, et al. Recognition of a dominant epitope in bovine heat-shock protein 70 in inner ear disease. Laryngoscope 1999;109:621–5.
61. Rauch SD, San Martin JE, Moscicki RA, Bloch KJ. Serum antibodies against heat shock protein 70 in Meniere's disease. Am J Otol 1995;16:648–52.
62. Rauch SD, Zurakowski D, Bloch DB, Bloch KJ. Anti-heat shock protein 70 antibodies in Meniere's disease. Laryngoscope 2000;110:1516–21.
63. Arnold W, Pfaltz R, Altermatt HJ. Evidence of serum antibodies against inner ear tissues in the blood of patients with certain sensorineural hearing disorders. Acta Otolaryngol (Stockh) 1985;99: 437–44.
64. Bachor E, ten Cate WJ, Gloddek B, Ehsani N. Immunohistochemical detection of humoral autoantibodies in patients with hearing loss in the last hearing ear. Laryngorhinootologie 2000;79:131–4.
65. Arbusow V, Strupp M, Dieterich M, et al. Serum antibodies against membranous labyrinth in patients with "idiopathic" bilateral vestibulopathy. J Neurol 1998;245:132–6.
66. Ottaviani F, Cadoni G, Marinelli L, et al. Anti-endothelial autoantibodies in patients with sudden hearing loss. Laryngoscope 1999;109:1084–7.
67. Helmchen C, Arbusow V, Jager L, et al. Cogan's syndrome: clinical significance of antibodies against the inner ear and cornea. Acta Otolaryngol (Stockh) 1999;119:528–36.
68. Sismanis A, Thompson T, Willis HE. Methotrexate therapy for autoimmune hearing loss: a preliminary report. Laryngoscope 1994;104:932–4.
69. Luetje CM. Theoretical and practical implications for plasmapheresis in autoimmune inner ear disease. Laryngoscope 1989;99:1137–46.
70. Parnes LS, Sun AH, Freeman DJ. Corticosteroid pharmacokinetics in the inner ear fluids: an animal study followed by clinical application. Laryngoscope 1999;109:1–17.
71. Chandrasekhar SS, Rubinstein RY, Kwartler JA, et al. Dexamethasone pharmacokinetics in the inner ear: comparison of route of administration and use of facilitating agents. Otolaryngol Head Neck Surg 2000;122:521–8.
72. Yang GS, Song HT, Keithley EM, Harris JP. Intratympanic immunosuppressives for prevention of immune-mediated sensorineural hearing loss. Am J Otol 2000;21:499–504.
73. Takasu T, Harris JP. Reduction of inner ear inflammation by treatment with anti-ICAM-1 antibody. Ann Otol Rhinol Laryngol 1997;106:1070–5.

CHAPTER 20

Meniere's Disease, Vestibular Neuronitis, Paroxysmal Positional Vertigo, and Cerebellopontine Angle Tumors

Jacob Johnson, MD, Anil K. Lalwani, MD

The value and function of the vestibular system may often be underestimated when considering the various special senses that we possess. However, of all of the special senses, unilateral loss of the vestibular system may cause the most significant detriment for our daily function and survival.[1] This chapter discusses the common disorders that affect the vestibular system and provides a framework for evaluation, diagnosis, and treatment of patients with vestibular disorders. The common disorders affecting the peripheral vestibular system include benign paroxysmal positional vertigo (BPPV), Meniere's disease, and vestibular neuronitis. The most common disorder of the vestibular nerve is a vestibular schwannoma (VS), accounting for 80% of all cerebellopontine angle (CPA) lesions.[2] This chapter also discusses some of the other common (meningiomas and epidermoids) and uncommon tumors of the CPA that present by injuring the cochleovestibular system.

SIMPLIFIED VESTIBULAR ANATOMY AND FUNCTION

A brief review of vestibular anatomy and function, nystagmus, and central compensation to vestibular injury will facilitate the discussion of peripheral vestibular disorders and explain the lack of significant vestibular symptoms in patients with CPA tumors. The vestibular system provides positional information at rest and positioning information during movement. This information allows us to have visual tracking, perception of space, sense of direction, and postural balance during movement and rest. The components of the vestibular system include the peripheral organs, the vestibular division of the eighth cranial nerve (CN), the vestibular nuclei, and the projections to and from the vestibular nuclei. The peripheral organs are endolymph-filled membranous sacs suspended in the perilymph of the bony labyrinth and consist of two otolith organs, the saccule and the utricle, and three orthogonally oriented semicircular ducts. The saccule is joined to the cochlear duct (scala media) via the ductus reuniens. The sensory receptors (vestibular hair cells) are located in the cristae of the semicircular ducts and the maculae of the saccule and utricle. The cristae respond to angular acceleration associated with head rotation, and the maculae are stimulated by linear acceleration caused by translational or tilting head movements. The cristae are normally not sensitive to gravity, and this information is transduced by the maculae. The vestibular organs are innervated by the superior (utricle, lateral and superior semicircular ducts) and inferior (saccule and posterior semicircular duct) vestibular nerves. The posterior semicircular duct is innervated by the posterior branch of the inferior vestibular nerve that passes through the foramen singulare. The superior and inferior vestibular nerves travel in the internal auditory canal (IAC) with the cochlear and facial nerves. The vestibular nerves join with the cochlear nerve to form CN VIII, which crosses the CPA to reach the vestibular nuclei. The vestibular nuclei are located in the pons and have primary projections to the oculomotor, spinal, cerebellar, autonomic, and cerebral areas.[3]

The vestibular nerves have a baseline tonic firing rate. The fundamental strategy of the vestibular system is to compare the relative change in the fir-

ing rate based on the response of paired peripheral organs located in each temporal bone for any given head movement. For example, a brief head rotation to the right in the horizontal plane causes deflection of the cupulae of the right and left horizontal semicircular ducts. The right cupula is deflected toward the ampullated end of its semicircular duct, which leads to an increase in the firing rate of the right vestibular nerve, and the opposite effect occurs at the left horizontal duct and left vestibular nerve (Ewald's laws). This information regarding right angular acceleration is processed in the vestibular nuclei and shared with the visual system to allow visual fixation during the head movement and the cerebellum and spinal areas to maintain balance and posture. The response of the visual and balance systems to head movements provides a way to study the vestibular system and assess its function during injury. The most studied response is the vestibular oculomotor-reflex (VOR). The VOR holds an image on the retina during brief head movement. Understanding the VOR allows us to predict how the eyes should move in response to head movement. In the above example of a brief head movement to the right, the right part of the vestibular system is stimulated relative to the left part of the vestibular system. The VOR is manifest by contraction of the right medial rectus and left lateral rectus muscles. This allows the eyes to stay fixed on a point by relatively moving the eyes to the left as the head moves to the right. A "fast" reset of the eyes by the opposing pair of eye muscles allows visual acquisition of a new object along the new line of sight. This rapid type of eye movement to acquire a new target is called a saccade. If the right semicircular duct is persistently stimulated as in a warm caloric test, these eye movements (slow phase to the left and fast phase to the right) become repetitive and are called jerk nystagmus. The fast phase of the eye movement is to the right; therefore, the nystagmus is called right-beating nystagmus. A few rules will allow predicting the expected nystagmus for stimulation of the various semicircular ducts. The eye movements occur in the same plane as the stimulated semicircular duct, and the fast phase is toward the side of the stimulated duct or away from an injured vestibular organ (Ewald's laws).[1]

Injury to the peripheral or central vestibular system causes asymmetry in the baseline input into the vestibular centers, and this causes vertigo, nystagmus, vomiting, and a sense of falling toward the side of the injury.[4] Vertigo is defined as the illusion of movement. However, the chief complaint of patients with injury to the vestibular system is usually not vertigo but dizziness. If the complaint is clarified to be vertigo, the duration, periodicity, and circumstance of the vertigo and the presence of other neurologic signs or symptoms allow for categorization of the vertigo. The proximity of the vestibular system to the auditory system often causes vertigo to be coupled with hearing loss. Knowing the duration of the vertigo or dysequilibrium and the presence or absence of hearing loss allows narrowing the differential diagnosis[5–7] (Table 20–1). The vertigo may be caused by injury of the peripheral or central parts of the vestibular system. Often the presence of other neurologic abnormalities leads to an investigation for a central cause of the vertigo. However, central vestibular injury caused by a mass lesion or stroke may mimic a peripheral vestibular disorder. In the section on vestibular neuronitis, the nystagmus characteristic of peripheral versus central vestibular disorders is presented.

The central compensation for vestibular injury occurs via the cerebellum. The cerebellum provides a "clamping" response to the injured vestibular system to reduce the effects of the abnormal vestibular signal. In an acute injury, such as vestibular neuronitis, the vertiginous response lasts 3 to 5 days, and then the central compensation is able to modulate the signal from the injured vestibular system.[8] In episodic insults, such as that occur in Meniere's disease, the central compensation is not able to be as effective, so with each new episode there are acute vertiginous symptoms. In a slowly evolving process such as a VS, the central compensation occurs in step with the vestibular dysfunction, and the patient may have minimal to no vestibular symptoms. The central compensation is enhanced by vestibular activity and delayed by prolonged use of medical vestibular suppression. This observation has led to the development of vestibular rehabilitation programs. Vestibular rehabilitation programs use three strategies: (1) habituation exercises, which facilitate central compensation by extinguishing pathologic responses to head motion; (2) postural control exercises; and (3) general conditioning exercises. Vestibular rehabilitation is critically important in the elderly since their ability to have optimal central compensation is diminished.[9]

TABLE 20–1. Differential Diagnosis of Vertigo Based on Time Frame of Vertigo and Presence or Absence of Hearing Loss

Duration of Vertigo	*Hearing Loss Absent*	*Hearing Loss Present*
Seconds	Benign paroxysmal positional vertigo	Perilymphatic fistula Cholesteatoma
Minutes	Vertebral/basilar artery insufficiency Migraines	
Hours	Vestibulopathy	Meniere's disease
Days	Vestibular neuronitis	Labyrinthitis
Weeks	Central nervous system lesions Lyme disease Multiple sclerosis	Vestibular schwannoma Autoimmune processes Psychogenic

COMMON PERIPHERAL VESTIBULAR DISORDERS

Millions of people present annually to their physicians with the complaint of dizziness. In addition to the vestibular system, dizziness may be caused by poor vision, decreased proprioception (diabetes mellitus), cardiovascular insufficiency, cerebellar or brainstem strokes, neurologic conditions (migraines, multiple sclerosis), metabolic disorders, and side effects of medications. The role of the otolaryngologist includes clarifying the subset of patients who have vertigo owing to injury to the vestibular system and differentiating central from peripheral vestibular disorders. The evaluation includes a complete head and neck and vestibular examination (Table 20–2). Diagnostic evaluation includes audiology, vestibular testing, and imaging. The majority of patients with peripheral vestibular disorders will have BPPV, Meniere's disease, or vestibular neuronitis. These patients generally improve with supportive or conservative care (medical or physical therapy). The small percentage of medically recalcitrant patients can then be helped with surgical intervention. The surgical interventions, in general, ablate the vestibular system and rely on central compensation and vestibular rehabilitation to improve the patient.

Benign Paroxysmal Positional Vertigo

Benign paroxysmal positional vertigo was first recognized by Bárány in 1921 and was further characterized by Dix and Hallpike in 1952.[10] Benign paroxysmal positional vertigo is one of the most common peripheral causes of vertigo. Patients complain of vertigo lasting seconds when in certain positions, with no associated hearing loss. The most common cause of this disorder is the presence of debris in the posterior semicircular duct. The diagnosis is made by noting a characteristic nystagmus when the patient is in the Dix-Hallpike position. The use of specific maneuvers to reposition the debris into the utricle provides relief for the majority of patients.

Epidemiology The average age of presentation is in the fifth decade, and there is no gender bias. The incidence may range from 10 to 100 cases per 100,000 persons per year.[11] Nearly 20% of patients seen at vertigo clinics will be given the diagnosis of BPPV. Ten to 15% of patients will have an antecedent history of vestibular neuronitis, and another 20% will have a history of head trauma.[12,13]

Pathogenesis Benign paroxysmal positional vertigo occurs because a semicircular duct has debris attached to the cupula or free floating in the endolymph. The crista of the semicircular duct is stimulated by the movement of these particles in response to gravity. As described earlier, the semicircular ducts normally sense angular acceleration, not gravity. The study of temporal bones from patients with BPPV shows basophilic deposits adherent to the cupula, and this finding was termed cupulolithiasis.[14] During posterior semicircular canal (PCSS)

TABLE 20–2. Vestibular Evaluation

1. Head and neck examination including evaluation of cranial nerves
2. Spontaneous and gaze-evoked nystagmus with Fresnel glasses
   - Direction: fixed-peripheral, changing-central
   - Form: jerk-peripheral, pendular-central
   - Fixation: suppression-peripheral, enhanced-central
3. Smooth pursuit: "Follow my fingers"
4. Saccades: "Look to my left or right finger when I say to"
   - Dysmetric: cerebellar
   - Slow: brainstem
   - Late: frontal lobe
   - Dysconjugate: multiple sclerosis
5. Head thurst
   - Normal: no refixation saccade
   - Abnormal: refixation saccade (peripheral)
6. Head shake: "10 degrees, 2 cycles/second, 20 seconds"
   - Normal: no nystagmus
   - Abnormal: horizontal nystagmus-peripheral, vertical nystagmus-central (brainstem)
7. Dynamic visual activity: "Look at Snellen chart with head shake"
   - Normal: < 3 line drop
   - Abnormal: 3 or more line drop—bilateral vestibular loss
8. Fixation suppression: "Look at your thumb during rotation"
   - Normal: no nystagmus
   - Abnormal: nystagmus-central (flocculus)
9. Positional testing: Dix-Hallpike maneuver
   - Normal: no nystagmus
   - Abnormal: downbeat, fatigable, rotatory nystagmus
10. Cerbellum: finger to nose, rapid alternating movements, heel to shin
11. Posture: Romberg's test

occlusion in patients with resistant BPPV, free-floating debris has been observed in the endolymph and has been termed canalolithiasis.[15] Electron microscopy of these particles shows that they are likely otoconia originating from the macula of the gravity-sensitive utricle.[16]

The cupula of the semicircular duct has the same specific gravity as the endolymph and so is not sensitive to gravity. However, the debris in the semicircular duct moves in response to gravity, and when the patient places the semicircular duct in a dependent position, the particles move and entrain endolymph with them and cause deflection of the cupula. The unexpected gravity-sensitive response from the semicircular duct causes vertigo. The majority of BPPV is caused by debris in the posterior semicircular duct, but debris may also enter the horizontal and superior semicircular ducts.[17]

Natural History The natural history of BPPV includes an acute onset and remission over a few months. However, up to 30% of patients may have symptoms for longer than 1 year. The majority of patients will improve with a repositioning maneuver. Patients may have unpredictable recurrences and remissions, and the rate of recurrence may be 10 to 15% per year.[12] These patients may be retreated with a repositioning maneuver. A subset of patients who have adapted by not using certain positions to avoid the vertigo or who have other balance disorders will benefit from balance rehabilitation therapy.

Symptoms and Signs The patient usually complains of a sudden onset of vertigo that lasts 10 to 20 seconds with certain head movements. The triggering movements include rolling over in bed into a lateral position, getting out of bed, looking up and back, and bending over. The vertigo may be associated with nausea. The patients have normal hearing, no spontaneous nystagmus, and normal neurologic evaluations.[10]

Diagnostic Evaluation *Dix-Hallpike Test.* Benign paroxysmal positional vertigo is diagnosed by observing a characteristic nystagmus when performing the Dix-Hallpike test[10] (Figure 20–1). For testing the right PSCC, the patient sits on the examination table and turns the head 45 degrees to the right. This places the PSCC in the sagittal plane. The examiner stands facing the patient on the patient's

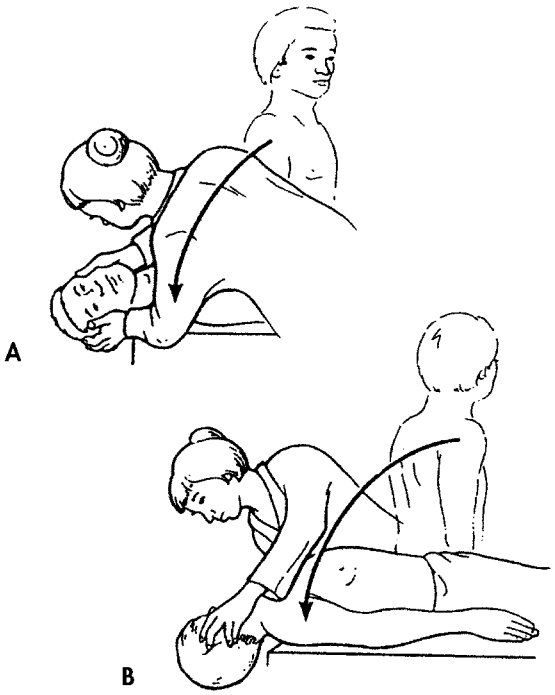


FIGURE 20–1. Benign paroxysmal positional vertigo is diagnosed by observing a characteristic nystagmus when performing the Dix-Hallpike test. For testing the right posterior semicircular canal, the patient sits on the examination table and turns his/her head to the right 45 degrees. The patient is moved by the examiner from the seated to supine position with the head slightly hanging over the edge of the table. The eyes are observed for the characteristic down-beating, rotatory nystagmus. The nystagmus has a 1- to 2-second latency and is fatigable. Reproduced with permission from Shepard NT, Telian SA. Practical management of the balance disorder patient. San Diego: Singular; 1996.

right side or stands behind the patient. The patient is then moved by the examiner from the seated to the supine position with the head slightly hanging over the edge of the table. The right ear is down, and the chin is pointing slightly up. The eyes are observed for the characteristic nystagmus. There is a latency of 1 to 2 seconds prior to the onset of the nystagmus and vertigo. The nystagmus is mixed with a torsional and vertical component and is geotropic (downbeat, rotatory nystagmus). The nystagmus follows Ewald's law for excitation of the dependent PSCC. The nystagmus is in the plane of the canal and the fast phase is toward the stimulated canal. The vertigo and nystagmus increase and then decrease within 20 seconds. The vertigo and nystagmus are reduced with repeated Dix-Hallpike tests, so the nystagmus is fatigable. All of these criteria need to be present to diagnose a patient with BPPV caused by debris in the PSCC.

Imaging. Imaging is reserved for patients who do not have the characteristic nystagmus, have associated neurologic findings, or do not respond to treatment. The imaging choice is a magnetic resonance imaging (MRI) scan with gadolinium contrast to evaluate the brainstem, CPA, and IAC. The MRI is the most sensitive and specific test to identify posterior fossa tumors.

Management *Repositioning Maneuver.* The primary management for BPPV includes maneuvers to reposition the debris into the utricle. The most widely used set of maneuvers was described by Epley[18] (Figure 20–2). The patient is taken through four moves starting in the sitting position with head turned at a 45-degree angle toward the affected side. First, the patient is placed into the Dix-Hallpike position (supine with the affected ear down) until the vertigo and nystagmus subside. Second, the patient's head is turned to the opposite side, causing the affected ear to be up and the unaffected ear to be down. Third, the whole body and head are then turned away from the affected side to a lateral decubitus position with the head in a face down position. The fourth step is to bring the patient back to a sitting position with the head turned toward the unaffected side. The maneuver may be repeated if the patient is still symptomatic. A bone vibrator may also be placed on the mastoid bone during the maneuvers to loosen the debris. Eighty percent of patients will be cured by a single repositioning maneuver. The maneuver may be repeated in resistant patients or patients with recurrences. Other maneuvers and exercises are available for resistant patients.[19]

Surgical Management. For a very small number of patients with intractable BPPV, surgical treatment is available. These patients have failed repositioning maneuvers and have no intracranial pathology on imaging studies. The primary surgical option is PSCC occlusion.[20] A standard mastoidectomy is performed, and the PSCC is fenestrated. The membranous duct is occluded with muscle, fascia, or bone

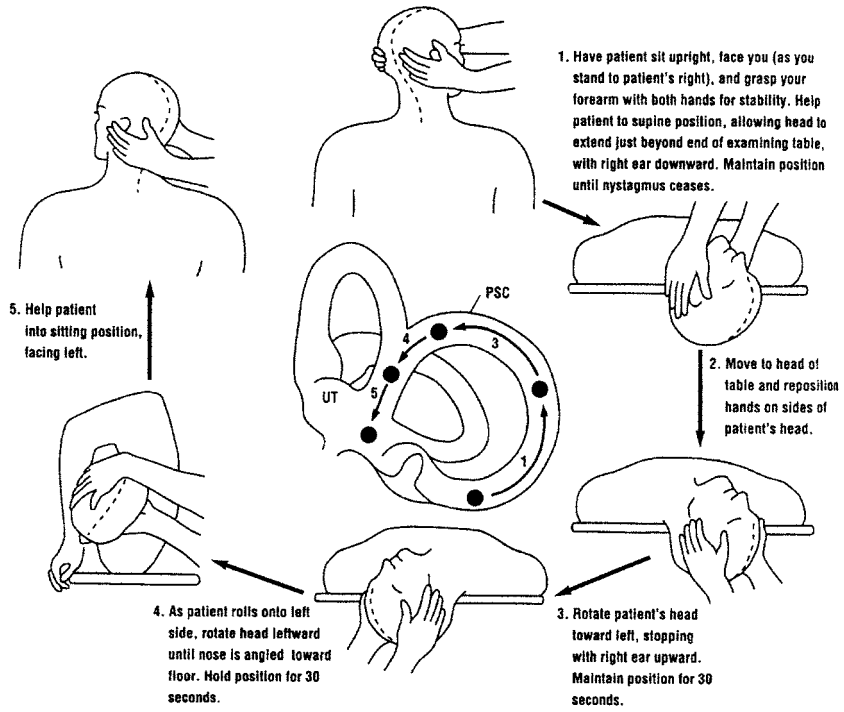


FIGURE 20–2. The primary management for benign paroxysmal positional vertigo includes maneuvers to reposition the debris in the posterior semicircular canal (PSC) into the utricle (UT). The Epley maneuver is shown in this diagram. Eighty percent of patients will be cured by a single repositioning maneuver. The maneuver may be repeated if the patient is still symptomatic. A bone vibrator may also be placed on the mastoid bone during the maneuvers to loosen the debris. Reproduced with permission from Foster CA, Baloh RW. Episodic vertigo. In: Rakel RE, editor. Conn's current therapy. 47th ed. Philadelphia: WB Saunders; 1995. p. 837–41.

pate or collapsed with a laser. The occlusion prevents debris and subsequent endolymph movement to deflect the cupula. There may be a temporary mixed hearing loss that usually recovers. The success rate for PSCC occlusion is high. A more technically challenging surgical option with increased risk to hearing involves ablating the nerve supply of the PSCC via a singular neurectomy.[21]

Meniere's Disease

History and Pathogenesis Meniere's disease or endolymphatic hydrops is an idiopathic inner ear disorder characterized by attacks of vertigo, fluctuating hearing loss, tinnitus, and aural fullness.[22] The history of Meniere's disease parallels our understanding of vestibular anatomy and physiology and provides a rationale for treatment of Meniere's disease and other vestibular disorders. The recognition that hearing loss originated in the inner ear has been long-standing. However, the recognition that vertigo was linked to the inner ear rather than only to central sources was first clearly described by Prosper Meniere in 1861.[23] In the decade following Meniere's description, the function and microscopic anatomy of the vestibular system were further refined. In 1869, Boettcher described the normal microscopic appearance of the utricle, saccule, and endolymphatic sac.[23] In 1870, Friedrich Goltz clarified the function of the vestibular system by concluding that the semicircular canals were responsible for mediating equilibrium only and were not involved with hearing. The next step in understanding the pathogenesis of Meniere's disease was provided by Knapp in 1871. He believed that Meniere's disease was caused by elevated intracochlear (endolymphatic) pressure. The clinician's ability to evaluate vertigo and nystagmus with caloric, rotational, and galvanic testing and to differentiate and categorize peripheral and central vestibular disorders was developed and introduced by Robert Bárány in the late nineteenth and early twentieth century.

The treatment strategies for Meniere's disease were developed based on the mentioned advances in vestibular anatomy and physiology. In 1902, Parry performed a CN VIII division for vertigo in a patient with suspected Meniere's disease.[23] In 1904, Milligan and Lake performed a labyrinthectomy for vertigo, providing a safer alternative than CN VIII section.[23] The effectiveness of CN VIII section for Meniere's disease was shown by Dandy in 1928.[23] These procedures caused hearing loss, so hearing preservation treatments were developed.[23] In 1931, McKenzie performed a selective vestibular neurectomy.[24] The the-

ory of increased endolymphatic pressure prompted Portman to open the endolymphatic sac (ES) via a transmastoid approach in 1926.[23] The patients had improvement in vertigo and hearing preservation. The medical management proposed by Frustenberg in 1934 to treat the increased endolymphatic pressure included a low-sodium diet and ammonia salts (acting as diuretics).[23]

In 1938, Hallpike and Cairns made the most significant advance in understanding the pathogenesis of Meniere's disease by describing the histopathology of Meniere's disease.[25] They noted gross dilation of the saccule and scala media with obliteration of the perilymph spaces of the vestibule and scala vestibuli and confirmed the concept of endolymphatic hydrops. Their descriptions caused the treatment strategies to focus on the labyrinth. In the 1940s and 1950s, Cawthorne popularized the transmastoid labyrinthectomy as the standard of care for Meniere's disease.[23] In 1948, Fowler used streptomycin, an aminoglycoside, to perform a chemical labyrinthectomy.[23] The development of the operative microscope by Nylén allowed House to perform the first microsurgical vestibular nerve sections in 1960.[26] The microsurgical approach improved preservation of hearing and maintenance of facial nerve function. The 1990s have seen a resurgence in the use of intratympanic gentamicin (a vestibular selective aminoglycoside) for the treatment of Meniere's disease.[23,27]

The cause of the endolymphatic hydrops seen by Hallpike and Crains in 1938 continues to elude us. The cause of Meniere's disease has been attributed to anatomic, infectious, immunologic, and allergic factors. The focus of most studies has been on the endolymphatic duct and sac, the basic premise being that there is increased endolymphatic fluid owing to impaired reabsorption of the endolymphatic fluid in the endolymphatic duct and sac. A histopathologic study has shown blockage in the longitudinal flow of endolymph in the endolymphatic duct, ES, utricular duct, saccular duct, and ductus reuniens.[28] Hebbar and colleagues have reported that the ES in patients with Meniere's disease is smaller, has less absorptive tubular epithelium, and has increased perisaccular fibrosis.[29] A blinded, controlled study, however, did not show any difference in the connective tissue or fibrosis surrounding the ES in Meniere's disease.[30] The vestibular duct has also been found to be smaller in Meniere's disease. In recent studies, a decrease in type II vestibular hair cells has been demonstrated in Meniere's disease. The role and significance of the decrease of these type II hair cells are currently not known.[31] The ES has been shown to be important in inner ear metabolic homeostasis. The ES secretes glycoprotein conjugates in response to osmotic challenges, and preliminary studies have shown alteration in glycoprotein metabolism in Meniere's disease. There has been no conclusive proof of an infectious agent related to Meniere's disease.[30,32]

The role of allergy and immune mechanisms in Meniere's disease is under investigation. The "seat" of immunity in the inner ear may be the ES. The ES is able to process antigens and mount a local antibody response. The ES may be vulnerable to immune injury because of the hyperosmolarity of its contents and the fenestrations in its vasculature. These two properties increase the risk of immune complex deposition and injury. Immunoglobulin (Ig) G deposition is seen in the ESs of patients undergoing ES shunt procedures. Meniere's patients also have elevated IgM complexes and C1q component of complement and low levels of IgA complexes in their serum. Patients with Meniere's disease have also shown vulnerability to autoimmune (cytotoxic) reactions. Thirty percent of patients with Meniere's disease had autoantibodies to an inner ear antigen by Western blot analysis. The response of some patients with Meniere's disease to corticosteroid therapy and the increased rate of expression of certain human leukocyte antigens (A3, Cw7, B7, and DR2) in Meniere's disease support the presence of an underlying immune mechanism.[33] A similar argument may be made regarding Meniere's disease and allergy. A significant percentage (50%) of patients with Meniere's disease have concomitant inhalant and/or food allergy, and treating these allergies with immunotherapy and diet modification has improved the manifestations of their allergy and their Meniere's disease. The fenestrated blood vessels of the ES may be vulnerable to vasoactive mediators, such as histamine, that are released during an IgE-mediated allergic reaction. Some have suggested a synergistic role of allergy or viral infection in potentiating the immune abnormalities in Meniere's disease.[34]

Epidemiology The incidence of Meniere's disease ranges from 10 to 150 cases in 100,000 persons per

year. There is no gender bias, and patients typically present in the fifth decade of life. A new diagnosis of Meniere's disease in someone less than 20 or over 70 years is unusual. There is no right or left ear predilection for the disease.[35]

Natural History Meniere's disease is characterized by remissions and exacerbations, making it difficult to predict its future behavior in any individual patient based on the patient's own history, diagnostic evaluations, or epidemiologic profiles. The initial manifestation may be vertigo or hearing loss, but within 1 year of onset, the typical syndrome (attacks of vertigo, tinnitus, fluctuating hearing loss, and aural fullness) is present. Longitudinal studies have shown that after 10 to 20 years, the vertigo attacks subside in most patients and the hearing loss stabilizes to a moderate-to-severe level (50 dB HL). Meniere's disease usually affects one ear initially, but the risk of developing Meniere's disease in the other ear appears to be linear with time. Twenty-five to 45% of patients may develop disease in the second ear.[35,36]

Symptoms and Signs Meniere's disease occurs as episodic attacks lasting for hours. The four symptoms and signs include unilateral, fluctuating sensorineural hearing loss; vertigo lasting minutes to hours; constant or intermittent tinnitus typically increasing in intensity before or during the vertiginous attack; and aural fullness. The acute attack is also associated with nausea and vomiting, and following the acute attack, patients feel exhausted for a few days. Table 20–3 shows the diagnostic scale for Meniere's disease created by the Committee on Hearing and Equilibrium of the American Academy of Otolaryngology-Head and Neck Surgery.[22] As emphasized in the diagnostic scale, the diagnosis of Meniere's disease is based on the longitudinal course of the disease rather than on a single attack.

Diagnostic Evaluation Meniere's disease is a clinical diagnosis. The diagnostic evaluation of Meniere's disease primarily includes audiometry and a fluorescent treponemal antibody absorption to rule out syphilis. Electrophysiologic studies, other serologic studies, and imaging are obtained as needed. The role of allergy testing continues to be defined. There is no diagnostic test for Meniere's disease.[37]

TABLE 20–3. Diagnostic Scale for Meniere's Disease of the American Academy of Otolaryngology-Head and Neck Surgery*

Certain Meniere's disease
Definitive Meniere's disease, plus histopathologic confirmation
Definitive Meniere's disease
Two or more episodes of vertigo of at least 20 min
Audiometrically documented hearing loss on at least one occasion
Tinnitus and aural fullness
Probable Meniere's disease
One definite episode of vertigo
Audiometrically documented hearing loss on at least one occasion
Tinnitus and aural fullness
Possible Meniere's disease
Episodic vertigo without documented hearing loss
Sensorineural hearing loss, fluctuating or fixed, with dysequilibrium but without definitive episodes

*In all scales, other causes must be excluded using any technical methods (eg, imaging, laboratory, etc). Adapted from Committee on Hearing and Equilibrium.[22]

Audiology. Audiologic assessment initially shows a low-frequency or low- and high-frequency (inverted V) sensorineural hearing loss. As the disease progresses, there is a flat sensorineural hearing loss.[36] A glycerol dehydration test involves measuring serial pure-tone thresholds and discrimination scores during diuresis.[38] The diagnosis of Meniere's disease is supported if there is improvement in the patient's hearing. The test is positive in only 50% of patients suspected of having Meniere's disease.

Electrocochleography. Electrocochleography measures the sound-evoked electrical potentials from the inner ear. The three phenomena measured from the external auditory canal, tympanic membrane, or promontory in response to clicks include the cochlear microphonic, summating potential, and action potential. The endolymphatic hydrops of Meniere's disease causes a larger summating potential, so the ratio of the summating potential to the

action potential (SP/AP) is elevated.[39] The SP/AP lacks the specificity or sensitivity to use it to diagnose Meniere's disease consistently or to predict its clinical course.[37]

Vestibular Testing. Vestibular testing (electronystagmography [ENG] with caloric testing) shows peripheral vestibular dysfunction. The caloric response decreases during the first decade of the disease and usually stabilizes at 50% of normal function.[40]

Serology. Fluorescent treponemal antibody absorption is mandatory in any patient given the diagnosis of an idiopathic disease since syphilis may perfectly imitate Meniere's disease. Autoimmune inner ear disease may present initially with a Meniere's picture. The distinguishing characteristics of an autoimmune inner ear disease include a more aggressive course and early bilateral involvement. Autoimmune serologic tests may also be helpful. A promising test is a Western blot looking for anticochlear antibodies that can bind to a 68 to 70 kD antigen.[41]

Imaging. Magnetic resonance imaging with gadolinium contrast allows the exclusion of retrocochlear pathology, such as a VS. An imaging scan is not mandatory with a classic course of Meniere's disease leading to a clinical diagnosis. Imaging should be obtained if the initial presentation or course is unusual or if surgical management is planned.

Management In considering the management of Meniere's disease, an observation by Dandy rings true: "The syndrome is one that lends itself well for a time to statistical conclusions because of the marked variations in the frequency of attacks. Some patients will go for months or even years between attacks and what is done last gets the credit for the free interval…"[24] Since the introduction of aminoglycoside therapy in 1948, no significant conceptual advances have been made in the treatment of Meniere's disease. The current treatments focus on relieving vertigo without further injuring the hearing. Hearing may be temporarily improved or stabilized by the current treatments, but the hearing does not have long-term stability.

Medical Treatment. The primary management of Meniere's disease involves a low-sodium diet (1,500 mg/d) and diuretics (hydrochlorothiazide).[42] Van Deelen and Huizing, in a hydrochlorothiazide, double-masked, placebo-controlled, crossover study, showed that diuretics seem to improve the vestibular complaints but have no effect on hearing or tinnitus.[43] Some patients benefit from dietary restrictions on caffeine, nicotine, alcohol, and theophylline-containing foods (chocolate). Acute attacks are managed with vestibular suppressants (meclizine, diazepam) and antiemetic medications (prochlorperazine suppository). The majority of patients are controlled with conservative management. Medically refractory patients with serviceable hearing may undergo intratympanic gentamicin therapy, ES surgery, or vestibular nerve section. Patients without serviceable hearing may undergo intratympanic gentamicin therapy or transmastoid labyrinthectomy.

Aminoglycoside. Intratympanic gentamicin was introduced by Schuknecht in 1957 and has seen a resurgence in the 1990s by the work of Nedzelski and others.[27] Intratympanic gentamicin is absorbed into the inner ear primarily via the round window and selectively damages the vestibular hair cells relative to the cochlear hair cells. Gentamicin may also decrease endolymph production by affecting dark cells in the stria vascularis.[44] Intratympanic gentamicin has nearly a 90% vertigo control rate with a follow-up of at least 2 years, and the extent of hearing loss depends on the protocol for gentamicin delivery. A variety of treatment protocols (daily, biweekly, weekly, monthly injections) using fixed-dose or titration end-point regimens exist, but few trends are present. Treatments are stopped if there is persistent hearing loss. Vertigo control is nearly always obtained if vestibular function is ablated. However, the risk of hearing loss increases as the total dose and frequency of gentamicin injections are increased. Current protocols are reducing the dose and frequency of injections to decrease hearing loss and still obtain vertigo control. Vertigo control may be obtained with some residual vestibular function present, and this residual function may be useful if patients develop bilateral Meniere's disease. Recent studies with weekly or monthly injections have shown nearly 90% vertigo control with 7 to 17% < 30 dB hearing loss and less than 1% ≥ 30 dB hearing loss.[27,44,45]

Surgery. Patients who have failed medical and gentamicin treatment may require surgical intervention.

Endolymphatic sac surgery and vestibular nerve sections preserve hearing, whereas labyrinthectomy ablates hearing.

Endolymphatic Sac. Endolymphatic sac surgery involves a mastoidectomy and locating the ES on the posterior fossa dura (Figure 20–3, A). The sac is medial to the sigmoid sinus and inferior to the PSCC. The ES is also located along an imaginary line (Donaldson's line) in the plane of the horizontal semicircular canal. The ES may be decompressed or have a shunt placed that communicates into the subarachnoid space or mastoid cavity. The "sham" study, a double-masked, placebo-controlled comparison of the endolymphatic mastoid shunt versus a cortical mastoidectomy, by Thomsen et al showed that ES surgery provides no greater benefit than performing a cortical mastoidectomy.[46] A 9-year follow-up showed 70% control of vertigo in both surgical groups.[47] Reanalysis of the sham study suggested a greater benefit in the ES surgery group, and a recent study with a 5-year follow-up showed an 88% functional level 1 or 2 response after an endolymphatic mastoid shunt operation.[48] The success rates have great variation in the literature, so the benefit of ES surgery is not universally accepted. Endolymphatic shunt surgery provides a nondestructive option for patients who fail medical or aminoglycoside therapy and have good hearing. The role of endolymphatic surgery is currently in decline with the renewed interest in intratympanic aminoglycoside therapy.

Vestibular Nerve Section. Vestibular nerve section provides a definitive treatment of unilateral Meniere's disease in patients with serviceable hearing. Ninety-five percent of patients achieve vertigo control, and hearing is preserved in over 95% of patients.[37] The vestibular neurectomy may be approached via a retrosigmoid or middle fossa approach (Figure 20–3, B). These procedures are described in detail in the CPA section. The risk to the facial nerve is less than 1% in the retrosigmoid approach and less than 5% in the middle fossa approach. The advantage of the middle fossa approach is sectioning the vestibular nerves in the IAC lateral to where they join the cochlear nerve in the CPA.[49] The patients are acutely vertiginous and have nystagmus (fast phase away from the operated ear) for a few days until central compensation takes effect. The hearing appears to degenerate in the postoperative patient in accordance with the natural history of Meniere's disease.

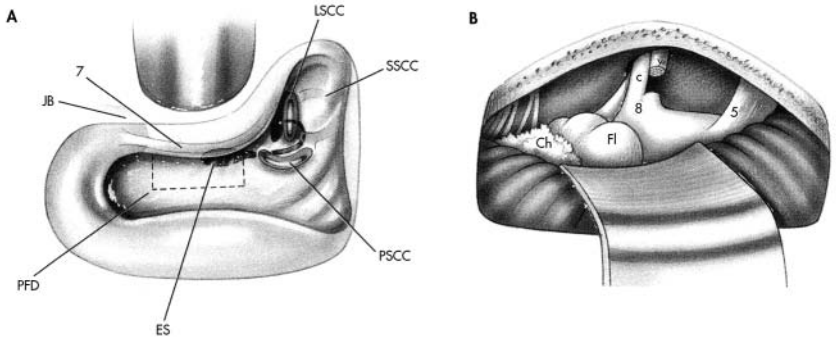


FIGURE 20–3. Two hearing-preserving operations for Meniere's disease include endolymphatic sac (ES) surgery and vestibular nerve section. *A*, Endolymphatic sac surgery (left ear) involves a mastoidectomy and locating the ES on the posterior fossa dura (PFD). The sac is located posterior to the mastoid segment of the facial nerve (7), inferior to the posterior semicircular canal (PSCC), and superior to the jugular bulb (JB). The ES may be decompressed or have a shunt placed that communicates into the mastoid cavity. Endolymphatic shunt surgery provides a nondestructive option for patients who fail medical or aminoglycoside therapy and have good hearing. LSCC and SSCC = lateral and superior semicircular canals. *B*, Vestibular nerve section provides a definitive treatment of vertigo in patients with serviceable hearing. The vestibular neurectomy may be approached via a retrosigmoid or middle fossa approach. The artwork shows a completed left retrosigmoid vestibular neurectomy with sectioning of the vestibular nerve (v) and preservation of the cochlear nerve (c). 8 = vestibulocochlear nerve near the pons; 5 = trigeminal nerve; Fl = flocculus; Ch = choroid plexus. Reproduced with permission from Jackler RK.[50]

Labyrinthectomy. A transmastoid labyrinthectomy with fenestration of the bony semicircular canals and vestibule and removal of the membranous neuroepithelium provides control of vertigo in nearly all patients with unilateral Meniere's disease and poor hearing. The rate of control may decline by 10 years owing to development of vertebral basilar artery insufficiency (aging), poorer vision, and development of Meniere's disease in the contralateral ear. The complete loss of unilateral vestibular function owing to the labyrinthectomy leads to unsteadiness in up to 30% of patients.[51]

Vestibular Neuronitis

A case of vestibular neuronitis was described in 1909 by Ruttin, and the term vestibular neuronitis was coined by Hallpike in 1949.[52] Vestibular neuronitis is the third most common cause of peripheral vestibular vertigo following BPPV and Meniere's disease.[53] The presentation includes acute vertigo. Like Meniere's disease, the pathogenesis is not known, but the majority of patients recover, with no sequelae. The primary role of the physician is to rule out a central cause of the acute vertigo. The treatment is primarily supportive care.

Epidemiology Vestibular neuronitis has no gender bias and typically affects middle-aged people. Less than half of the patients will have an antecedent or concurrent viral illness.[52]

Pathogenesis The proposed causes of vestibular neuronitis include viral infection, vascular occlusion, and immune mechanisms. The evidence to support a viral cause is limited by lack of pathologic tissue to study, and the evidence based on seroconversion remains unconvincing. Therefore, support for the viral cause remains rare and circumstantial. The study of the available temporal bones of patients with vestibular neuronitis shows a spectrum of injury. There may be no abnormality seen to significant degenerative changes in the vestibular nerve, Scarpa's ganglia, and vestibular neuroepithelium. The injury is often seen in the superior vestibular nerve. In terms of a vascular cause, there was no evidence of vascular occlusion affecting the vestibular system in any of the temporal bones studied.[52,54]

Natural History The natural history of vestibular neuronitis includes an acute attack of vertigo that lasts a few days with complete or at least partial recovery within a few weeks to months. Some patients (15% in one study) may have significant vestibular symptoms even after 1 year.[55] Recurrent attacks in the same or contralateral ear have been reported but are unusual. Some patients may later develop BPPV. The majority of patients have complete recovery. The few patients who have persistent vestibular symptoms should have vestibular rehabilitation therapy.

Symptoms and Signs The presentation of vestibular neuronitis includes the sudden onset of vertigo with nausea and vomiting. The patient has normal hearing and a normal neurologic examination. The patient may have postural instability toward the injured ear but is still able to walk without falling. The patient usually does not have a headache. The patient has spontaneous nystagmus characteristic of an acute peripheral vestibular injury. The nystagmus is usually horizontal with a torsional component and is suppressed by visual fixation. The reduction in vestibular signal in the injured ear leads to relative vestibular excitation in the opposite ear. The result is that the slow phase of nystagmus is toward the injured ear and the fast phase is away from the injured ear. The nystagmus is intensified by looking toward the fast phase and is decreased by looking toward the slow phase or toward the injured ear. This principle is Alexander's law. The direction of the nystagmus does not change with changes in the direction of gaze.[56]

Diagnostic Evaluation The diagnosis of vestibular neuronitis is based on the constellation of findings described above. The need for further evaluation is necessary only if there is a concern for a central cause of the acute vertigo or the acute vertigo does not substantially improve in 48 hours.[56] The primary central cause for acute vertigo lasting days is a brainstem or cerebellar stroke. In the majority of cases, there will be other neurologic findings: diplopia, dysmetria, dysarthria, motor and sensory deficits, abnormal reflexes, inability to walk without falling, and central nystagmus. Central nystagmus is not affected by visual fixation and may change directions with changes in gaze. Purely vertical or purely torsional nystagmus is highly suggestive of a central disorder.[57] In the event of an isolated inferior cerebellar stroke, the presentation may be indistinguishable from

vestibular neuronitis. Twenty-five percent of patients with risk factors for stroke who present with vertigo, nystagmus, and postural instability have had an inferior cerebellar stroke.[58] Therefore, patients with significant risk factors for stroke should have an imaging study if they present with acute vertigo, nystagmus, and postural instability.

Vestibular Testing. Vestibular testing in vestibular neuronitis shows an absent or reduced caloric response in the injured ear in the majority of patients. The caloric responses eventually become normal in 42% of patients.[59]

Imaging. Magnetic resonance imaging with emphasis on identification of both infarction and hemorrhage in the brainstem and cerebellum is obtained in patients with risk factors for stroke, patients with additional neurologic abnormalities, and patients who do not show improvement within 48 hours. The other option is a computed tomographic (CT) scan with thin cuts evaluating the brainstem, cerebellum, and fourth ventricle.

Management *Supportive Care.* The primary management includes symptomatic and supportive care during the acute phase of the illness. Patients are given vestibular suppressants and antiemetics to control the vertigo, nausea, and vomiting. These medications are withdrawn as soon as possible to not interfere with the central vestibular compensation.[8]

Cerebellopontine Angle Tumors

The CPA consists of a potential cerebrospinal fluid (CSF)-filled space in the posterior cranial fossa bounded by the temporal bone, cerebellum, and brainstem. The CPA is traversed by CNs V to XI, and the facial (VII) and vestibulocochlear (VIII) nerves are most prominent. Cerebellopontine angle tumors account for 10% of all intracranial tumors (Table 20–4). Nearly 90% of all CPA tumors are VSs (acoustic neuromas) and meningiomas.[60] Other CPA lesions include congenital rest lesions (epidermoids, arachnoid cysts, lipomas), schwannomas of other CNs, intra-axial tumors, metastases, vascular lesions (paragangliomas, hemangiomas), and lesions extending from the skull base (cholesterol granulomas, chordomas).[61] Cerebellopontine angle lesions become clinically symptomatic by causing compression of the neurovascular structures in and around the CPA. The classic description of these symptoms by Cushing includes initially unilateral hearing loss, vertigo, nystagmus, altered facial sensation, facial pain that later progresses to facial palsy, vocal cord palsy, dysphagia, diplopia, respiratory compromise, and death (Table 20–5). This description highlights the untreated natural history of CPA tumors.[62]

TABLE 20–4. Cerebellopontine Angle (CPA) Lesions

Common CPA lesions
Schwanommas (cranial nerves VIII, VII, V)
Meningiomas
Epidermoids
Congenital rest lesions
Epidermoids
Arachnoid cysts
Lipomas
Vascular lesions
Hemangiomas
Paragangliomas (glomus jugulare tumors)
Aneurysms
Hemangioblastomas
Intra-axial tumors
Medulloblastomas
Astrocytomas
Gliomas
Fourth ventricle tumors
Hemangioblastomas
Lesions extending from the skull base
Cholesterol granulomas
Glomus complex tumors
Chordomas
Chondrosarcomas
Metastases

The treatment of CPA tumors includes observation, surgical removal, and irradiation. The key clinical developments in the management of CPA

TABLE 20–5. Cerebellopontine Angle Syndrome

Unilateral hearing loss
Tinnitus
Vertigo
Hypesthesia and neuralgia
Nystagmus
Facial palsy
Vocal cord palsy
Dysphagia
Diplopia
Respiratory compromise
Death

tumors include placing an emphasis on earlier diagnosis by Cushing in 1917 and a dogma change in 1925 by Dandy from rapid tumor enucleation to meticulous, hemostatic total tumor removal.[63] Fundamental technical advances allowing diagnosis and surgical removal include the advent of modern imaging and the introduction of the surgical microscope by House and Doyle in 1961. House helped create the field of neurotology and brought the surgical management of CPA tumors into the modern era.[23,24]

The focus of this section includes understanding the anatomy of the CPA; describing the pathogenesis of VSs, meningiomas, and epidermoids; and providing a clinical context to diagnose, counsel, and treat patients with various CPA tumors.

Anatomy The CPA is roughly triangular shaped in the axial plane and is filled with CSF (Figure 20–4). The superior boundary is the tentorium, and the inferior boundary is the cerebellar tonsil and medullary olives. The anterior border is the posterior dural surface of the petrous bone and clivus, and the posterior border is the ventral surface of the pons and cerebellum. The medial border is the cisterns of the pons and medulla, and the apex is the region of the lateral recess of the fourth ventricle. The lateral opening of the fourth ventricle, the foramen of Luschka, opens into the CPA. Cranial nerves V to XI traverse the cephalic and caudal extent of the CPA. The central structures crossing the CPA to and from the IAC are the facial (CN VII) and vestibulocochlear (CN VIII) nerves, respectively.[64]

Cranial nerves VII and VIII are covered with central myelin provided by neuroglial cells as they cross the CPA and carry a sleeve of posterior fossa dura into the IAC. The transition to peripheral myelin made by Schwann cells occurs at the medial opening of the IAC. The vestibulocochlear nerve divides into three nerves: the cochlear nerve and the superior and inferior vestibular nerves in the lateral extent of the CPA or medial part of the IAC. The IAC is divided into four quadrants by a vertical crest, called Bill's bar, and a transverse crest. The CN VII comes to lie in the anterosuperior quadrant and is anterior to the superior vestibular nerve and superior to the cochlear nerve, whereas the inferior vestibular nerve lies in the posteroinferior quadrant and is inferior to the superior vestibular nerve and posterior to the cochlear nerve (see Figure 20–4). The anterior inferior cerebellar artery (AICA) is the main artery in the CPA and is the source of the labyrinthine artery. The labyrinthine artery courses via the IAC and is an end artery for the hearing and balance organs. The AICA has a variable relationship to CN VII and VIII and to the IAC.[64]

Common Cerebellopontine Angle Lesions The three most common tumors of the CPA are schwannomas, meningiomas, and epidermoids. Each of these tumors has a similar clinical presentation, and they are primarily differentiated by their imaging characteristics. The treatment for these three lesions is surgical resection without injuring the neurovascular structures of the CPA.

Vestibular Schwannomas. Vestibular schwannomas are nerve sheath tumors of the superior and inferior vestibular nerves.[65] They arise in the medial part of the IAC or lateral part of the CPA and cause clinical symptoms by displacing, distorting, or compressing adjacent structures in the CPA.

Epidemiology. Vestibular schwannomas, incorrectly called acoustic neuromas, are by far the most common tumor involving the CPA.[2] Vestibular schwannomas make up 80% of CPA tumors and 8% of all intracranial tumors. Various epidemiologic studies

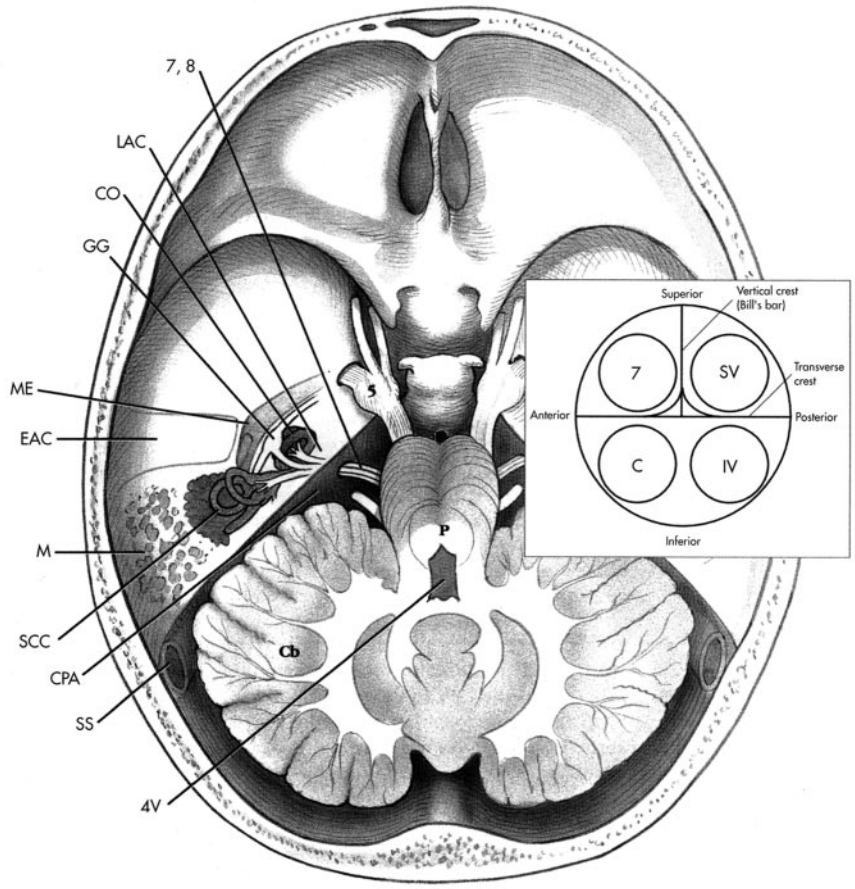


FIGURE 20–4. In the axial plane, the cerebellopontine angle (CPA) has a triangular shape. The CPA is bounded by the cerebellum (Cb), petrous bone, and brainstem. Cranial nerves V (5) to XI traverse the cephalic and caudal extent of the CPA. The central structures crossing the CPA to and from the internal auditory canal (IAC) are the facial (7) and vestibulocochlear nerves (8), respectively. The vestibulocochlear nerve divides into three nerves: the cochlear nerve (C) and the superior (SV) and inferior vestibular (IV) nerves in the lateral extent of the CPA or medial part of the IAC. The *inset diagram* shows a cross-section of the IAC. The IAC is divided into four quadrants by a vertical crest, called Bill's bar, and a transverse crest. The facial nerve comes to lie in the anterosuperior quadrant and is anterior to the superior vestibular nerve and superior to the cochlear nerve, whereas the inferior vestibular nerve lies in the posteroinferior quadrant and is inferior to the superior vestibular nerve and posterior to the cochlear nerve. CO = cochlea; GG = geniculate ganglion; ME = middle ear; EAC = external auditory canal; M = mastoid; SCC = semicircular canals; SS = sigmoid sinus; 4V = fourth ventricle; P = pons. Reproduced with permission from Jackler RK.[50]

have shown an incidence of 10 in 1 million individuals per year.[66] This figure correlates with the 2,000 to 3,000 people diagnosed with VSs each year in the United States. There is no gender bias. The age of presentation is 40 to 60 years of age. Ninety-five percent of VSs occur in a sporadic fashion. The remaining 5% of patients have neurofibromatosis 2 (NF2) or familial VSs. The age of presentation is earlier in nonsporadic VSs, and patients usually present in the second or third decade of life.

PATHOGENESIS. Vestibular schwannomas originate from the Schwann cells of the superior or inferior vestibular nerves at the transition zone of the peripheral and central myelin. This transition zone occurs in the lateral part of the CPA or medial part of the IAC. Therefore, VSs most often arise in the IAC and occasionally arise in the CPA. Schwannomas rarely arise from the cochlear nerve and are rarely malignant. The propensity to develop from the vestibular nerves may be owing to the vestibular

ganglia in the IAC having the highest concentration of Schwann cells.[67]

Recent studies have improved our molecular understanding of VSs. Vestibular schwannomas occur owing to mutations in the gene for the tumor suppressor protein merlin, located on chromosome 22q12.[68,69] Merlin is a cytoskeletal protein and may have a role in cell-cell contact inhibition.[70] The formation of VSs requires mutations of both copies of the merlin gene. One functioning merlin gene will prevent the formation of VSs. Somatic mutations in both copies of the merlin gene will result in sporadic VSs. The probabilities of two spontaneous, independent mutations at one locus predict a unilateral VS presenting in the fourth to sixth decade of life.[71]

In contrast, familial VSs occurring in NF2 require only one somatic mutation event. In NF2, patients inherit one mutated merlin gene and one normal merlin gene. A mutation in the normal allele leads to bilateral VSs by the age of 20. Therefore, NF2 is autosomal recessive at the gene level since disease expression requires mutations in both alleles of the gene, but the inheritance is autosomal dominant (pseudodominant) since inheritance of one mutated allele often leads to the disease state. Neurofibromatosis 2 is a central form of neurofibromatosis, with patients having central nervous system tumors including schwannomas, meningiomas, and gliomas. The great majority of these patients will develop bilateral VSs. In comparison, patients with NF1 (von Recklinghausen's disease) have intra- and extracranial tumors, and less than 5% of these patients will form unilateral VSs.[72,73] Interestingly, the cranial evaluation of two children recovered from a 4,000-year-old burial site showed widened IACs. These siblings likely had VSs owing to NF2.[74]

Genetic screens for the NF2 mutation have been developed and are the basis for genetic counseling for family members of NF2 patients. The severity of mutation involving the merlin gene in NF2 can predict the severity of disease manifestation.[71] The study of VSs in NF2 has provided insight into not only this disease but also the cause of sporadic VSs and has demonstrated a very interesting genetic inheritance mechanism.

PATHOLOGY. ***Gross.*** Vestibular schwannomas have a smooth surface with a yellow to gray color. The tumor is usually solid, with occasional cystic components, and therefore has a firm to soft texture depending on the solid to cystic components.

Microscopic. Microscopic evaluation shows the surface connective tissue capsule to be only 3 to 5 μm in thickness.[75] The classic histologic findings include areas of densely packed cells with spindle-shaped nuclei and fibrillar cytoplasm intermixed with hypocellular areas containing vacuolated, pleomorphic cells. These dense regions are called Antoni A areas, and the hypocellular regions are called Antoni B areas. Within the Antoni A areas, the palisades of the nuclei are termed Verocay bodies.[76] Histologic sections also show prominent, thick-walled vessels. As stated earlier, the cell of origin is the Schwann cell, so VS sections stain with S-100 immunoperoxidase.

NATURAL HISTORY. The natural history of VSs includes a slow rate of growth in the IAC and then into the cistern of the CPA. Studies show that periods of growth are intermixed with periods of quiescence. The average growth rate is 1.8 mm/year.[77] This slow growth causes progressive and often insidious symptoms and signs as there is displacement, distortion, and compression of the structures first in the IAC and then in the CPA. This slow growth via cellular proliferation provides a predictable progression of symptoms and signs. Occasionally, the tumor may undergo rapid expansion owing to cystic degeneration or hemorrhage into the tumor. A rapid expansion causes rapid movement along the subsequently described phases of VS symptoms and may cause rapid neurologic deterioration.

The initial intracanalicular growth affects the vestibulocochlear nerve in the rigid IAC and causes unilateral hearing loss, tinnitus, and vertigo or dysequilibrium. These three symptoms are the typical presenting complaints of not only patients with VS but also of those with other lesions of the CPA. Interestingly, the motor component of the facial nerve is resistant to injury during this phase of growth, and patients have normal facial function. The tumor then grows into the CPA cistern and grows freely without causing significant new symptoms because structures in the CPA are initially displaced without injury (Figure 20–5, A). As the tumor approaches 3 cm, the tumor abuts on the boundaries of the CPA and results in a new set of symptoms and signs. Compression of the CN V causes corneal and midface numbness or pain. Further distortion of CN

VIII and now CN VII causes further hearing loss and dysequilibrium, as well as facial weakness or spasms. Brainstem distortion leads to narrowing of the fourth ventricle (Figure 20–5, B). Further growth leads to the final clinical spectrum described by Cushing in the CPA syndrome.[62] The patient develops cerebellar signs owing to compression of the flocculus and cerebellar peduncle. The patient also develops obstructive hydrocephalus owing to closure of the fourth ventricle. The increasing intracranial pressure manifests in ocular changes, headache, mental status changes, nausea, and vomiting. If the VS continues to grow without intervention, death occurs owing to respiratory compromise.[78]

Symptoms and Signs. Hearing loss is present in 95% of patients with VSs.[78,79] Conversely, 5% of patients will have normal hearing, so unilateral vestibular or facial complaints without hearing loss do not rule out retrocochlear disease. Of the patients with hearing loss, most will have slowly progressive hearing loss with noise distortion. Twenty percent will have an episode of sudden hearing loss.[80] The improvement of hearing loss with or without treatment does not rule out retrocochlear disease. The level of hearing loss is not a clear predictor of tumor size.

Tinnitus is present in 65% of patients. The tinnitus is most often constant, with a high-buzzing pitch. This symptom is often not reported by patients because of the focus on the accompanying hearing loss. Similarly, because of the central compensation for the slowly evolving vestibular injury, patients tolerate and adapt well to the dysequilibrium. The majority of patients will have self-limiting episodes of vertigo. The dysequilibrium is initially mild and constant and often does not prompt a medical visit. Dysequilibrium is present in 60% of patients.

Facial and trigeminal nerve dysfunction occurs after the auditory and vestibular impairments. Patients usually have midface (V2) numbness and often have an absent corneal reflex. The motor supply of the trigeminal nerve to the muscles of masti-

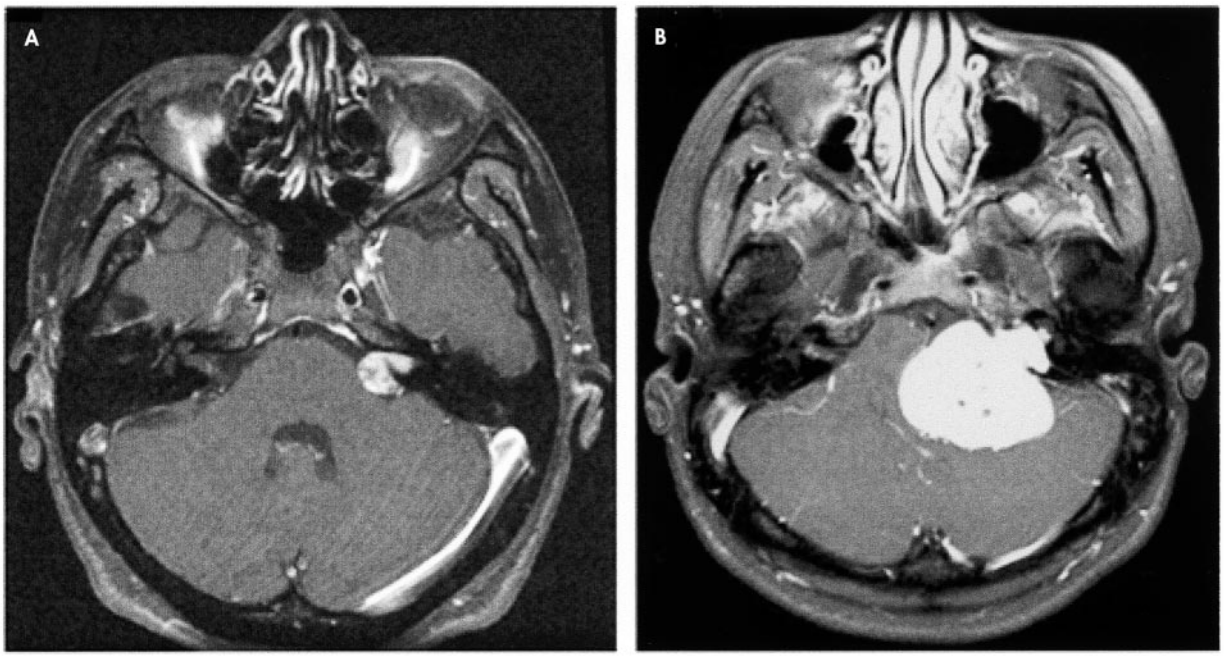


FIGURE 20–5. Axial T_1-weighted magnetic resonance images with gadolinium show the progressive growth pattern of a vestibular schwannoma. Vestibular schwannomas are centered in the internal auditory canal (IAC), expand the IAC, are globular in shape, and enhance with gadolinium. *A*, Vestibular schwannoma involves the IAC and cistern of the cerebellopontine angle. The tumor has solid and cystic components. *B*, Vestibular schwannoma has expanded the IAC and is causing significant compression of the brainstem and cerebellum. The brainstem distortion leads to narrowing of the fourth ventricle. (Courtesy of Nancy J. Fischbein.)

cation is rarely affected. The sensory component of the facial nerve is first affected, which causes numbness of the posterior part of the external auditory canal and is referred to as Hitzelberger's sign. Facial weakness or spasm occurs in 17% of patients and usually leads to a diagnosis of VS within 6 months.[78,79]

With large VSs or tumors that have undergone rapid expansion, patients will have visual complaints of decreased visual acuity and diplopia owing to compromise of CN II, IV, or VI. Hydrocephalus leads to complaints of headache, altered mental status, nausea, and vomiting; on examination, patients have increased intracranial pressure and papilledema. Compression of the lower CNs IX and X causes dysphagia, aspiration, and hoarseness, and examination reveals a poor gag reflex and vocal cord paralysis.

Diagnostic Evaluation. The average patient will require 4 years from the onset of symptoms to diagnosis. The diagnostic dilemma lies in choosing the appropriate patient to pursue audiologic and imaging studies. Although 5% of patients with VSs may present in an atypical fashion, the majority of patients will present with complaints of unilateral hearing loss or hearing distortion, unilateral tinnitus, vertigo or dysequilibrium, or facial numbness, weakness, or spasm. These patients with unilateral auditory, vestibular, and facial complaints are not a diagnostic dilemma and need to undergo careful evaluation to rule out retrocochlear disease. The initial step in the evaluation includes an audiologic assessment. If the audiologic assessment suggests a retrocochlear lesion, then imaging of the CPA is performed to rule out a retrocochlear lesion. Vestibular testing lacks specificity in diagnosing VSs.

Audiology. The standard auditory evaluation should include pure-tone audiometry, speech discrimination scores (SDSs), acoustic reflex thresholds, and acoustic reflex decay. Pure-tone audiometry of patients with VSs shows asymmetric, down-sloping, high-frequency, sensorineural hearing loss in almost 70% of patients. The hearing may also be normal, or the hearing loss may involve only the low frequencies or have a flat, trough, or peak configuration.[81] Retrocochlear hearing loss causes SDSs to be lower than predicted by the pure-tone thresholds. This out-of-proportion depression of SDSs is further accentuated when retested at a higher speech intensity. This phenomenon is called rollover. Poor SDSs are present in about one half of the patients with VSs. An abnormal SDS should trigger an imaging evaluation, but a normal SDS does not rule out a VS. Loss of the acoustic reflex or acoustic reflex decay is noted in most patients with VSs, but normal acoustic reflexes do not preclude a VS. In summary, asymmetric hearing loss, out-of-proportion deterioration in the SDS, and abnormal acoustic reflexes mandate an imaging evaluation to rule out a retrocochlear lesion. The majority of patients with VSs will have one of these abnormalities on audiologic analysis. The minority of the patients with subtle complaints and normal auditory evaluations fall into an area of diagnostic dilemma.[82] The likelihood of having a retrocochlear lesion may be very small, but the patient or physician may feel compelled by the patient's complaints to obtain audiometric testing. There is no clear management algorithm, and the physician must make a decision whether imaging evaluation is needed. Regardless of the decision, the most important part of the diagnostic plan should include patient counseling regarding the future symptoms and signs of CPA tumors and the need for planned follow-up.

Vestibular Testing. Vestibular testing does not provide a sensitive or specific means of diagnosing VSs. The most common test obtained to evaluate vestibular complaints is ENG with caloric testing. An ENG in a patient with a VS will show reduced caloric response in the affected ear. The extent of vestibular function present predicts the amount of postoperative vertigo. The location of the VS on the inferior or superior vestibular nerve may also be predicted by the ENG since the ENG primarily evaluates the lateral semicircular canal. The lateral semicircular canal is innervated by the superior vestibular nerve.[83]

Auditory Brainstem Response. Auditory brainstem response (ABR) is the measured electrical response of the cochlea and its brainstem pathway to short-duration broad-band clicks. The evoked response is a characteristic waveform with five identifiable peaks (I to V). The absolute latency or timing of each wave is recorded. In patients with VSs, the ABR is partially or completely absent, or there is a delay in the latency of wave V on the affected side. The delay may be an absolute delay based on normative data or a delay compared to the latency of wave V on the other side. An interaural delay of wave

V greater than 0.2 ms is considered abnormal. Overall, ABR has a sensitivity of > 90% and a specificity of 90% in detecting VS. However, when only considering intracanalicular tumors, 18 to 33% of tumors will be missed.[84] As the detection limits and costs of imaging studies have improved, the role of ABR in diagnosis of VS has dramatically declined. In the future, there will likely be no clear role for ABR in VS diagnosis.

Imaging. Magnetic resonance imaging with gadolinium contrast provides the gold standard for VS diagnosis or exclusion.[85] The MRI scan also allows for surgical planning. Typically, a series of T_1-weighted images in which CSF is dark and fat is bright, a T_1-weighted image with gadolinium contrast, and a T_2-weighted image in which CSF is bright are obtained. The images are in 2 to 4 mm contiguous or overlapping slices. The various lesions within the CPA may be differentiated based on their varying imaging and enhancing characteristics. The MRI characteristics of a VS include a hypointense globular mass centered over the IAC on a T_1-weighted image with enhancement when gadolinium is added. Vestibular schwannomas are iso- to hypointense on T_2-weighted images[86] (see Figure 20–5). The two detractors to the many advantages of MRI, including lack of radiation exposure, are the time and cost of each scan. These two issues are being addressed by developing imaging sequences that focus on the CPA and do not use contrast administration. One example is a high-resolution fast spin echo T_2-weighted sequence.[87] This scanning strategy uses CSF as the contrast agent. Fast spin echo T_2-weighted scans can be done in 15 to 20 minutes and cost less compared with a standard MRI with intravenous contrast administration. The drawback is a slight loss in sensitivity owing to false-negative readings of very small tumors. The steady improvements in the quality and cost of MRIs have made MRI the preferred diagnostic modality for CPA tumors.

In situations for which MRI cannot be used or is not accessible, a CT scan with iodine contrast or an ABR offers a reasonable alternate screening modality. A CT scan with contrast provides consistent identification of CPA tumors over 1.5 cm or CPA tumors that at least have a 5 mm CPA component. Vestibular schwannomas appear as ovoid masses centered over the IAC with nonhomogenous enhancement. A CT scan with contrast can miss intracanalicular tumors unless there is bony expansion of the IAC.

MANAGEMENT. ***Surgery.*** The primary management for VSs is surgical removal. The roles of observation and radiotherapy are currently for patients who cannot tolerate a surgical procedure or have a life span of less than 5 years. The role of radiotherapy is expanding. The majority of patients will undergo surgical excision because the natural history of VSs is to grow and eventually cause significant compression of the structures surrounding the CPA, as described above. In addition, the ability to accomplish total tumor removal with the least morbidity is generally related to the size of the tumor, so significant surgical delay is not recommended.

Surgical approaches to the CPA include translabyrinthine, retrosigmoid, and middle fossa craniotomies (Figure 20–6, A). The appropriate approach for a particular patient is based on the hearing status, size of the tumor, extent of IAC involvement, and experience of the surgeon (Table 20–6). The approaches are either hearing preserving or ablating. The retrosigmoid and middle fossa approaches are hearing preserving. However, they have limitations of exposure to all aspects of the CPA and IAC. The middle fossa approach is well suited to patients with good hearing and a tumor less than 2 cm. The retrosigmoid approach is well suited to patients with good hearing and a tumor less than 4 cm and not involving the lateral part of the IAC. The lateral part of the IAC is usually only directly accessible via a retrosigmoid approach by removing the PSCC. The violation of the PSCC will lead to hearing loss. The translabyrinthine approach causes total hearing loss and so is appropriate for patients with poor hearing (pure-tone average > 30 dB) or patients with good hearing and tumors not accessible by the hearing-preserving approaches. The surgical technique for these procedures will be briefly described. Three critical issues inherent to all three techniques are the extent of exposure of the IAC and CPA, identification and preservation of the facial nerve, and extent of brain retraction. These operations use electrophysiologic monitoring of CN VII and an ABR in hearing-preserving approaches. The use of high-speed drills, craniotomies, and operating microscopes has brought VS removal into the modern age.[50,88]

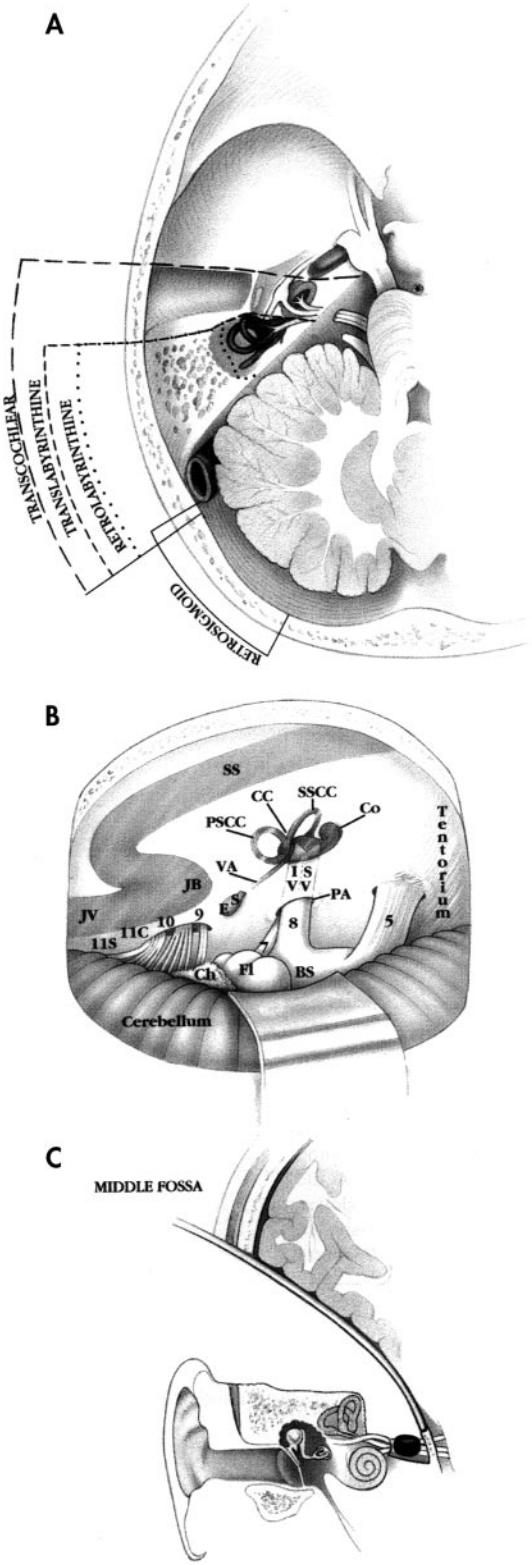


FIGURE 20–6. The surgical approaches to the cerebellopontine angle (CPA) include transcochlear, translabyrinthine, retrolabyrinthine, retrosigmoid, and middle fossa craniotomies. *A*, The axial view of the skull at the level of the internal auditory canal (IAC) and CPA shows the extent of resection needed for these four posterior fossa craniotomies. The main posterior fossa approaches to CPA lesions are the translabyrinthine and retrosigmoid craniotomies. The translabyrinthine approach requires a mastoidectomy, decompression of the sigmoid sinus, and vestibular labyrinthectomy and provides excellent access to the IAC and CPA. However, labyrinthectomy leads to complete hearing loss. *B*, The retrosigmoid craniotomy is positioned posterior to the sigmoid sinus and requires cerebellar retraction. The retrosigmoid approach is a versatile approach with a panoramic view of the CPA from the foramen magnum inferiorly to the tentorium superiorly. The medial two-thirds of the IAC are also accessible without violating the inner ear, therefore preserving hearing. SS = sigmoid sinus; JB = jugular bulb; JV = jugular vein; Co = cochlea; SSCC = superior semicircular canal; PSCC = posterior semicircular canal; CC = common crus; VA = vestibular aqueduct; ES = endolymphatic sac; SV = superior vestibular nerve; IV = inferior vestibular nerve; 5, 7 to 11 = cranial nerves V, VII to XI; S = spinal; C = cranial; BS = brainstem; Fl = flocculus; Ch = choroid plexus. *C*, The middle fossa approach provides access to the IAC by extradural elevation of the temporal lobe. This exposure allows removal of intracanalicular tumors while preserving hearing. An intracanalicular vestibular schwannoma is depicted in the diagram. Reproduced with permission from Jackler RK.[50]

TABLE 20–6. Surgical Approaches to the Cerebellopontine Angle and Their Indications

Hearing preserving
Retrosigmoid: good hearing and tumor not involving the lateral internal auditory canal
Middle fossa: good hearing and tumor less than 2 cm, internal auditory canal tumors
Hearing ablating
Translabyrinthine: poor hearing or larger tumors not accessible by other approaches

Translabyrinthine Approach. The primary approach for removal of VS is the translabyrinthine approach. The boundaries of the approach include the facial nerve and cochlear aqueduct anteriorly, middle fossa dura superiorly, posterior fossa dura posteriorly, and jugular foramen inferiorly (see Figure 20–6, A). These boundaries are approached via the familiar postauricular incision. A complete canal up mastoidectomy is accomplished, with identification of the incus, tegmen, sigmoid sinus, and facial nerve. A compete labyrinthectomy is then performed with medial skeletonization of the middle and posterior fossa dura and decompression of the sigmoid sinus to the jugular foramen. After bony skeletonization of the IAC, the dura of the IAC is opened and the facial nerve is identified medial to the vertical crest (Bill's bar). Once the facial nerve is identified in the fundus or lateral aspect of the IAC, tumor removal occurs from lateral to medial along the IAC. In large tumors, the tumor is debulked internally, and then the tumor capsule is removed from the surrounding structures including the facial nerve. After tumor removal, abdominal fat is placed into the defect.

The three advantages of the translabyrinthine approach are the ability to remove tumors of all sizes, minimal retraction of the brain, and the ability to visualize directly and preserve the facial nerve. The rate of facial nerve preservation is 97%. The rate of CSF leak presenting under the incision or draining through the nose via the eustachian tube is 5%. The majority of these CSF leaks resolve with conservative management, which includes mastoid dressing and fluid restriction.[89] There is a minimal risk of meningitis associated with a CSF leak.

Retrosigmoid Approach. The retrosigmoid approach is a modification of the traditional suboccipital approach used by neurosurgeons to address most posterior fossa lesions. The retrosigmoid approach is a versatile approach with a panoramic view of the CPA from the foramen magnum inferiorly to the tentorium superiorly (Figure 20–6, B). The medial two-thirds of the IAC are also accessible without violating the inner ear; therefore, hearing is preserved.

The surgical technique starts with a curvilinear skin incision 6 cm behind the ear over the retromastoid region. The soft tissue and posterior nuchal musculature are elevated to expose the mastoid and retromastoid bone. A 5 × 5 cm craniotomy is performed with the sigmoid sinus as the anterior boundary and the transverse sinus the superior boundary. Elevation of a bone plate is technically difficult, so the bone may be removed by drilling. The bone fragments are collected and will be replaced during closure. The bone fragments will reform a bone plate and prevent adherence of the musculature to the dura. If decompression of the sigmoid sinus is needed for exposure, a mastoidectomy may also be performed. The dura is then opened along the sigmoid sinus, and the cerebellum is seen. The CSF from the cisterna magnum should be released prior to retracting the cerebellum. Medial retraction of the cerebellum allows visualization of the CPA. To address the IAC component of the tumor, the posterior bone of the IAC should be removed. The bone dust created is carefully confined and removed to prevent meningeal irritation. The extent of IAC skeletonization is limited by the proximity to the inner ear. The endolymphatic duct and sac serve as landmarks to the proximity of the PSCC and allow preservation of the inner ear and hearing. The facial nerve is normally anterior to the tumor, or its position is ascertained with facial nerve monitor-

ing. The tumor removal is as previously described. After tumor removal and hemostasis, air cells along the IAC and mastoid are closed with bone wax or bone cement to eliminate paths for CSF leak. A fat or muscle graft may also be placed into the petrosal defect to prevent CSF leak. The dura mater is closed, and the bone plate or bone pate is replaced. The musculature and soft tissue are meticulously closed.[90]

The primary advantage of the retrosigmoid approach relative to the translabyrinthine approach is the ability to preserve hearing in properly selected tumors. If hearing preservation is not an issue, the retrosigmoid approach allows a versatile approach to the CPA and IAC. The relative disadvantages compared with the translabyrinthine approach include persistent postoperative headache, increased difficulty in resolving CSF leaks, need for cerebellar retraction, and inability to have direct access to the facial nerve. The combination of intradural drilling leading to meningeal irritation by bone dust and dissection of suboccipital musculature causes nearly 10% of patients to have a persistent, severe postoperative headache.[91] In the case of extensive pneumatization of the IAC and mastoid, the air cells may be difficult to seal completely, and the inability to address the aditus ad antrum or the eustachian tube causes CSF leaks to be persistent despite conservative treatment. The extent of cerebellar retraction is minimal in small tumors, but the amount of retraction increases with larger tumors. The surgical control of the facial nerve is adequate in the retrosigmoid approach, but the exposure of the facial nerve is superior in the translabyrinthine approach.

Middle Fossa Approach. The middle fossa approach provides a hearing-preserving approach to intracanalicular tumors with a less than 1 cm cisternal component. The surgical technique involves an inverted U-shaped incision centered over the ear. The temporal muscle is reflected inferiorly to expose the squamous portion of the temporal bone. A 5×5 cm temporal craniotomy is performed and is centered over the zygomatic root. Extradural elevation of the temporal lobe is accomplished to reveal the temporal bone. The greater superficial petrosal nerve leading to the geniculate ganglion reveals the anterior, lateral boundary of the IAC, and the arcuate eminence reveals the posterior boundary of the IAC (Figure 20–6, C). These landmarks may be difficult to identify, and the IAC dura may have to be identified medially by drilling toward the porus acousticus. Once the IAC is identified and well skeletonized medially, the bone removal continues laterally. However, the extent of IAC skeletonization laterally is limited by the basal turn of the cochlea anteriorly and superior semicircular canal posteriorly. The IAC dura is opened posteriorly to avoid injury to the facial nerve. The tumor is dissected free of the facial nerve and removed in a medial to lateral direction. Any air cells are sealed, and the dural defect is covered with a fat or muscle plug. The craniotomy bone flap is replaced, and the incision is closed.

The middle fossa approach is unique compared with the posterior fossa craniotomies (translabyrinthine, retrosigmoid) because the entire IAC is accessible without violating the inner ear. This exposure allows removal of intracanalicular tumors while preserving hearing. The limitations of the middle fossa approach include tumors with a greater than 1 cm cisternal component. In situations of hearing preservation, an extended middle fossa approach with further removal of bone around the IAC and elevation or division of the superior petrosal sinus and tentorium allows improved exposure into the CPA. The relative merits of the procedure with increased temporal lobe retraction and limited access to the posterior fossa in the event of bleeding relative to a retrosigmoid approach continue to be defined. The disadvantages of the middle fossa approach include temporal lobe retraction and possible poor surgical position of the facial nerve relative to the tumor. Temporal lobe retraction may cause transient speech and memory disturbances and auditory hallucinations. The facial nerve, especially if the tumor originates from the inferior vestibular nerve, will be between the surgeon and the tumor. The increased manipulation of the facial nerve during tumor removal increases the risk of transient facial paresis.

Operative Complications. The intraoperative complications for all three approaches include vascular injury, air embolism, parenchymal brain injury, and cranial nerve injury.[92] The AICA originates from the basilar artery and supplies the labyrinthine artery and the lower portion of the cerebellum and the vein of Labbé, which can be the only venous drainage of the temporal lobe, are vulnerable during VS surgery. In the event of an air

embolism via an open vein, the patient should be placed into a left lateral and Trendelenburg position to trap the air in the right ventricle and then the air can be aspirated via a central venous catheter. The cerebellum during a retrosigmoid craniotomy and the temporal lobe during a middle fossa craniotomy are at risk from retraction injury.

Postoperative Complications. Postoperative complications include hemorrhage, stroke, venous thromboembolism, syndrome of inappropriate antidiuretic hormone, CSF leak, and meningitis.[92] Postoperative hemorrhage will manifest as neurologic and cardiovascular deterioration and will require evacuation. Studies have shown that postoperative low-molecular-weight heparin in addition to compression stockings and intermittent pneumatic compression devices may reduce the risk of thromboembolism in high-risk patients (elderly, obese) without increasing the risk of an intracranial bleed.[93] The most common complication is a CSF leak, which occurs in 10 to 15% of patients either via the wound or via a pneumatic pathway to the eustachian tube. The majority of these leaks resolve with conservative care, which includes placing wound sutures at the leak site, replacing the mastoid dressing, decreasing intracranial pressure with acetazolamide, restricting fluid intake, and resting in bed. Some patients will also require a lumbar subarachnoid drain, and a very few patients will need surgical reexploration.[94] A related complication is meningitis. Meningitis occurs in 2 to 10% of patients and may be aseptic, bacterial, or lipoid owing to irritation from the fat graft. The distinction between aseptic and bacterial meningitis is necessary because the treatment for aseptic meningitis is a corticosteroid with taper and for bacterial meningitis is antibiotics.[95] Delayed meningitis should be considered bacterial and is likely to be caused by a CSF leak.

Prognosis and Rehabilitation. Patients are most concerned about deafness, imbalance, and facial nerve weakness. The most important factors for hearing preservation are tumor size and preoperative hearing level. Hearing preservation ranges from 20 to 70%. Patients with good contralateral hearing tolerate the unilateral loss or may be helped with a CROS (contralateral routing of signals) hearing aid, and patients with poor contralateral hearing may be rehabilitated with a cochlear implant if the cochlear nerve fibers are preserved. Almost half of patients will have vertigo or imbalance beyond the postoperative period, but these symptoms have a minimal impact on daily activities. The rapidity of vestibular compensation after unilateral vestibular loss is determined by the patient's efforts to exercise and challenge the vestibular system. Patients who continue to have dysequilibrium in the extended postoperative period should have vestibular rehabilitation therapy. Facial nerve function is also best predicted by tumor size. In smaller tumors, over 90% of patients will have House-Brackmann grade 1 or 2 function (grade 1 is normal and grade 6 is complete paralysis). If all sizes are considered, approximately 80% of patients will ultimately have grade 1 or 2 function.[96]

The rehabilitation of facial nerve injury is based on general principles of nerve injury, recovery, and rehabilitation. If the nerve is transected intraoperatively, the nerve should be repaired primarily if possible or with a greater auricular interposition graft. Postoperative function may be predicted in an anatomically intact nerve by the intraoperative stimulability of the nerve. If the nerve can be stimulated at less than 0.2 V, there is an over 85% chance of grade 1 or 2 function at 1 year.[97] The lack of facial function (grade 6) at 1 year and no reinnervation potentials on electromyography (EMG) should lead to a hypoglossal-facial transposition, interposition nerve graft, or a cross-facial graft. If facial rehabilitation has been delayed and there is electrical silence of the facial muscle on EMG, muscle transpositions with the temporalis or masseter muscle to the lip give improved tone and symmetry to the lower face. The upper face can be rehabilitated with a brow lift and gold weight for the eyelid. The eye must be protected with lubrication, ointment, and eye bubble if there is incomplete eye closure or lack of sensation to the cornea owing to trigeminal nerve involvement. Lack of eye care will lead to corneal injury and blindness.

In summary, all three approaches have mortality rates of less than 1% with an over 90% rate of tumor removal and facial nerve preservation.[92] The translabyrinthine approach has facial nerve preservation rates as high as 98%. The retrosigmoid approach allows 50% hearing preservation, and the middle fossa approach allows up to 70% hearing preservation.[98,99] The recurrence rate is less than 1.5%, and the majority of patients feel that they can return to full preoperative activities by 3 months.[100]

These three approaches allow the management of most CPA tumors. A small subset of patients will not be surgical candidates and will require observation or radiation therapy.

Observation. The predictable correlation between VS size and significant neurologic symptoms and the relative slow growth of VSs allow observation to be a management option for VSs. Patients may be observed if their life expectancy is shorter than the growth time required for the VS to cause significant neurologic symptoms. The growth pattern of the VS should be assessed in these patients with a second radiologic evaluation in 6 months and then yearly radiologic evaluations. Studies have shown that 15 to 24% of patients undergoing conservative management will require surgery or stereotactic radiation.[101,102] If the growth rate in the first year exceeds 2 to 3 mm per year, the patient will likely need treatment for the VS. The patient should understand the opportunity costs of conservative management, including having to treat a tumor that is larger and less amenable to hearing preservation procedures and/or stereotactic radiation.

Stereotactic Radiation. The goal of stereotactic radiation is to prevent further growth of the VS while preserving hearing and facial nerve function. This goal directly differs from the goal of complete tumor removal in microsurgical therapy. The mechanism of stereotactic radiation relies on delivering radiation to a specific intracranial target by using several precisely collimated beams of ionizing radiation. The beams take various pathways to the target tissue, creating a sharp dose gradient between the target tissue and the surrounding tissue. The ionizing radiation causes necrosis and vascular fibrosis, and the time course of effect is over 1 to 2 years. There is an expected transient swelling of the tumor for 1 to 2 years. The ionizing radiation is most commonly delivered using a 201-source cobalt 60 gamma-knife system. The standard linear accelerator can also be adapted to deliver stereotactic radiation. The practical aspects include that the patient wears a stereotactic head frame, computer-assisted radiation planning using a MRI, and a single treatment for delivery of the radiation.[103]

The success of stereotactic radiation to arrest tumor growth depends on the dose of radiation delivered. However, the rate of cranial neuropathies, including hearing loss, is decreased by lowering the radiation dose. The current trend has been to lower the marginal radiation dose, and long-term tumor control with these current dosing plans is under investigation.[104] Since VSs have a slow growth rate, these studies will require 5- to 10-year follow-ups to be confident about tumor control. Studies have shown control rates from 85 to over 95%. The hearing preservation rate decreases each year after radiation and stabilizes after 3 years at 50%. The rate of facial nerve dysfunction varies from 3 to 50% based on the radiation dose at the margin of the tumor and the length of the facial nerve in the radiation field. Approximately 20% of patients will have trigeminal neuropathy. The persistence and extent of these neuropathies with the lower dosing protocols continue to be studied. Hydrocephalus is also a complication of radiation.[105,106]

As the long-term effectiveness and sequelae of stereotactic radiation are further defined, the indications for radiation therapy will become further defined. Radiation therapy is useful in patients in whom arrest of tumor growth is acceptable. These patients have either short life expectancies or high surgical risk. Stereotactic radiation may improve hearing preservation in patients with 2 to 3 cm VSs compared to microsurgery. Radiation therapy in large tumors (> 3 cm) or tumors causing brain compression will exacerbate symptoms owing to initial tumor swelling.[107]

Nonvestibular Schwannomas. Vestibular schwannomas represent over 95% of all CPA schwannomas. In addition to CN VIII, schwannomas of CNs V, VII, IX, X, XI, and XII can involve the CPA.[108] Cerebellopontine angle schwannomas share clinical, pathologic, and imaging characteristics. The primary treatment, similar to that for VSs, is surgical resection. The surgical approach is based on the location of the schwannoma and the patient's hearing status. Resection of cranial nerve schwannomas may lead to significant CN dysfunction, so preoperative CN function and postoperative rehabilitation are important issues to consider.

Facial Nerve Schwannomas. Facial nerve schwannomas most commonly occur at the geniculate ganglion but can involve any portion of the facial nerve. Similar to VSs, facial nerve schwannomas present with hearing loss, tinnitus, and imbalance (CPA symptoms). Facial nerve symptoms, such as facial spasm or weakness, usually present with larger

tumors. Audiovestibular testing shows abnormality in acoustic reflex testing because of impairment of CN VII motor supply to the stapedius muscle. Facial nerve schwannomas are associated with reduction of electroneuronographic potentials on the ipsilateral side even when there is no clinically evident palsy. Imaging often does not allow differentiation between a VS and a facial nerve schwannoma. Distinguishing features on imaging of facial nerve schwannomas include expansion of the fallopian canal, extension from the geniculate ganglion into the middle fossa, and location of the schwannoma in the anterosuperior portion of the IAC (a position eccentric to the axis of the IAC).[108] These schwannomas may be observed until facial nerve function has deteriorated or a neurologic complication becomes imminent as resection usually requires division and grafting of the facial nerve. Mimetic function following interposition grafting is poor.[109] The best facial function attainable is a House-Brackmann grade 3.

Trigeminal Nerve Schwannomas. Trigeminal nerve schwannomas initially present with ipsilateral facial hypesthesia, paresthesias, neuralgia, and difficulties with chewing.[110] Trigeminal nerve schwannomas arise from the gasserian ganglion in the middle fossa and grow posteriorly to involve the CPA or arise from the root of the nerve and directly involve the anterior part of the CPA and Meckel's cave. These tumors frequently involve both the middle and posterior fossae, and a combined approach may be necessary for resection.

Lower Cranial Nerve Schwannomas. Schwannomas of CNs IX, X, and XI cause smooth enlargement of the jugular foramen and may grow superiorly into the CPA or inferiorly into the parapharyngeal space. Cranial nerve IX, X, and XI schwannomas produce symptoms based on their CN functions and so cause hypesthesia and weakness of the palate, vocal cord, and shoulder, respectively. Patients will present with dysphagia, hoarseness, and shoulder weakness. Cerebellopontine angle involvement will also lead to CPA symptoms. Schwannomas of CN XII cause hemiatrophy of the tongue and expansion of the hypoglossal canal. The treatment is surgical removal and rehabilitation of the patient's functional deficit.[111]

Meningiomas. Meningiomas are the second most common CPA tumors and account for 3 to 10% of tumors at this location.[112] Compared with schwannomas, meningiomas are a more heterogeneous group of tumors in regard to pathology, anatomic location, and treatment outcome. The majority of these tumors are benign and slow growing, and 1% will become symptomatic. Meningiomas differ in pathogenesis, anatomic location, and imaging characteristics from VSs but are nearly indistinguishable in terms of clinical presentation and audiovestibular testing. Meningiomas are primarily managed by surgical excision.

Pathogenesis. Meningiomas arise from arachnoid villi cap cells and are located along dura, venous sinuses, and neurovascular foramina. Meningiomas are most commonly sporadic but may occur in familial syndromes such as NF2, Werner's syndrome, and Gorlin's syndrome. More than one-third of patients with NF2 have meningiomas. Molecular studies have shown deletions in chromosome 22 in nearly 75% of meningiomas. Specifically, mutations in the NF2 gene, merlin, have been shown in 30 to 35% of meningiomas. Although the majority of meningiomas are benign, 5% are malignant. Chromosomal abnormalities in 1p, 6q, 9p, 10q, and 14q are seen in more aggressive or malignant meningiomas.[113]

Pathology. ***Gross.*** Meningiomas may be globular or flat (en plaque) and have a firm, gritty consistency with a thin capsule. They are densely adherent to the dura and may either cause reactive hyperostosis of the underlying bone or infiltrate the bone via haversian canals. Hyperostosis is seen in 25% of meningiomas.[114]

Microscopic. Meningiomas have an epithelial and mesenchymal differentiation. These tumor cells generally have a uniform distribution. Some meningiomas may demonstrate a papillary pattern interspersed in the uniform distribution. Calcification of the tips of these papillary whorls leads to the formation of calcospherite or psammoma bodies. The ratio of epithelial to mesenchymal differentiation and the presence of psammoma bodies are the basis for histologic classification of meningiomas. The four groups include meningotheliomatous (syncytial), transitional, fibrous, and angiomatous.[115] The syncytial type is the most common category. The syncytial type is usually benign and lacks papillary whorls or psammoma bodies. The transitional type has an increased amount of connective tissue and

psammoma bodies relative to the syncytial type. The transitional type lies histologically between the syncytial and fibrous types. The fibrous type has increased connective tissue, mesenchymal differentiation, and some psammoma bodies. The angiomatous types may actually be sarcomas rather than meningiomas and have a more aggressive growth pattern with bone invasion. The microscopic identification of a malignant meningioma includes increased mitotic figures, atypical mitosis, necrosis, lack of cellular uniformity, and brain invasion.[116]

SYMPTOMS AND SIGNS. From the onset of symptoms, meningiomas take an average of 5 years to be diagnosed. The typical patient is a woman in her fourth or fifth decade of life. Unlike VSs, there is a 2 to 1 female bias.[117] Meningiomas presenting in younger patients or multiple meningiomas in the same patient should prompt an evaluation for NF2. The most common complaints are the same as those attributable to VSs and include unilateral hearing loss (80%), vertigo or imbalance (75%), and tinnitus (60%). Symptoms and signs more common with meningiomas relative to VSs include trigeminal neuralgia (7 to 22%), facial paresis (11 to 36%), lower CN deficits (5 to 10%), and visual disturbances (8%).[112,117,118]

DIAGNOSTIC EVALUATION. ***Auditory-Vestibular Evaluation.*** The hearing loss has no characteristic audiometric pattern. The SDSs suggest a retrocochlear lesion in 50% of patients. The ABR may be normal in 25% of patients.[119]

Imaging. Imaging provides the diagnosis of meningioma and allows the differentiation between meningioma and VS[120] (Table 20–7). A CT scan without contrast shows an iso- or hyperdense mass with areas of calcification in 10 to 26% and provides information regarding hyperostosis or bony invasion. Meningiomas enhance homogeneously with CT contrast, and 90% of meningiomas can be detected by contrast-enhanced CT.[114] Magnetic resonance imaging is the study of choice for the diagnosis of meningiomas. Meningiomas are hypo- to isointense on T_1-weighted images and have a variable intensity on T_2-weighted images. Areas of calcification appear dark on both T_1-weighted and T_2-weighted images (Figure 20–7). Unlike VSs, meningiomas are broad based (sessile) and are usually not centered over the porus acusticus internus. The broad-based attachment to the petrous wall leads to an obtuse bone-tumor angle. There is no widening of the IAC. Also, unlike VSs, meningiomas more commonly herniate into the middle fossa. T_1-weighted enhanced images can show an enhancing dural tail (meningeal sign) adjacent to the bulk of the tumor in 50 to 70% of meningiomas.[121]

MANAGEMENT. The two primary management options are observation and surgery. Observation is indicated in patients with limited life expectancy or in whom the expected morbidity of surgical excision is not justified. Surgical treatment ideally consists of total meningioma removal, excision of a cuff of surrounding dura, and drilling of the underlying bone. The surgical approach is based on the tumor location and the patient's hearing status. In contrast to VSs, the anatomic location of posterior fossa meningiomas is varied (Figure 20–8). The site of the meningioma is a major determinant of the types of morbidity from the tumor and the success of treatment. A simple classification differentiates whether the tumor is present medial or lateral to the IAC. Meningiomas medial to the IAC are more common. These meningiomas commonly arise along the inferior petrosal sinus and may involve the petrous apex, lateral part of the clivus, and Meckel's cave. Meningiomas lateral to the IAC involve the sigmoid sinus, jugular bulb, and superior petrosal sinus. In an uncommon pattern, meningiomas may be centered on the IAC and closely mimic a VS. Internal auditory canal meningiomas may invade the inner or middle ear. Meningiomas may also be superior to the IAC and be considered midpetrosal meningiomas.[112]

Meningiomas lateral to the IAC are reached via a retrosigmoid approach. The facial nerve in lateral meningiomas is most often displaced anteriorly and so does not lie between the surgeon and the tumor. Therefore, the facial nerve is less traumatized during tumor removal. The retrosigmoid approach also allows hearing preservation. Limited intracanalicular meningiomas may be managed by the middle fossa approach, especially if hearing preservation is possible. Meningiomas involving the IAC and having poor hearing are reached via the translabyrinthine approach. If the tumor invades the cochlea and has anteromedial extension to the clivus or Meckel's cave, a transcochlear approach should be considered[122] (see Figure 20–6, A). The transcochlear approach sacrifices hearing and requires rerouting

TABLE 20–7. Radiologic Differentation of Meningioma versus Vestibular Schwannoma

	Meningioma	*Vestibular Schwannoma*
Shape	Sessile	Globular
Internal auditory canal	Eccentric, extrinsic, and not eroded	Centered, penetrates, and eroded
Calcification	Present	Absent
Hyperostosis	Present	Absent
Tumor-bone angle	Obtuse	Acute
Meningeal sign	Present	Absent

of the facial nerve. Sixty percent of CPA meningiomas involve the middle fossa and may require a combined middle and posterior fossa craniotomy. The type of posterior fossa craniotomy in the combined approach will depend on the need for hearing preservation and the extent of surgical exposure required.

Total tumor removal is accomplished in 70 to 85% of patients with meningiomas. Incomplete tumor removal is often associated with adherence of the meningioma to the brainstem or cavernous sinus involvement.[123] The long-term recurrence after total tumor removal is between 10 and 30%, whereas that of subtotal removal is over 50%.[124] Hearing preservation is more likely than in VSs and approaches 70%. The facial nerve function has a 17% rate of deterioration from preoperative levels.[125] The other complications include gait disturbance and CSF leak (8%). The operative mortality is between 1 and 9%.[126]

Adjunctive therapies include external beam and stereotactic radiation therapy. Radiation therapy should be considered in cases of inoperable tumors, subtotal resection, recurrent tumors, and malignant tumors. Retrospective reviews of patients in whom subtotal tumor removals were carried out have shown that external beam radiation reduces recurrence rates.[127] The role of stereotactic radiation therapy in meningiomas continues to be defined.

EPIDERMOIDS. Epidermoids are much less common than VSs or meningiomas. They account for approximately 5% of CPA lesions. Epidermoids are slow growing and often grow to a large size before causing CPA symptoms because they initially grow around structures via paths of least resistance rather than cause compression. Epidermoids are treated by surgical excision, but total removal is more difficult than with VSs because epidermoids become adherent to normal structures.[128]

PATHOGENESIS. Epidermoids likely develop from ectodermal inclusions that become trapped during embryogenesis. These ectodermal inclusions lead to a keratinizing squamous cell epithelium in the CPA. The squamous cell epithelium produces a cyst filled with sloughed keratinaceous debris.[129]

PATHOLOGY. The gross appearance is that of a nodular cyst. The cyst is lined with squamous cell epithelium, and the cyst is filled with lamella of desquamated keratinaceous debris. The majority of epidermoids are benign, but there are a few reports of squamous cell carcinoma arising in epidermoids.[130]

SYMPTOMS AND SIGNS. The presentation is similar to other CPA lesions, with hearing loss being the most common symptom. Epidermoids have a higher rate of preoperative facial (40%) and trigeminal (50%) nerve involvement relative to VSs. Patients may present with hemifacial spasm, facial hypesthesia, neuralgia, or wasting of the muscles of mastication.[131]

DIAGNOSTIC EVALUATION. ***Auditory-Vestibular Evaluation.*** Auditory-vestibular testing does not provide any distinguishing patterns for epidermoids.

Imaging. On CT, epidermoids are hypodense compared to brain. A distinguishing characteristic relative to VSs and meningiomas is that epidermoids show no enhancement with intravenous contrast.

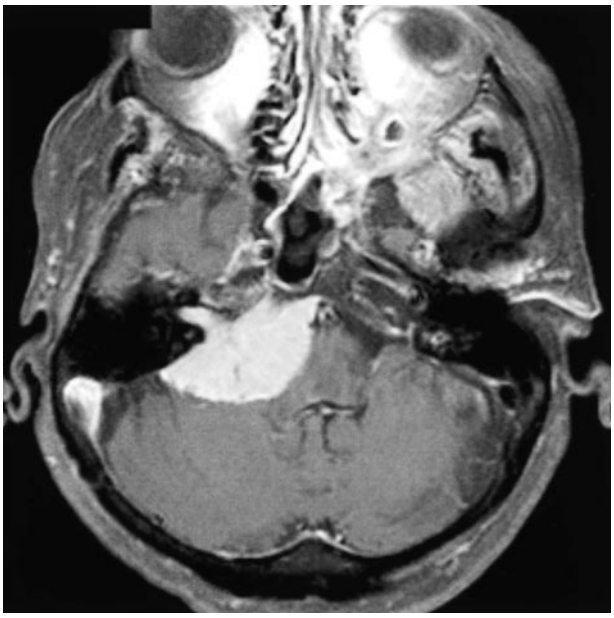


FIGURE 20–7. Axial T_1-weighted magnetic resonance image with gadolinium enhancement demonstrating a cerebellopontine angle meningioma. Unlike vestibular schwannomas, meningiomas are broad based (sessile), have a dural tail, and are eccentric to the internal auditory canal. The broad-based attachment to the petrous wall leads to an obtuse bone-tumor angle. (Courtesy of Nancy J. Fischbein.)

Epidermoids have irregular borders, are not centered on the IAC, and do not usually widen the IAC (Figure 20–9). Epidermoids have characteristics similar to CSF on MRI (hypointense on T_1-weighted images and hyperintense on T_2-weighted images) and again do not enhance with gadolinium contrast.[132]

MANAGEMENT. The primary treatment of epidermoids is surgical. The approaches include retrosigmoid in hearing preservation and translabyrinthine in patients with significant hearing loss. Any extension into the middle fossa can usually be removed via the posterior fossa craniotomy. The ability to remove the tumor completely is limited by the propensity of epidermoids to adhere to neurovascular structures. Attempts at complete tumor removal may increase the rate of postoperative transient or permanent CN palsies. Total tumor removal is accomplished in less than 50% of patients, and the recurrence rate may be as high as 50%. Nevertheless, the majority of patients have excellent or good postoperative function.[133]

Congenital Rest Lesions Congenital rest lesions involving the CPA include epidermoids, arachnoid cysts, and lipomas. These lesions occur owing to errors in embryogenesis. These errors allow vestigial structures to remain and grow during adult life. These lesions are not aggressive but rather are slow growing and have a tendency to envelop the neurovascular structures in the CPA. The presenting symptoms are very similar if not indistinguishable from those of VSs, and only imaging allows their differentiation. The imaging characteristics on CT scan are very similar and include a well-encapsulated, hypodense mass that does not enhance with contrast. Magnetic resonance imaging allows differentiation based on the signal characteristics of desquamated epithelium, CSF, or fat. The treatment is surgical, but total removal is more difficult than in VSs and is not always necessary.[61]

ARACHNOID CYSTS. Arachnoid cysts are CSF-filled cysts surrounded by an epithelial lining. This epithelial lining originates from a duplication of the arachnoid membrane and has secretory capabilities. The rate of growth is unpredictable, and patients may present with arachnoid cysts in the CPA or IAC at any age. The key imaging point is that these cysts match the signal intensity of CSF on every MRI sequence and do not enhance with gadolinium. The treatment involves diuretics, shunting procedures, and marsupialization of the cyst into the subarachnoid space.[134]

LIPOMAS. Lipomas are rare lesions of the CPA and IAC. They are caused by congenital malformations that lead to proliferation of adipocytes in subarachnoid cisterns or ventricles. Magnetic resonance imaging parallels the intensities of fat, so lipomas are hyperintense on T_1-weighted images, show no enhancement with gadolinium, are hypointense on T_2-weighted images, and become hypointense on T_1-weighted images with fat suppression. The neurovascular structures of the CPA travel through the lipoma. Therefore, the surgical treatment of these lesions, if they become symptomatic, is conservative debulking.[135]

Vascular Lesions A variety of vascular lesions may involve the CPA directly or via extension. These include paragangliomas (glomus jugulare tumors), hemangiomas, and aneurysms.

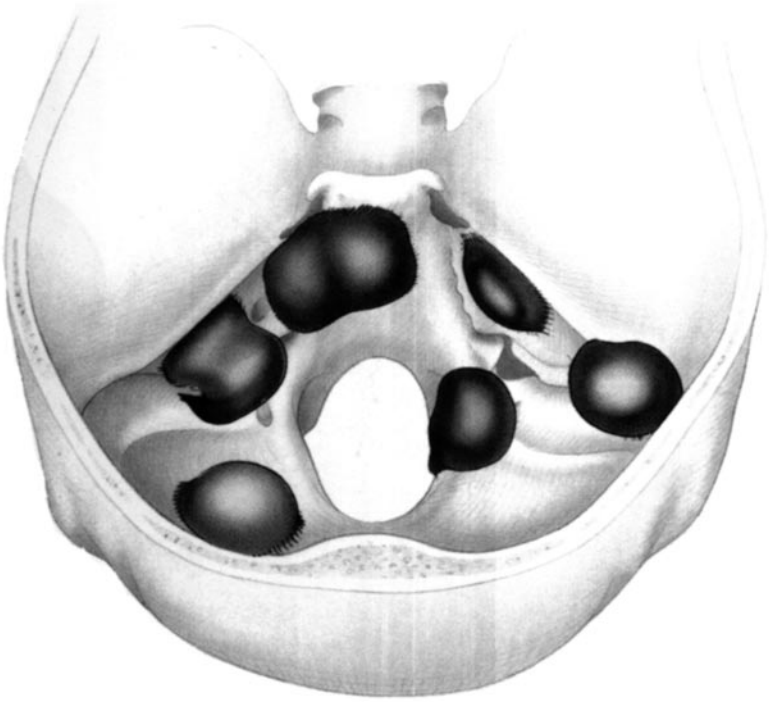


FIGURE 20–8. Meningiomas arise from arachnoid villi cap cells and are located along dural edges, venous sinuses, and neurovascular foramina. The site of the meningioma is a major determinant of types of morbidity from the tumor and the success of treatment. A simple classification differentiates whether the tumor is present medial or lateral to the internal auditory canal (IAC). Meningiomas medial to the IAC commonly arise along the inferior petrosal sinus and may involve the petrous apex, lateral part of the clivus, and Meckel's cave. Meningiomas lateral to the IAC involve the sigmoid sinus, jugular bulb, and superior petrosal sinus. In an uncommon pattern, meningiomas may be centered on the IAC or superior to the IAC and considered a midpetrosal meningioma. Reproduced with permission from Jackler RK.[50]

Paragangliomas (Glomus Jugulare Tumors) Glomus jugulare tumors arise from paraganglionic tissues (chief cells) and have intracranial extension in 15% of patients. These tumors are slow growing and present initially with pulsatile tinnitus and conductive hearing loss. Further growth at the jugular foramen causes lower cranial neuropathies, and then intracranial extension into the posterior fossa may lead to sensorineural hearing loss and dizziness.[136] Paragangliomas have a variable appearance on CT scans, but bone algorithms show the extent of temporal bone involvement. Paragangliomas cause irregular expansion of the jugular foramen, whereas lower CN schwannomas cause smooth enlargement of the jugular foramen. Paragangliomas have a salt and pepper appearance on T_2-weighted MRI and have pronounced enhancement with gadolinium. Magnetic resonance imaging also shows the extent of intracranial involvement and the presence of flow voids.[137] Magnetic resonance angiography (MRA) and traditional angiography provide information on involvement of the great vessels and allow preoperative embolization of larger tumors. The treatment requires surgical excision. Jugular bulb involvement may be addressed by a transmastoid neck approach. Extension to the carotid artery or intracranial extension will require an infratemporal fossa approach.[138]

Hemangiomas. Hemangiomas of the temporal bone often involve the geniculate ganglion and the internal auditory meatus. Hemangiomas are benign, slow-growing vascular hamartomas. Hemangiomas involving the geniculate ganglion cause progressive facial paresis. Patients also complain of facial twitches, tinnitus, and facial pain. Hearing loss, if present, is conductive owing to middle ear involvement. Facial paresis occurs sooner with hemangiomas than with facial nerve schwannomas.

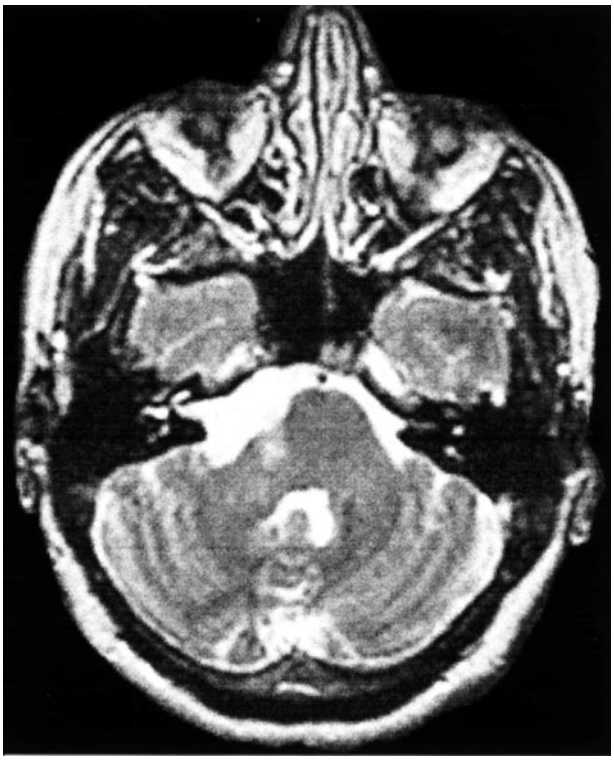


FIGURE 20–9. Axial T_2-weighted magnetic resonance image (MRI) of an epidermoid involving the clivus, petrous bone, internal auditory canal (IAC), and cerebellopontine angle. Epidermoids have irregular borders, are not centered on the IAC, and do not usually widen the IAC. Epidermoids have imaging characteristics similar to cerebrospinal fluid on MRI (hypointense on T_1-weighted images and hyperintense on T_2-weighted images) and, in contrast to vestibular schwannomas and meningiomas, do not enhance with gadolinium contrast. (Courtesy of Nancy J. Fischbein.)

Computed tomographic imaging shows a small, soft tissue mass at the geniculate ganglion with surrounding smooth or irregular enlargement of the fallopian canal. The small size of the soft tissue mass, irregular bony erosion, and presence of calcium in the tumor are all suggestive of geniculate ganglion hemangioma versus facial nerve schwannoma. Magnetic resonance imaging shows isointense T_1-weighted images, intense enhancement, and hyperintense T_2-weighted images.[139] Hemangiomas in the IAC present similarly to VSs, and a preoperative differentiation may be very difficult. The treatment involves surgical removal when there is significant facial nerve dysfunction. Because hearing is often intact, the middle fossa approach provides good surgical exposure and allows hearing preservation. The chance of an intact facial nerve after removal of a hemangioma is higher than a facial nerve schwannoma. Nevertheless, facial nerve anastomosis or grafting is often required.[140]

Aneurysms. Aneurysms and vascular anomalies of posterior circulation (AICA, posterior inferior cerebellar artery, carotid artery, vertebral artery, basilar artery) are rare but produce CPA symptoms by causing compression of neurovascular structures. Aneurysms are seen as enhancing lesions on CT scan. In addition to MRI, MRA and angiography allow characterization of these vascular lesions.[141]

Intra-axial Tumors Intra-axial tumors of the brainstem (glioma), cerebellum (medulloblastoma, astrocytoma, hemangioblastoma), and fourth ventricle (ependymoma, choroid plexus papilloma) may extend into the CPA and present with CPA symptoms.[142] The CPA extension occurs owing to exophytic growth and growth into the CPA via the foramina of Luschka and rarely owing to extra-axial origin directly in the CPA from an embryonic rest. These intra-axial tumors involving the CPA must be differentiated from extra-axial CPA masses to counsel and treat the patient appropriately. The differentiation is based primarily on imaging characteristics. Imaging characteristics suggestive of an extra-axial tumor include bony changes, widening of the subarachnoid cistern, displacement of brain and blood vessels away from the skull or dura, and sharp definition of the tumor margin. Characteristics suggestive of intra-axial tumors include an irregular and poorly defined brain-tumor margin, widening of a foramen of Luschka, brain edema out of proportion to the CPA component of the tumor, and hydrocephalus.[143] The management of intra-axial lesions includes angiography, conservative surgical resection, and adjunctive therapy.

Lesions Extending from the Skull Base Lesions involving the skull base, specifically the petrous apex (cholesterol granulomas), clivus (chordomas), and petro-occipital fissure (chondrosarcomas), may grow posteriorly and laterally to involve the CPA. Each of these lesions has characteristic presentations and imaging criteria but at times may present solely

with CPA symptoms. The treatment is primarily surgical.

Cholesterol Granulomas. The petrous apex of the temporal bone lies anterior and medial to the inner ear and posterior and lateral to the clivus and contains pneumatized air cells in one-third of temporal bones. Obstruction of these air cells leads to inflammation and hemorrhage into the air cells. The phagocytosis of red cells leads to deposition of cholesterol crystals and a foreign body reaction in the petrous apex. This process leads to a cholesterol granuloma that may extend beyond the petrous apex to involve the CPA. Patients may have CPA symptoms in addition to symptoms of headache and sixth nerve dysfunction. The key differential diagnosis in the petrous apex and CPA of a cholesterol granuloma is an epidermoid. The distinguishing factor is that cholesterol granulomas are hyperintense on T_1-weighted and T_2-weighted images, whereas epidermoids are hypointense on T_1-weighted images. The treatment is surgical drainage rather than complete excision when symptomatic. The petrous apex may be accessed via a transmastoid or transcanal approach.[144,145]

Chordomas. Chordomas arise from remnants of the notochord, and skull base chordomas occur at the clivus (spheno-occipital synchondrosis). Patients usually present with headache and diplopia, but posterolateral extension to the CPA may lead to CPA symptoms. Computed tomographic scans show an isodense mass with bone destruction, intratumor calcifications, and marked enhancement with contrast. Magnetic resonance imaging shows hypointense T_1-weighted images, marked gadolinium enhancement, and hyperintense T_2-weighted images. The midline location and bony destruction without sclerosis are characteristic of chordomas. The treatment is complete surgical excision and radiation for subtotal removal or recurrence.[146]

Chondrosarcomas. The main differential diagnosis of chordomas is chondrosarcomas.[146] These tumors arise along the spheno-occipital synchondrosis and are laterally located relative to chordomas. Chondrosarcomas may arise from embryonal cartilage rests located at the skull base synchondrosis. Tumor growth laterally will cause involvement of the IAC and CPA. Hearing loss may very well be the presenting complaint. The differential from chordoma may require immunohistochemistry stains since chondrosarcomas, unlike chordomas, do not stain positively with epithelial tissue markers.[147]

Metastases Metastatic disease from lung, breast, skin (melanoma), prostate, nasopharynx, lymphoma, and kidney are the most common extra-axial malignant neoplasms of the CPA.[148] In contrast to benign lesions of the CPA, malignant lesions of the CPA cause rapid progression of CPA symptoms. On imaging, the lesions are small and isointense to brain on T_1-weighted and T_2-weighted MRI and enhance with gadolinium. There is a high likelihood of parenchymal brain metastases and bilateral CPA lesions. The extent of treatment is based on the extent of the metastatic and primary disease and includes multimodality treatment with surgical biopsy or resection, radiation, and chemotherapy. The other aspect of treatment includes relieving symptoms of hydrocephalus or brainstem compression. The primary extra-axial malignancies of the CPA are exceedingly rare and include lymphomas, squamous cell carcinomas, malignant VSs, and malignant meningiomas.[61]

REFERENCES

1. Shepard NT, Solomon D. Functional operation of the balance system in daily activities. Otolaryngol Clin North Am 2000;33:455–69.
2. Eldridge R, Parry D. Vestibular schwannoma (acoustic neuroma). Consensus development conference. Neurosurgery 1992;30:962–4.
3. Parent A. Carpenter's human neuroanatomy. 9th ed. Baltimore: Williams and Wilkins; 1996.
4. Fetter M, Zee DS. Recovery from unilateral labyrinthectomy in rhesus monkey. J Neurophysiol 1988;59:370–93.
5. Walker MF, Zee DS. Bedside vestibular examination. Otolaryngol Clin North Am 2000;33:495–506.
6. Rosenberg ML, Gizzi M. Neuro-otologic history. Otolaryngol Clin North Am 2000;33:471–82.
7. Solomon D. Distinguishing and treating causes of central vertigo. Otolaryngol Clin North Am 2000; 33:579–601.
8. Brandt T, Strupp M, Arbusow V, Dieringer N. Plasticity of the vestibular system: central compensation and sensory substitution for vestibular deficits. Adv Neurol 1997;73:297–309.

9. Shepard NT, Telian SA. Programmatic vestibular rehabilitation. Otolaryngol Head Neck Surg 1995; 112:173–82.
10. Dix R, Hallpike CS. The pathology, symptomatology, and diagnosis of certain common disorders of the vestibular system. Proc R Soc Med 1952;54:341–54.
11. Froehling DA, Silverstein MD, Mohr DN, et al. Benign positional vertigo: incidence and prognosis in a population-based study in Olmsted County, Minnesota. Mayo Clin Proc 1991;66:596–601.
12. Furman JM, Cass SP. Benign paroxysmal positional vertigo. N Engl J Med 1999;341:1590–6.
13. Baloh RW, Honrubia V, Jacobson K. Benign positional vertigo: clinical and oculographic features in 240 cases. Neurology 1987;37:371–8.
14. Schuknecht HF. Cupulolithiasis. Arch Otolaryngol 1969;90:765–78.
15. Parnes LS, McClure JA. Free-floating endolymph particles: a new operative finding during posterior semicircular canal occlusion. Laryngoscope 1992; 102:988–92.
16. Welling DB, Parnes LS, O'Brien B, et al. Particulate matter in the posterior semicircular canal. Laryngoscope 1997;107:90–4.
17. Honrubia V, Baloh RW, Harris MR, Jacobson KM. Paroxysmal positional vertigo syndrome. Am J Otol 1999;20:465–70.
18. Epley JM. The canalith repositioning procedure: for treatment of benign paroxysmal positional vertigo. Otolaryngol Head Neck Surg 1992;107:399–404.
19. Brandt T, Daroff RB. Physical therapy for benign paroxysmal positional vertigo. Arch Otolaryngol 1980;106:484–5.
20. Parnes LS, McClure JA. Posterior semicircular canal occlusion for intractable benign paroxysmal positional vertigo. Ann Otol Rhinol Laryngol 1990; 99:330–4.
21. Gacek RR. Technique and results of singular neurectomy for the management of benign paroxysmal positional vertigo. Acta Otolaryngol (Stockh) 1995; 115:154–7.
22. Committee on Hearing and Equilibrium. Guidelines for the diagnosis and evaluation of therapy in Meniere's disease. Otolaryngol Head Neck Surg 1995;113:181–5.
23. Lustig LR, Lalwani A. The history of Meniere's disease. Otolaryngol Clin North Am 1997;30:917–45.
24. Jackler RK, Whinney D. A century of eighth nerve surgery. Otol Neurotol 2001;22:401–16.
25. Hallpike CS, Cairns H. Observation on the pathology of Meniere's syndrome. J Laryngol Otol 1938; 53:625–54.
26. House WF. Surgical exposure of the internal auditory canal and its contents through the middle cranial fossa. Laryngoscope 1961;71:1363–85.
27. Harner SG, Driscoll CL, Facer GW, et al. Long-term follow-up of transtympanic gentamicin for Meniere's syndrome. Otol Neurotol 2001;22:210–4.
28. Schuknecht HF, Ruther A. Blockage of longitudinal flow in endolymphatic hydrops. Eur Arch Otorhinolaryngol 1991;248:209–17.
29. Hebbar GK, Rask-Andersen H, Linthicum FH. Three-dimensional analysis of 61 human endolymphatic ducts and sacs in ears with and without Meniere's disease. Ann Otol Rhinol Laryngol 1991;100:219–25.
30. Wackym PA, Sando I. Molecular and cellular pathology of Meniere's disease. Otolaryngol Clin North Am 1997;30:947–60.
31. Tsuji K, Velazquez-Villasenor L, Rauch SD, et al. Temporal bone studies of the human peripheral vestibular system. Meniere's disease. Ann Otol Rhinol Laryngol Suppl 2000;181:26–31.
32. Gibson WP, Arenberg IK. Pathophysiologic theories in the etiology of Meniere's disease. Otolaryngol Clin North Am 1997;30:961–7.
33. Dornhoffer JL, Arenberg IK. Immune mechanisms in Meniere's syndrome. Otolaryngol Clin North Am 1997;30:1017–26.
34. Derebery MJ. The role of allergy in Meniere's disease. Otolaryngol Clin North Am 1997;30:1007–16.
35. Green JD, Blum DJ, Harner SG. Longitudinal followup of patients with Meniere's disease. Otolaryngol Head Neck Surg 1991;104:783–8.
36. Katsarkas A. Hearing loss and vestibular dysfunction in Meniere's disease. Acta Otolaryngol (Stockh) 1996;116:185–8.
37. Ruckenstein MJ. Vertigo and dysequilibrium with associated hearing loss. Otolaryngol Clin North Am 2000;33:535–62.
38. Klockhoff I, Lindblom U. Endolymphatic hydrops revealed by glycerol test. Preliminary report. Acta Otolaryngol (Stockh) 1966;61:459–62.
39. Schmidt PH, Eggermont JJ, Odenthal DW. Study of Meniere's disease by electrocochleography. Acta Otolaryngol Suppl (Stockh) 1974;316:75–84.
40. Stahle J, Friberg U, Svedberg A. Long-term progression of Meniere's disease. Acta Otolaryngol Suppl (Stockh) 1991;485:78–83.

41. Gottschlich S, Billings PB, Keithley EM, et al. Assessment of serum antibodies in patients with rapidly progressive sensorineural hearing loss and Meniere's disease. Laryngoscope 1995;105:1347–52.
42. Devaiah AK, Ator GA. Clinical indicators useful in predicting response to the medical management of Meniere's disease. Laryngoscope 2000;110(11): 1861–5.
43. van Deelen GW, Huizing EH. Use of a diuretic (Dyazide) in the treatment of Meniere's disease. A double-blind cross-over placebo-controlled study. ORL J Otorhinolaryngol Relat Spec 1986;48:287–92.
44. Kaplan DM, Nedzelski JM, Chen JM, Shipp DB. Intratympanic gentamicin for the treatment of unilateral Meniere's disease. Laryngoscope 2000;110: 1298–305.
45. Atlas JT, Parnes LS. Intratympanic gentamicin titration therapy for intractable Meniere's disease. Am J Otol 1999;20:357–63.
46. Thomsen J, Bretlau P, Tos M, Johnsen NJ. Placebo effect in surgery for Meniere's disease. A double-blind, placebo-controlled study on endolymphatic sac shunt surgery. Arch Otolaryngol 1981;107:271–7.
47. Bretlau P, Thomsen J, Tos M, Johnsen NJ. Placebo effect in surgery for Meniere's disease: nine-year follow-up. Am J Otol 1989;10:259–61.
48. Pensak ML, Friedman RA. The role of endolymphatic mastoid shunt surgery in the managed care era. Am J Otol 1998;19:337–40.
49. LaRouere MJ. Surgical treatment of Meniere's disease. Otolaryngol Clin North Am 1996;29: 311–22.
50. Jackler RK. Atlas of neurotology and skull base surgery. St. Louis: Mosby; 1996.
51. Levine SC, Glasscock M, McKennan KX. Long-term results of labyrinthectomy. Laryngoscope 1990;100: 125–7.
52. Nadol JB. Vestibular neuritis. Otolaryngol Head Neck Surg 1995;112:162–72.
53. Strupp M, Brandt T. Vestibular neuritis. Adv Otorhinolaryngol 1999;55:111–36.
54. Schuknecht HF, Kitamura K. Second Louis H. Clerf Lecture. Vestibular neuritis. Ann Otol Rhinol Laryngol Suppl 1981;90:1–19.
55. Matsuo T, Sekitani T. Vestibular neuronitis: neurotological findings and progress. ORL J Otorhinolaryngol Relat Spec 1985;47:199–206.
56. Hotson JR, Baloh RW. Acute vestibular syndrome. N Engl J Med 1998;339:680–5.
57. Hood JD, Korres S. Vestibular suppression in peripheral and central vestibular disorders. Brain 1979; 102:785–804.
58. Norrving B, Magnusson M, Holtas S. Isolated acute vertigo in the elderly; vestibular or vascular disease? Acta Neurol Scand 1995;91:43–8.
59. El-Kashlan HK, Telian SA. Diagnosis and initiating treatment for peripheral system disorders: imbalance and dizziness with normal hearing. Otolaryngol Clin North Am 2000;33:563–78.
60. Mahaley MS, Mettlin C, Natarajan N, et al. Analysis of patterns of care of brain tumor patients in the United States: a study of the Brain Tumor Section of the AANS and the CNS and the Commission on Cancer of the ACS. Clin Neurosurg 1990;36:347–52.
61. Eisenman D, Voigt EP, Selesnick SH. Unusual tumors of the internal auditory canal and cerebellopontine angle. In: Jackler RK, Driscoll CLW, editors. Tumors of the ear and temporal bone. Philadelphia: Lippincott, Williams and Wilkins; 2000. p. 236–75.
62. Cushing H. Tumors of the nervus acusticus and the syndrome of the cerebellopontile. Philadelphia: WB Saunders; 1917.
63. Dandy W. Results of removal of acoustic neuromas by the unilateral approach. Arch Surg 1941;42:1027–33.
64. Lang J. Clinical anatomy of the cerebellopontine angle and internal acoustic meatus. Adv Otorhinolaryngol 1984;34:8–24.
65. Stewart TJ, Liland J, Schuknecht HF. Occult schwannomas of the vestibular nerve. Arch Otolaryngol 1975;101:91–5.
66. Tos M, Charabi S, Thomsen J. Clinical experience with vestibular schwannomas: epidemiology, symptomatology, diagnosis, and surgical results. Eur Arch Otorhinolaryngol 1998;255:1–6.
67. Tallan EM, Harner SG, Beatty CW, et al. Does the distribution of Schwann cells correlate with the observed occurrence of acoustic neuromas? Am J Otol 1993;14:131–4.
68. Rouleau GA, Wertelecki W, Haines JL, et al. Genetic linkage of bilateral acoustic neurofibromatosis to a DNA marker on chromosome 22. Nature 1987;329: 246–8.
69. Rouleau GA, Merel P, Lutchman M, et al. Alteration in a new gene encoding a putative membrane-organizing protein causes neuro-fibromatosis type 2. Nature 1993;363:515–21.
70. Shaw RJ, McClatchey AI, Jacks T. Regulation of the neurofibromatosis type 2 tumor suppressor protein,

merlin, by adhesion and growth arrest stimuli. J Biol Chem 1998;273:7757–64.
71. Irving RM. The molecular pathology of tumours of the ear and temporal bone. J Laryngol Otol 1998; 112:1011–8.
72. National Institutes of Health Consensus Development Conference. Neurofibromatosis conference statement. Arch Neurol 1988;45:575–8.
73. Kanter WR, Eldridge R, Fabricant R, et al. Central neurofibromatosis with bilateral acoustic neuroma; genetic, clinical and biochemical distinctions from peripheral neurofibromatosis. Neurology 1980;30: 851–9.
74. Pirsig W, Ziemann-Becker B, Tesschler-Nicola M. Acoustic neuroma: four thousand years ago. In: Tos M, Thomson J, editors. Proceedings of the First International Conference on Acoustic Neuroma. Amsterdam: Kugler; 1992. p. 7–12.
75. Kuo TC, Jackler RK, Wong K, et al. Are acoustic neuromas encapsulated tumors? Otolaryngol Head Neck Surg 1997;117:606–9.
76. Nager GT. Acoustic neurinomas. Pathology and differential diagnosis. Arch Otolaryngol 1969;89:252–79.
77. Selesnick SH, Johnson G. Radiologic surveillance of acoustic neuromas. Am J Otol 1998;19:846–9.
78. Matthies C, Samii M. Management of 1000 vestibular schwannomas (acoustic neuromas): clinical presentation. Neurosurgery 1997;40:1–10.
79. Selesnick SH, Jackler RK, Pitts LW. The changing clinical presentation of acoustic tumors in the MRI era. Laryngoscope 1993;103:431–6.
80. Selesnick SH, Jackler RK. Atypical hearing loss in acoustic neuroma patients. Laryngoscope 1993; 103:437–41.
81. Neary WJ, Newton VE, Laoide-Kemp SN, et al. A clinical, genetic and audiological study of patients and families with unilateral vestibular schwannomas. II. Audiological findings in 93 patients with unilateral vestibular schwannomas. J Laryngol Otol 1996;110:1120–8.
82. Lustig LR, Rifkin S, Jackler RK, Pitts LH. Acoustic neuromas presenting with normal or symmetrical hearing: factors associated with diagnosis and outcome. Am J Otol 1998;19:212–8.
83. Huygen PL, Hoogland GA. Vestibular and ocular motor manifestations of cerebellopontine angle tumours. Adv Otorhinolaryngol 1984;34:57–70.
84. Wilson DF, Hodgson RS, Gustafson MF, et al. The sensitivity of auditory brainstem response testing in small acoustic neuromas. Laryngoscope 1992;102: 961–4.
85. Saeed SR, Woolford TJ, Ramsden RT, Lye RH. Magnetic resonance imaging: a cost-effective first line investigation in the detection of vestibular schwannomas. Br J Neurosurg 1995;9:497–503.
86. Curtin HD, Hirsch WL. Imaging of acoustic neuromas. Otolaryngol Clin North Am 1992;25:553–607.
87. Shelton C, Harnsberger HR, Allen R, King B. Fast spin echo magnetic resonance imaging: clinical application in screening for acoustic neuroma. Otolaryngol Head Neck Surg 1996;114:71–6.
88. Jackler RK. Overview of surgical neurotology. In: Jackler RK, Brackmann DE, editors. Neurotology. St. Louis: Mosby; 1994. p. 651–84.
89. Jackler RK, Pitts LH. Selection of surgical approach to acoustic neuroma. Otolaryngol Clin North Am 1992;25:361–87.
90. Catalano PJ, Jacobowitz O, Post KD. Prevention of headache after retrosigmoid removal of acoustic tumors. Am J Otol 1996;17:904–8.
91. Ruckenstein MJ, Harris JP, Cueva RA, et al. Pain subsequent to resection of acoustic neuromas via suboccipital and translabyrinthine approaches. Am J Otol 1996;17:620–4.
92. Samii M, Matthies C. Management of 1000 vestibular schwannomas (acoustic neuromas): surgical management and results with an emphasis on complications and how to avoid them. Neurosurgery 1997;40:11–23.
93. Agnelli G, Piovella F, Buoncristiani P, et al. Enoxaparin plus compression stockings compared with compression stockings alone in the prevention of venous thromboembolism after elective neurosurgery. N Engl J Med 1998;339:80–5.
94. Gormley WB, Sekhar LN, Wright DC, et al. Acoustic neuromas: results of current surgical management. Neurosurgery 1997;41:50–60.
95. Ross D, Rosegay H, Pons V. Differentiation of aseptic and bacterial meningitis in postoperative neurosurgical patients. J Neurosurg 1988;69:669–74.
96. Wiegand DA, Ojemann RG, Fickel V. Surgical treatment of acoustic neuroma (vestibular schwannoma) in the United States: report from the Acoustic Neuroma Registry. Laryngoscope 1996;106:58–66.
97. Lalwani AK, Butt FY, Jackler RK, et al. Facial nerve outcome after acoustic neuroma surgery: a study from the era of cranial nerve monitoring. Otolaryngol Head Neck Surg 1994;111:561–70.

98. Irving RM, Jackler RK, Pitts LH. Hearing preservation in patients undergoing vestibular schwannoma surgery: comparison of middle fossa and retrosigmoid approaches. J Neurosurg 1998;88:840–5.
99. Rowed DW, Nedzelski JM. Hearing preservation in the removal of intracanalicular acoustic neuromas via the retrosigmoid approach. J Neurosurg 1997;86: 456–61.
100. Samii M, Matthies C. Management of 1000 vestibular schwannomas (acoustic neuromas): hearing function in 1000 tumor resections. Neurosurgery 1997;40:248–62.
101. Deen HG, Ebersold MJ, Harner SG, et al. Conservative management of acoustic neuroma: an outcome study. Neurosurgery 1996;39:260–6.
102. Glasscock ME, Pappas DG, Manolidis S, et al. Management of acoustic neuroma in the elderly population. Am J Otol 1997;18:236–42.
103. Lunsford LD, Linskey ME. Stereotactic radiosurgery in the treatment of patients with acoustic tumors. Otolaryngol Clin North Am 1992;25:471–91.
104. Flickinger JC, Kondziolka D, Niranjan A, Lunsford LD. Results of acoustic neuroma radiosurgery: an analysis of 5 years' experience using current methods. J Neurosurg 2001;94:1–6.
105. Kondziolka D, Lunsford LD, McLaughlin MR, Flickinger JC. Long-term outcomes after radiosurgery for acoustic neuromas. N Engl J Med 1998; 339:1426–33.
106. Kaylie DM, Horgan MJ, Delashaw JB, McMenomey SO. A meta-analysis comparing outcomes of microsurgery and gamma knife radiosurgery. Laryngoscope 2000;110:1850–6.
107. Driscoll CLW. Vestibular schwannnoma. In: Jackler RK, Driscoll CLW, editors. Tumors of the ear and temporal bone. Philadelphia: Lippincott, Williams and Wilkins; 2000. p. 172–218.
108. Lalwani AK. Meningiomas, epidermoids, and other nonacoustic tumors of the cerebellopontine angle. Otolaryngol Clin North Am 1992;25: 707–28.
109. Storper IS, Glasscock ME, Jackson CG, et al. Management of nonacoustic cranial nerve neuromata. Am J Otol 1998;19:484–90.
110. Samii M, Migliori MM, Tatagiba M, Babu R. Surgical treatment of trigeminal schwannomas. J Neurosurg 1995;82:711–8.
111. Moffat DA, Ballagh RH. Rare tumours of the cerebellopontine angle. Clin Oncol 1995;7:28–41.
112. Thomas NW, King TT. Meningiomas of the cerebellopontine angle. A report of 41 cases. Br J Neurosurg 1996;10:59–68.
113. Irving RM. Meningiomas of the internal auditory canal and cerebellopontine angle. In: Jackler RK, Driscoll CLW, editors. Tumors of the ear and temporal bone. Philadelphia: Lippincott, Williams and Wilkins; 2000. p. 219–35.
114. Valavanis A, Schubiger O, Hayek J, Pouliadis G. CT of meningiomas on the posterior surface of the petrous bone. Neuroradiology 1981;22:111–21.
115. Russell DS, Rubinstein LJ. Pathology of tumors of the nervous system. 4th ed. Baltimore: Williams and Wilkins; 1977.
116. Langford LA. Pathology of meningiomas. J Neurooncol 1996;29:217–21.
117. Matthies C, Carvalho G, Tatagiba M, et al. Meningiomas of the cerebellopontine angle. Acta Neurochir Suppl 1996;65:86–91.
118. Baguley DM, Beynon GJ, Grey PL, et al. Audiovestibular findings in meningioma of the cerebellopontine angle: a retrospective review. J Laryngol Otol 1997;111:1022–6.
119. Hart MJ, Lillehei KO. Management of posterior cranial fossa meningiomas. Ann Otol Rhinol Laryngol 1995;104:105–16.
120. Lalwani AK, Jackler RK. Preoperative differentiation between meningioma of the cerebellopontine angle and acoustic neuroma using MRI. Otolaryngol Head Neck Surg 1993;109:88–95.
121. Tokumaru A, O'Uchi T, Eguchi T, et al. Prominent meningeal enhancement adjacent to meningioma on Gd-DTPA-enhanced MR images: histopathologic correlation. Radiology 1990;175: 431–3.
122. Langman AW, Jackler RK, Althaus SR. Meningioma of the internal auditory canal. Am J Otol 1990;11:201–4.
123. Arriaga M, Shelton C, Nassif P, Brackmann DE. Selection of surgical approaches for meningiomas affecting the temporal bone. Otolaryngol Head Neck Surg 1992;107:738–44.
124. Mirimanoff RO, Dosoretz DE, Linggood RM, et al. Meningioma: analysis of recurrence and progression following neurosurgical resection. J Neurosurg 1985;62:18–24.
125. Schaller B, Heilbronner R, Pfaltz CR, et al. Preoperative and postoperative auditory and facial nerve function in cerebellopontine angle menin-

giomas. Otolaryngol Head Neck Surg 1995;112: 228–34.
126. Kallio M, Sankila R, Hakulinen T, Jaaskelainen J. Factors affecting operative and excess long-term mortality in 935 patients with intracranial meningioma. Neurosurgery 1992;31:2–12.
127. Miralbell R, Linggood RM, de la Monte S, et al. The role of radiotherapy in the treatment of subtotally resected benign meningiomas. J Neurooncol 1992; 13:157–64.
128. Yamakawa K, Shitara N, Genka S, et al. Clinical course and surgical prognosis of 33 cases of intracranial epidermoid tumors. Neurosurgery 1989;24:568–73.
129. Nager GT. Epidermoids involving the temporal bone: clinical, radiological and pathological aspects. Laryngoscope 1975;85:1–21.
130. Abramson RC, Morawetz RB, Schlitt M. Multiple complications from an intracranial epidermoid cyst: case report and literature review. Neurosurgery 1989;24:574–8.
131. Mohanty A, Venkatrama SK, Rao BR, et al. Experience with cerebellopontine angle epidermoids. Neurosurgery 1997;40:24–30.
132. Uchino A, Hasuo K, Matsumoto S, et al. Intracranial epidermoid carcinoma: CT and MRI. Neuroradiology 1995;37:155–8.
133. Samii M, Tatagiba M, Piquer J, Carvalho GA. Surgical treatment of epidermoid cysts of the cerebellopontine angle. J Neurosurg 1996;84:14–9.
134. Haberkamp TJ, Monsell EM, House WF, et al. Diagnosis and treatment of arachnoid cysts of the posterior fossa. Otolaryngol Head Neck Surg 1990; 103:610–4.
135. Schuhmann MU, Ludemann WO, Schreiber H, et al. Cerebellopontine angle lipoma: a rare differential diagnosis. Skull Base Surg 1997;7:199–205.
136. O'Leary MJ, Shelton C, Giddings NA, et al. Glomus tympanicum tumors: a clinical perspective. Laryngoscope 1991;101:1038–43.
137. Olsen WL, Dillon WP, Kelly WM, et al. MR imaging of paragangliomas. AJR Am J Roentgenol 1987;148:201–4.
138. Green JD, Brackmann DE, Nguyen CD, et al. Surgical management of previously untreated glomus jugulare tumors. Laryngoscope 1994;104: 917–21.
139. Lo WW, Shelton C, Waluch V, et al. Intratemporal vascular tumors: detection with CT and MR imaging. Radiology 1989;171:445–8.
140. Shelton C, Brackmann DE, Lo WW, Carberry JN. Intratemporal facial nerve hemangiomas. Otolaryngol Head Neck Surg 1991;104:116–21.
141. Papazian M, Paparella M, Hames E, Frisk J. Aneurysms of the temporal bone. Ear Nose Throat J 1993;72: 474–84.
142. House JL, Burt MR. Primary CNS tumors presenting as cerebellopontine angle tumors. Am J Otol 1985; Suppl:147–53.
143. Ahn MS, Jackler RK. Exophytic brain tumors mimicking primary lesions of the cerebellopontine angle. Laryngoscope 1997;107:466–71.
144. Sabin HI, Bordi LT, Symon L. Epidermoid cysts and cholesterol granulomas centered on the posterior fossa: twenty years of diagnosis and management. Neurosurgery 1987;21:798–805.
145. Brodkey JA, Robertson JH, Shea JJ, Gardner G. Cholesterol granulomas of the petrous apex: combined neurosurgical and otological management. J Neurosurg 1996;85:625–33.
146. Gay E, Sekhar LN, Rubinstein E, et al. Chordomas and chondrosarcomas of the cranial base: results and follow-up of 60 patients. Neurosurgery 1995; 36:887–97.
147. Brooks JJ. Immunohistochemistry of soft tissue tumors: progress and prospects. Hum Pathol 1982; 13:969–74.
148. Yuh WT, Mayr-Yuh NA, Koci TM, et al. Metastatic lesions involving the cerebellopontine angle. AJNR Am J Neuroradiol 1993;14:99–106.

CHAPTER 21

Presbyacusis

John H. Mills, PhD, Paul R. Lambert, MD

Presbyacusis is defined generally as the hearing loss associated with increasing chronologic age. It is a common problem today and will become even more significant in the future. Currently, according to the National Center for Health Statistics, some 20 million Americans have impaired hearing, and approximately 75% of these people are over age 55. In the 25-year period between 1976 and 2000, the number of persons below age 75 increased by 23%, the number between age 75 and 84 increased by 57%, and the number over age 84 increased by 91%. Indeed, by the year 2030, the elderly will comprise 32% of the population, an increase of 250%. As many as 60 to 75% of these older persons will have clinically significant hearing loss. Because of the increasing number of older Americans, high-prevalence problems such as age-related hearing loss and other chronic disabling conditions (balance, vision, arthritis) will place extensive and novel demands on the health care system.

BACKGROUND

The medical-scientific history of presbyacusis is relatively recent and rich. Although Toynbee in 1849 was perhaps the first to write about age-related hearing loss and to prescribe a treatment (application of solutions of silver nitrate or mercurous chloride to the external auditory canal),[1] Zwaardemaker in 1891 is credited with the first accurate description of presbyacusis.[2] With the clever use of Dalton whistles, Zwaardemaker demonstrated that older persons had far more difficulty detecting high-pitched sounds than younger persons. By the use of bone-conduction testing, he correctly considered age-related hearing loss to be of cochlear origin. Since the 1930s, anatomic and histopathologic changes have been observed at nearly every location in the aging auditory system from the external ear with collapsing ear canals to neural degeneration extending from the auditory nerve to the auditory brainstem and temporal lobe. Before reviewing the pertinent literature on presbyacusis, it is necessary to review terminology.

TERMINOLOGY

There are several definitions of the term presbyacusis. Indeed, even the spelling is an issue, that is, presbyacusis versus presbycusis. Here we use "presbycusis." It is used loosely to describe a seemingly endless list of genetic, environmental, and disease states that can cause hearing loss in an older person. Often presbyacusis is used to refer to hearing loss purely caused by the natural process of aging. More often, hearing loss is called presbyacusis if the person is beyond the fifth decade, without consideration of disease, genetics, or other factors. In many studies, older persons are considered homogeneous and are grouped regardless of hearing levels or are grouped if their hearing loss exceeds an arbitrary criterion.

Some of the confusion associated with the use of the term can be eliminated or reduced by the appropriate use of the terms presbyacusis, socioacusis, and nosoacusis.[3] A generic definition of presbyacusic is age-related hearing loss (or threshold shift) that is the effect of aging in combination with lifelong exposures to nonoccupational noise, ototoxic agents, diet, drugs, and other miscellaneous factors. A more restrictive, precise definition of presbyacusis is hearing loss that increases as a function of chronologic age and is attributable to "aging" per se. This "purely aging" hearing loss probably has a genetic basis. It may be correlated with age-related deterioration or declines in other senses, especially balance, vision, and touch.

Socioacusis is defined as the hearing loss produced primarily by exposure to nonoccupational noise in combination with lifestyle factors such as

diet and exercise. Nosoacusis is the hearing loss attributable to diseases with ototoxic effects. Thus, the hearing loss assessed in a person into the fifth decade and beyond is the combined result of aging, nosoacusis, socioacusis, and possibly, for many persons, exposure to occupational noise. It is important for medical and legal reasons to differentiate presbyacusis from socioacusis, nosoacusis, and occupational hearing loss.

ADDITIVITY OF PRESBYACUSIS, SOCIOACUSIS, AND NOSOACUSIS

Most laboratory (with animals) and field studies indicate that hearing losses caused by exposure to noise are additive (in decibels) with the hearing loss attributed to presbyacusis (in the strict sense of the word).[4] A small sensorineural hearing loss (SNHL) of 25 dB at age 25, for example, seemingly has little social or medical relevance; however, by age 70, a hearing loss of 25 dB caused almost totally by the aging process is added to the existing SNHL. The result is a moderate-to-severe SNHL of 50 dB. In other words, a seemingly minor hearing loss at a young age becomes a severe loss when the effects of presbyacusis become evident.

Rules for combining hearing losses attributable to presbyacusis, nosoacusis, and socioacusis are not always straightforward. In the medicolegal assessment of hearing loss, it is assumed that presbyacusic effects (aging plus socioacusis plus nosoacusis) add in decibels to the hearing loss produced by noise. As stated above, this approach is supported by some laboratory and field data for groups of subjects under a limited set of exposure conditions; however, for individual subjects and for complicated noise exposures, the procedures for allocation of hearing loss into different components are controversial.[4]

EPIDEMIOLOGY

Over the past 25 years or so, there has been a significant epidemiologic effort to describe hearing levels as a function of chronologic age. Two sets of data have survived extensive scrutiny and are now part of an international standard, International Standards Organization (ISO) 1999, "Acoustics: Determination of Occupational Noise Exposure and Estimation of Noise-Induced Hearing Impairment." Figure 21–1, A, shows age-related permanent threshold shifts at audiometric frequencies for males and females. These data are referred to as Data Base A in the ISO 1999 standard and are considered to represent highly screened subjects. That is, the data in Figure 21–1, A, represent "pure aging" effects as well as socioacusis. Most of the hearing loss attributable to occupational hearing loss has been eliminated, although a small nosoacusic effect may be present. Figure 21–1, B, gives epidemiologic data, referred to as Data Base B in the ISO 1999 standard, which are considered to reflect "pure aging" effects, socioacusis, some nosoacusis, and some effects attributable to occupational hearing loss. As shown in Figure 21–1, hearing loss in the higher frequencies is measurable by age 30; it increases systematically to age 60 (and beyond), is largest at 4 and 6 kHz, and is much larger in males than in females. There are also significant differences between Figure 21–1, A (Data Base A), and Figure 21–1, B (Data Base B), presumably owing to differences in subject selection. Small effects caused by occupational noise and nosoacusis were eliminated in Data Base A. Significant debate exists currently about the appropriateness and validity of Data Base A and B, particularly in a medicolegal context involving the assessment of occupational noise-induced hearing loss.

Data from Figure 21–1, A and B, have been replotted in Figure 21–2 to show the average hearing loss at 0.5, 1, 2, and 3 kHz as a function of chronologic age. The mean hearing loss at these particular audiometric test frequencies is especially meaningful because this combination of test frequencies comprises those used in the computation of hearing handicap as recommended by the American Academy of Otolaryngology-Head and Neck Surgery (AAO). As Figure 21–2 shows, hearing loss increases systematically from age 20 through age 75. There are at least two remarkable features to Figure 21–2. One is that the largest difference between the worst case (males, Data Base B) and best case (females, Data Base A) is about 5 dB. One could suggest, perhaps strongly, that much of the debate about Data Base A and B is "much ado about nothing," that is, 5 dB. This line of thinking is supported by the fact that, even at age 75, the hearing loss is only about 20 dB. According to the AAO 1979 definitions of hearing handicap, the average hearing loss at 0.5, 1, 2, and 3 kHz must exceed 25 dB to be considered a hearing handicap. In other words, according to the data of

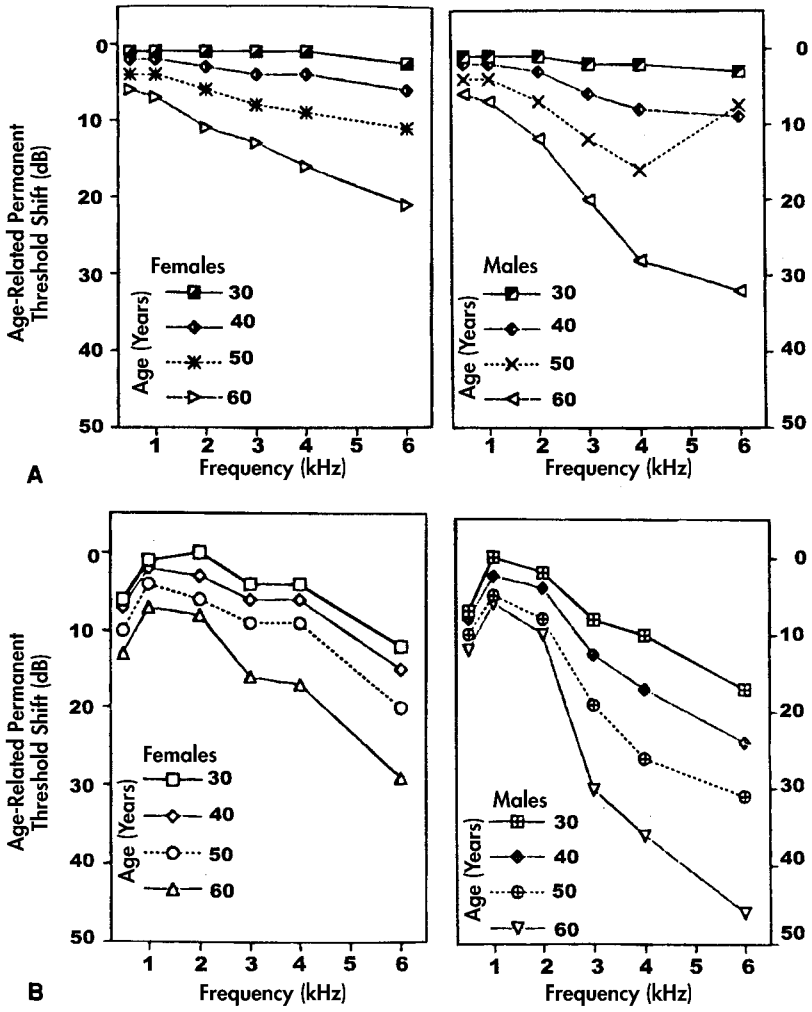


FIGURE 21–1. Epidemiologic data showing hearing levels and chronologic age from an international standard, ISO 1999. The top panel, *A*, is referred to as Data Base A and represents a sample that was highly screened to eliminate the potential influence of the effects of noise. The bottom panel, *B*, is also from ISO 1999 and is referred to as Data Base B. This sample is less well screened than *A* and may contain data from persons with a history of exposure to noise.

Figure 21–2 and the AAO 1979 definition of hearing handicap, substantially less than 50% of the population, male or female, have a hearing handicap even at age 75.

Other less selective epidemiologic data suggest that the prevalence of hearing handicap among older persons is substantially higher than that indicated by the epidemiologic data used in the ISO 1999 standard. Using 1,662 subjects from the famous Framingham study of cardiovascular disease, Gates et al reported that 55% met or exceeded the AAO 1979 medicolegal definition of hearing handicap.[5] The difference between the results of Gates et al and those of ISO 1999 reflect sampling differences and the inclusion of subjects older than age 75 in the Gates et al sample.

In addition to an audiologic assessment, hearing handicap can be assessed more subjectively using questionnaires such as the Hearing Handicap Inventory for the Elderly.[6] The correspondence between objective, audiologic measures of hearing handicap and subjective measures from self-reports and questionnaires suggests that the AAO measure of handicap tends to overestimate handicap for older persons with mild-to-moderate hearing losses and to underestimate handicap for older persons with severe hearing losses. Of course, there are many reasons for discrepancies between subjective and audiologic estimates of hearing handicap. One general rule emerges from questionnaires and epidemiologic studies of hearing handicap in older persons, namely, that the pure-tone average (PTA) of 0.5, 1, 2, and 3 kHz must exceed about 30 dB before most older persons consider themselves to have a hearing problem. Thus, there are indeed millions of older hearing-handicapped persons in the United States.

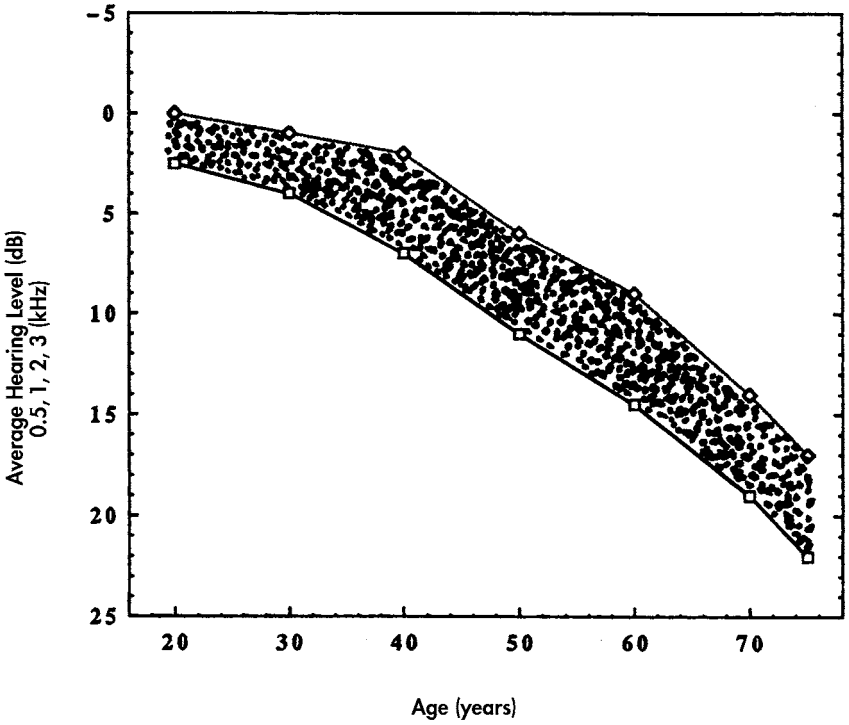


FIGURE 21–2. Data from Figure 21–1 have been replotted to show the average hearing level at 0.5, 1, 2, and 3 kHz as a function of age. The *shaded area* represents the difference between the best case, that is, the women of Data Base A, and the worst case, that is, the men of Data Base B.

Although much of the discussion has centered on the audiometric frequencies from 0.5 to 3 kHz, one of the more dramatic features of age-related hearing loss is the decline of auditory sensitivity at higher frequencies.[7,8] Figure 21–3, A, shows increasing hearing loss from age 20 to 29 to age 50 to 59 for test frequencies at and above 8 kHz. By age 50 to 59, the hearing loss at 16 kHz is greater than 60 dB. Figure 21–3, B, shows longitudinal threshold changes (the same group of observers) over the age range of 70 to 81 years, including thresholds for test frequencies up to 8 kHz. Hearing levels at 4 and 8 kHz exceed 50 dB by age 70 and at 8 kHz exceed 70 dB by age 79. Hearing loss at frequencies above 8 kHz is even more pronounced and is not predictable by hearing levels in the 1 to 4 kHz range.[9] Clearly, there is a dramatic age-related decline in auditory sensitivity at high frequencies. It appears that hearing loss starts in the second decade (or earlier) in the frequency range between 16 and 20 kHz, proceeds systematically in magnitude in the 16 kHz region, and spreads systematically downward in the frequency domain. By the eighth decade, the hearing loss is moderate to severe, even at 1 and 2 kHz.

The interpretation of epidemiologic data that show a loss of high-frequency sensitivity at 16 to 20 kHz, when 10 and 19 year olds are compared with 20 to 29 year olds, is not straightforward. One point of view is that these age-related changes in hearing thresholds, as well as systematic declines in the number of outer hair cells starting at birth and continuing throughout the lifetime of the individual,[10] are genetically determined, age-dependent events. That is, these age-related events are totally endogenous in origin. They are independent of exogenous factors such as diet, exposure to environmental noise (socioacusis), and disease (nosoacusis). As part of the Framingham heart study, Gates and colleagues have estimated the role of genetics in age-related hearing loss.[11] Their heritability estimates suggest that as much as 55% of the variance associated with age-related hearing loss is attributable to the effect of genes. These heritability estimates of age-related hearing loss are similar in magnitude to those reported for hypertension and hyperlipidemia, are much stronger in women than in men, and can be used to support the point of view that age-related changes in the ear and hearing reflect the combined effect of genetics, socioacusis, and nosoacusis. In some views of presbyacusis, socioacusis is given a major role, almost surely because of the famous Mabaan study.[12]

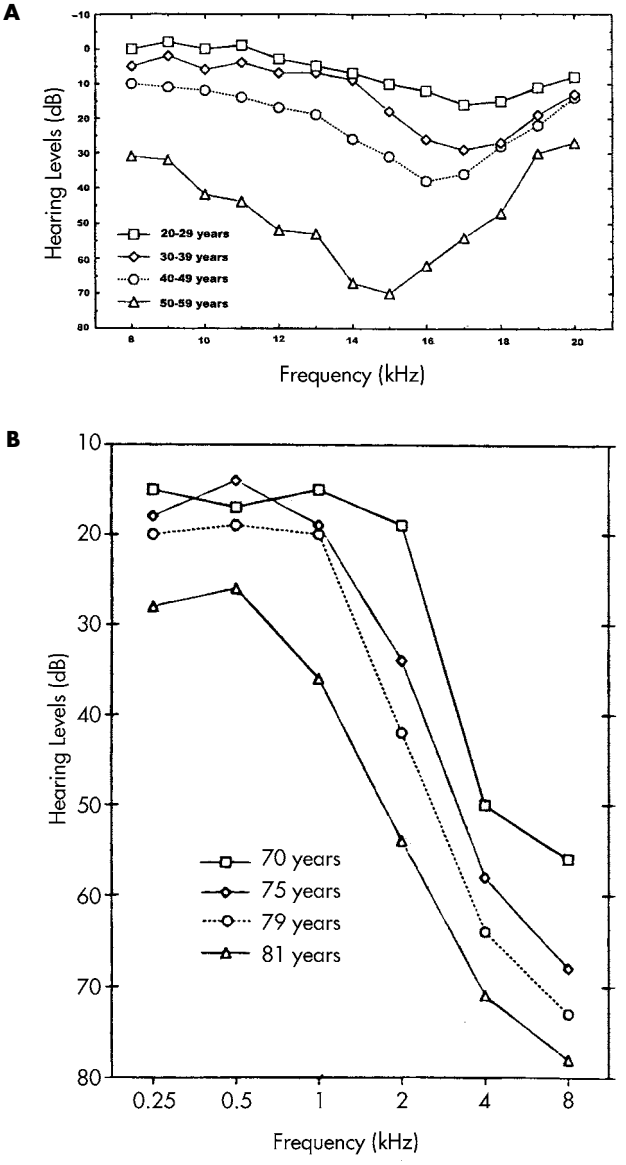


FIGURE 21–3. Hearing loss at high frequencies. *A*, Hearing levels above 8 kHz over the age range 20 to 59 years. *B*, Hearing levels from 0.25 to 8 kHz covering the age range 70 to 81 years, measured longitudinally. Adapted from Stelmachowicz PG et al[7] and Moller MB.[8]

A hearing survey of Mabaans, a tribe located in a remote and undeveloped part of Africa, showed exceptionally good auditory sensitivity for males and females in their sixth to ninth decades. In addition to the lack of exposure to occupational noise, no exposure to firearms, and exposure only to low levels of environmental noise, the Mabaans were reported to lead a low-stress lifestyle, have a low-fat diet, and have a low prevalence of cardiovascular disease. This study was quoted widely in lay and professional publications and became the scientific basis for the thesis that persons in Western industrialized societies were at risk of deafness because of socioacusis. Later analysis of the Mabaan data indicated few differences between hearing levels of Mabaans and hearing levels of well-screened individuals from industrialized societies. One possible exception was that the hearing levels of Mabaan males over the age of 60 were slightly better in the high frequencies than their Western male counterparts. Additional criticism of the original Mabaan data involved the accuracy of the age of the Mabaan subjects. Apparently, the exact ages of the Mabaan subjects could not be confirmed. Lastly, important epidemiologic details of the Mabaan sample could not be determined; thus, the validity of the sampling was questionable (ie, a highly biased sample). It is possible, for example, that only the healthiest subjects were available and that these subjects also had the best hearing. Thus, the Mabaan data are equivocal in their support of the thesis of profound effects of socioacusis and nosoacusis on age-related hearing levels.

Whereas the Mabaan data are equivocal in their support of the role of nosoacusis and socioacusis in age-related hearing loss, experiments with laboratory animals clearly support the idea of an age-related hearing loss that is independent of socioacusis, nosoacusis, and exposure to occupational noise. In light of the complexities of the genetic, environmental, and disease-state factors that may affect the hearing of older persons, there are several ongoing efforts to develop animal models of presbyacusis. An appropriate animal model may allow identification and control of the most pertinent variables as well as permit systematic studies of the biologic basis of presbyacusis. It is also important to examine laboratory studies with humans and make comparisons with laboratory animal data when possible.

LABORATORY STUDIES: HUMANS AND ANIMALS

Of the many investigations of presbyacusis, perhaps the most quoted and most extensive source of histopathologic data is the temporal bone studies of Schuknecht and colleagues.[13] From a number of studies of human temporal bones, they have identi-

fied four types or categories of presbyacusis: (1) sensory, characterized by atrophy and degeneration of the sensory and supporting cells; (2) neural, typified by loss of neurons in the cochlea and central nervous system (CNS); (3) metabolic, characterized by atrophy of the lateral wall of the cochlea, especially the stria vascularis; and (4) mechanical, where the inner ear changes its properties with a resulting inner ear conductive hearing loss. Each of these categories was hypothesized to have a characteristic audiometric configuration (ie, sloping, flat) and therefore to be audiometrically identifiable. In subsequent studies, considerable difficulty was encountered in correlating the audiometric configuration with histopathologic observations and in differentiating one type of presbyacusis from another on the basis of the audiometric configuration.[14] The problem is that the audiometric configurations of older persons do not form clearly defined categories, and histopathologic changes at the level of light and electron microscopy are almost always observed at multiple sites in a given aging ear.

In 1993, Schuknecht and Gacek revised the traditional categories of presbyacusis described above.[15] The revision is based on hundreds of observations of human temporal bones and is summarized by the following:

> 1) Sensory cell losses are the least important type of loss in the aged; 2) neuronal losses are constant and predictable expressions of aging; 3) atrophy of the stria vascularis is the predominant lesion of the aging ear; 4) no anatomical correlation for a gradual descending hearing loss …reflect a cochlear conductive loss; and 5) 25% cannot be classified using light microscopy.

This conclusion by Schuknecht and Gacek is important for at least two reasons. One is that the significance of sensory cell losses is de-emphasized in presbyacusis. The second is that the dominance of strial degeneration is emphasized. These two points bring a consensus to human temporal bone results and those obtained from experiments with animals.

In many species of animals, auditory thresholds estimated from electrophysiologic potentials arising from the auditory nerve and brainstem increase as a function of chronologic age.[16,17] The age-related decline in auditory function occurs in animals who are born and reared in acoustically controlled environments where sound levels rarely exceed 40 dBA. In addition to control of the acoustic environment, none of the animals received antibiotics or other drugs. Conductive hearing loss was eliminated as a potential source of the measured hearing loss. Clearly, environmental noise, drugs, disease, and trauma did not have a causative role in the hearing losses observed in the aging animals. Thus, the notion that presbyacusis is not an aging effect per se but the combined effect of socioacusis and nosoacusis is not supported by laboratory experiments with several species of animals.

Perhaps the most dramatic feature of age-related hearing loss in laboratory animals is the variability between animals. For example, in a longitudinal study of aging gerbils, some animals, at the one extreme, had normal or nearly normal auditory sensitivity from 1 to 16 kHz, whereas at the other extreme, some animals did not respond to signals presented at levels of 80 dB SPL. The variability of this magnitude is remarkable given that chronologic age, environment (temperature, humidity, air quality), acoustic history, and diet of the animals were virtually identical. These data and others involving different inbred strains of mice suggest a strong genetic role for age-related hearing loss, at least in laboratory animals.

In addition to a qualitative/quantitative correspondence between age-related hearing loss measured in aging rodents and audiograms of 60- to 70-year-old humans,[16] many of the histopathologic observations made on human temporal bones have been confirmed and extended on experimental animals.[17–19] In aging rodents, there is usually a small loss of sensory (outer) hair cells in the most basal and the most apical regions of the cochlea. Nerve loss, as indicated by loss of spiral ganglion cells, is pronounced throughout the cochlea. The lateral wall of the cochlea, including the stria vascularis, usually shows degeneration originating in both the base and apex and extending to midcochlear regions as the animal ages; however, in some animals, there is a patchy loss of stria vascularis. Thus, in several species of animals, anatomic studies of cochlear material demonstrate at least three of the histopathologic conditions described in humans by Schuknecht and colleagues, namely, sensory, neural, and metabolic presbyacusis.[13] In animals, all three types are usually present in each animal.

As with humans, the most prominent (dominant) type of presbyacusis in laboratory animals

(with the exception of mice, C57) is the metabolic category.[18–23] In addition to age-related, systematic degeneration of the stria vascularis starting in both the apex and base of the cochlea and proceeding to the midcochlear region, preceding and accompanying this degeneration is a loss of the protein Na, K-adenosine triphosphatase. Sometimes the loss of this important protein, which is detectable using immunohistochemical techniques, occurs when the stria vascularis appears normal under light and electron microscopic examination. Thus, a normal-appearing stria vascularis in an older ear is no assurance that the stria vascularis is normal. Accordingly, the prevalence of "metabolic presbyacusis" may prove to be substantially higher in humans when the appropriate immunohistochemical techniques are applied to the study of human temporal bones of the aged.

The stria vascularis and underlying spiral ligament have a prominent role in generating electrochemical gradients and regulating fluid and ion homeostasis in the cochlea. As mentioned above, these tissues undergo age-related changes in a number of species, including humans. In accordance with its name, the stria is heavily vascularized and has an extremely high metabolic rate. Conceivably, alterations in strial microvasculature could compromise cochlear blood flow and ultimately lead to strial degeneration. Changes in blood flow in aging and in noise-exposed ears are inconsistent. Histopathologic studies on aging gerbil, on the other hand, have provided strong evidence for vascular involvement in age-related hearing loss. Morphometric analyses of lateral wall preparations stained to contrast blood vessels (Figure 21–4) have revealed losses of strial capillary area in aged animals.[19,20] The vascular pathology first presented as small focal lesions mainly in the apical and lower basal turns and progressed with age to encompass large regions at both ends of the cochlea. Remaining strial areas were highly correlated with normal microvasculature and with the endocochlear potential (described below).[20] Not surprisingly, areas of complete capillary loss invariably correlated with regions of strial atrophy. Subsequent ultrastructural analysis has revealed a significant thickening of the basement membrane,[21] which is accompanied by an increase in the deposition of laminin and an abnormal accumulation of immunoglobulin, as shown histochemically.[22,23] Thus, considerable support exists for the major involvement of strial microvasculature in age-related degeneration of the stria vascularis; however, the question of what constitutes the initial injury remains unanswered. Although it is tempting to speculate that atrophy of the stria vascularis occurs secondarily to vascular insufficiency resulting from capillary necrosis, the reverse could very well be true.

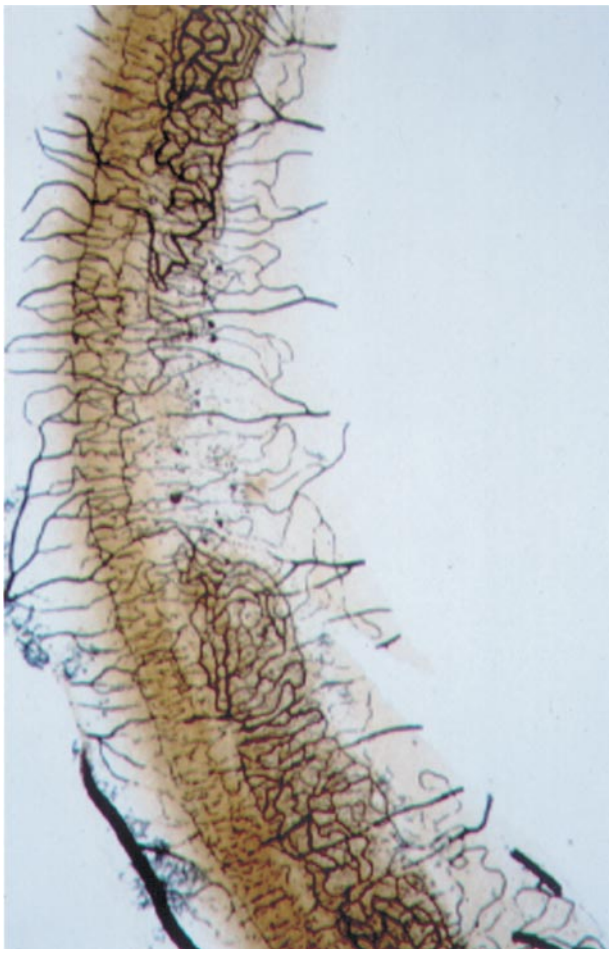


FIGURE 21–4. Surface preparation of stria vascularis/lateral wall dissection from an old gerbil stained with diaminobenzidine for endogeneus peroxidase to contrast with blood vessels (× 100 original magnification). The strial capillary bed overlies vessels of the spiral ligament. The slide shows substantial atrophy of the capillary bed. Adapted from Grattan MA et al.[19]

The most prominent changes in the physiologic properties of the aging ear are losses of auditory nerve function as indicated by increased thresholds of the compound action potential (CAP) of the auditory nerve.[24] Slopes of input-output functions of the CAP in aging animals are decreased even when the loss of auditory thresholds is only 5 to

10 dB. In other words, as the signal intensity is increased, the amplitude of the CAP increases by a fraction of that observed in young animals with normal hearing or young animals with noise-induced or drug-induced hearing losses. These shallow input-output functions of the CAP are also reflected in shallow input-output functions of auditory brainstem responses (ABRs). Thus, what appears to be abnormal function of the auditory brainstem in older animals reflects only the abnormal output of the auditory nerve. The reduced amplitudes of action potentials observed in aging ears are probably reflective of asynchronous or poorly synchronized neural activity in the auditory nerve. The pathologic basis of asynchronized activity in the auditory nerve is unknown but probably involves the nature of the synapse between individual auditory nerve fibers and the attachment to the inner hair cell, primary degeneration of spiral ganglion cells, and reductions in the endolymphatic potential, which are described below.

A second major difference between the presbyacusic ear and the ear with noise-induced hearing loss is the endolymphatic potential, that is, the 80 to 90 mV DC resting potential in the scala media of the cochlea.[18] In aging animals, this DC potential is often reduced significantly and proportionally to losses/degeneration of the stria vascularis. Moreover, in experimental animals, hearing losses can be reduced by introducing a DC voltage into scala media and raising a low endolymphatic potential, that is, 15 mV, to a value approaching 60 to 70 mV. Degeneration of the stria vascularis with the resultant decline in the endocochlear potential has given rise to the "dead battery theory" of presbyacusis.

A third difference between the physiology of the aging ear and the ear with noise-induced or drug-induced injury of the cochlea involves the nonlinear phenomenon of two-tone rate suppression, that is, activity of the auditory nerve elicited by one tone is suppressed or eliminated by the addition of a second tone of different frequency.[24] In ears of quiet-reared aging animals, the mechanism of two-tone rate suppression appears to remain intact. In contrast, in noise-induced or drug-induced injury of the cochlea, a reduction or complete loss of two-tone rate suppression may be the first indicator of injury of the cochlea, usually outer hair cells. These data are an excellent indicator that the pathologic basis of age-related hearing loss is fundamentally different from the pathologic basis of most forms of noise-induced or drug-induced hearing loss.

Otoacoustic emissions, both transient evoked and distortion product, are nonlinear phenomena that are assumed to reflect the integrity of sensory cells, especially outer hair cells. Given this assumption, one would expect to find very high correlations between loss of outer hair cells and changes in otoacoustic emissions; however, inconsistent relations among distortion products, threshold measures, and sensory cell pathology have been reported from widely different types of experiments. Indeed, there are reports of normal emissions in the presence of missing outer hair cells as well as reduced emissions in the presence of a complete complement of outer hair cells.[25] In the aging ear of experimental animals with a minimal, if any, loss of outer hair cells, distortion-product emissions are reduced somewhat in amplitude but are clearly present and robust.[26] In aging human ears, transient emissions are present in about 80% of persons with a PTA hearing level better than 10 dB, present in about 50% with PTA of 11 to 26 dB, and absent in about 80% with a PTA greater than 26 dB. Thus, the presence of a transient otoacoustic emission suggests excellent hearing levels for most persons, whereas its absence reveals very little. Indeed, the major application of otoacoustic emissions may be in hearing screening of infants.

A recent development is the association of mitochondrial deoxyribonucleic acid (mtDNA) deletions with SNHL and age-related hearing loss. Ueda et al, using DNA specimens extracted from peripheral blood leukocytes, found a higher rate of mtDNA deletions in patients with SNHL than in controls.[27] Seidman, in a rat study, tied mtDNA deletions to presbyacusis and mtDNA deletions to reactive oxygen metabolites (ROMs).[28] These are highly toxic molecules that can damage mitochondrial DNA, resulting in the production of specific mtDNA deletions. Thus, compounds that block or scavenge ROMs should attenuate age-related hearing loss. Rats were assigned to treatment groups including controls, caloric restriction, and treatment with several antioxidants, including vitamins E and C, and were allowed to age in a controlled environment. The calorie-restricted groups maintained the best hearing, the lowest quantity of mtDNA deletions in brain and ear tissues, and the least amount of outer hair cell loss. The antioxidant-treated subjects had better hearing than the controls and a slight trend

for fewer mtDNA deletions. The controls had the poorest hearing, the most mtDNA deletions, and the most outer hair cell loss. These data suggest that nutritional and pharmacologic strategies may prove to be an effective treatment that would limit age-related increases in ROM production, reduce mtDNA deletions, and thus reduce age-related hearing loss. Reactive oxygen metabolites (or reactive oxygen species) and oxidative stress have also been implicated in noise-induced hearing loss, ototoxic hearing loss, and cumulative injury that presents as age-related hearing loss. It is speculated that there is a genetic impairment of antioxidant protections that leads to the production of both age-related and noise-induced hearing loss by placing the cochlea in a state of vulnerability.[29,30] Although an age-related hearing loss gene has been identified, the murine Ahl mutation,[31,32] the discovery of the molecular/genetic basis of presbyacusis, noise-induced hearing loss, and SNHL in general are truly just beginning.

In summary, animal experiments conducted under strict conditions show age-related declines in auditory function that indicate a pure aging hearing loss. Most of the pathologic anatomy associated with this "pure aging" hearing loss is in the cochlea, where the dominant pathology is degeneration of the stria vascularis. In experimental animals, most of the hearing loss can be accounted for by anatomic, physiologic, and biochemical changes in the auditory periphery. Indeed, there is no need to include the auditory CNS in an explanation of age-related hearing loss, even when the criterion measure of age-related hearing loss is derived from potentials arising from the auditory brainstem. Nearly all changes observed in auditory brainstem potentials can be explained by alterations in the auditory periphery.

Explanations of age-related changes in auditory brainstem potentials in terms of alterations in the auditory periphery run counter to the current dogma that there is a significant, perhaps dominant, component of presbyacusis that involves degeneration of the auditory CNS. Indeed, there are significant, age-related changes in the auditory CNS involving subtle to gross changes in anatomy and neurochemistry. One prominent change is a reduction in γ-aminobutyric acid (GABA) affecting neural inhibition. Another is increases in the release of aspartate in the cochlear nucleus following cochlear injury and leads to increased excitatory neurotransmission. The role of the CNS in age-related hearing loss is almost surely substantial but remains largely unknown.

PERCEPTION OF AUDITORY SIGNALS AND SPEECH

In addition to age-related declines in auditory sensitivity, there are age-related declines in differential sensitivity for intensity, frequency, and time. Until recently, these age-related declines in the basic properties of the ear and hearing were almost always measured in older persons with significant hearing losses. Thus, it was nearly impossible to separate the effects of a hearing loss from the effects of aging. Recently, however, both discrimination in intensity and discrimination in frequency have been shown to decline with age only at low frequencies and independently of any hearing loss.[33] These results are important because they are negative effects measured in the presence of normal hearing and may very well represent age-related declines in information processing capability. It remains unclear whether these age-related declines represent age-related effects of the auditory periphery or CNS.

The term "phonemic regression" was coined to describe a disproportionate difficulty in speech perception relative to the magnitude of hearing loss of older persons. Later, many studies of speech discrimination and other complex listening tasks showed results with older subjects that were difficult to explain solely on the basis of the audiogram alone. With the ready availability of many published articles showing degenerative changes in the auditory brainstem and cortex of older persons, there evolved a viewpoint that a significant component of age-related hearing loss is attributable to the decline of the auditory CNS. Of course, the auditory CNS is involved in age-related hearing loss; however, the extent of the involvement of the CNS in presbyacusis is currently receiving much debate. Here we examine the CNS issue beginning with data on speech discrimination by older persons.

Perhaps the largest qualitative/quantitative description of speech discrimination as a function of age and hearing loss is provided by Jerger from the clinical records of 2,162 patients.[34] Figure 21–5 shows percent correct on phonetically balanced (PB) words as a function of chronologic age. The parameter on the figure is the PTA of 0.5, 1, and

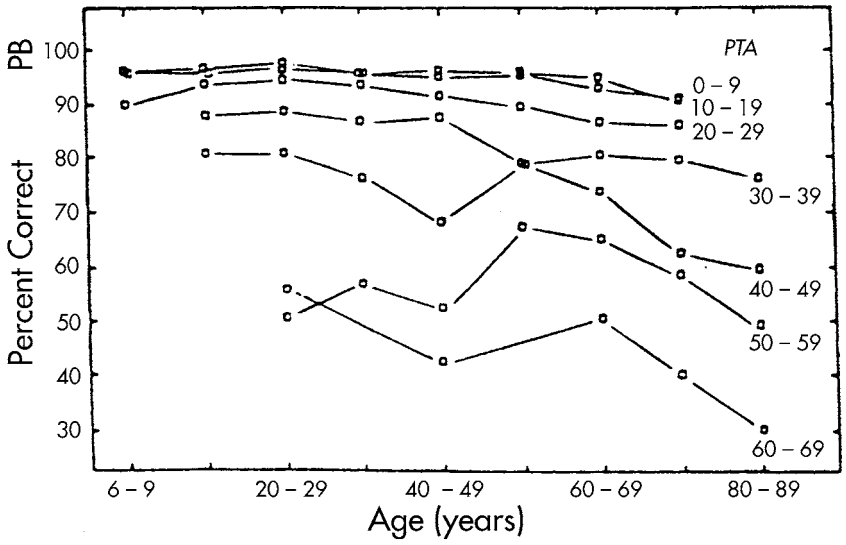


FIGURE 21–5. Speech discrimination as a function of age with average hearing loss at 0.5, 1, and 2 kHz as parameter. PB = phonetically balanced; PTA = puretone average. Adapted from Jerger J.[34]

2 kHz. Figure 21–5 shows that speech discrimination declines systematically as a function of chronologic age; however, the decline with age is dramatically dependent on hearing loss. That is, for subjects with very little hearing loss (less than 30 dB), the decline with age is measurable but small through age 70. On the other hand, for subjects with moderate-to-severe hearing losses, the decline with age is noteworthy, particularly for persons between the ages of 45 and 85 with hearing losses of 40 to 49 dB, 50 to 59 dB, and 60 to 69 dB. In other words, when hearing loss as indicated by the PTA is held constant, speech discrimination scores decreased between the ages of 40 and 80 for moderate (greater than 40 dB) to severe (60 dB) hearing losses. In contrast, age had very little effect as long as the PTA is less than about 30 to 39 dB. A remarkable and perhaps the most noteworthy feature of Figure 21–5 is that persons over the age range of 50 to 80 perform as well as 25 year olds as long as their average hearing loss at 0.5, 1, and 2 kHz is less than 35 dB or so. This fact runs counter to the stereotype of 70- to 80-year-old persons.

Performance on tests of speech discrimination by older persons is shown (Figure 21–6) for PB words, three versions of the Speech Perception in Noise (SPIN) test, and synthetic sentence identification (SSI).[35] Subjects in whom hearing levels were nearly identical (± 3 dB) were placed into three age groups over the range from 55 to 84 years. Statistical analysis of the speech discrimination data showed that performance on all tests were not affected by age. That is, when hearing levels were equated, there were no age-related declines in performance on tests of speech discrimination in persons over the age range of 55 to 84 years. In additional analyses using partial correlations, Dubno et al showed significant gender effects, that is, significant declines with age for males in word recognition, SSI, and SPIN in high-context sentences.[36] Age-related declines were not observed for females.

There are many additional studies of speech discrimination using background noises, degraded speech signals, reverberation, and other variations that resulted in more difficult listening tasks than the usual speech discrimination test, which is done in quiet listening conditions.[36,37] Many of these studies show age-related effects; however, the interpretation of many of these data showing age-related declines in auditory behavior is not straightforward because the subjects usually have significant hearing loss, as indicated by the audiogram. Accordingly, some persons believe that there is a large CNS component to presbyacusis, whereas others believe and have shown that speech discrimination by older persons is predictable given the audiometric hearing loss and the audibility of the speech material. Indeed, as much as 95% of the variance in speech discrimination results can be accounted for on the basis of the audiogram.[38] In other cases for which predictions of performance by older persons under noisy listening conditions are in error by 20 to 25%, it is still not

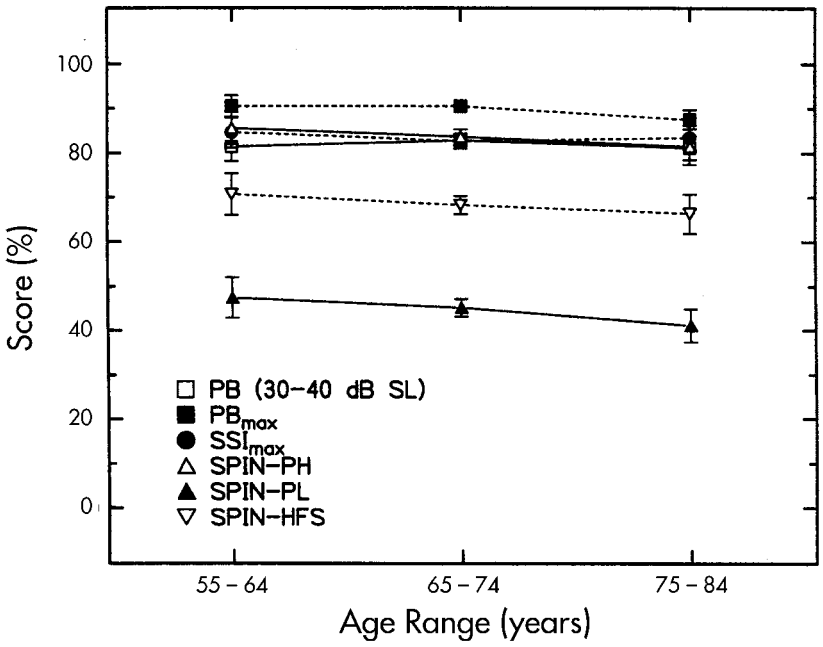


FIGURE 21–6. Speech discrimination as a function of age for different speech material. Hearing level is the same for each age group. PB = phonetically balanced; SSI = synthetic sentence identification; SPIN = Speech Perception in Noise; PH = high probability; PL = low probability. Adapted from Dubno JR et al.[35]

necessary to invoke the involvement of the CNS, although this is usually the interpretation.

Although there are large individual differences in auditory behavior among individuals that are predictable by individual differences in auditory thresholds, there are many age-related declines in auditory behavior that are not strongly associated with auditory thresholds. This is most evident using binaural listening tasks conducted under acoustically stressful conditions. In many binaural experiments using older subjects, age-related declines are observed that appear to be independent of peripheral hearing loss; however, it is often difficult to separate truly central processing disorders from disorders initiated by a pathologic input from the aging cochlea.

AUDITORY EVOKED POTENTIALS

With the development of straightforward techniques to measure evoked potentials arising from the auditory nerve, brainstem, and auditory cortex, there has been significant progress in evaluating the aging auditory nervous system of human subjects. In regard to the ABRs, most studies show age-related declines in the amplitude of wave V, which should be interpreted to reflect peripheral hearing loss rather than changes in the brainstem. Even in older persons with excellent hearing levels, ABR waveforms are of "poor quality" and reduced amplitude, probably reflecting pathology of the cochlea and auditory nerve as well as a reduction in synchronized neural activity. For young subjects with normal and abnormal hearing, auditory thresholds measured behaviorally are about 10 dB better than auditory thresholds estimated from the ABR. For aging subjects, this behavioral/ABR disparity is 20 dB, reflecting the poor quality/low amplitude of the ABR. Other evoked potentials show inconsistent results. The P300 arising from the auditory cortex shows age-related declines, whereas the amplitude-modulated following response may remain unaffected by aging. In contrast, the frequency-modulated following response may be enhanced in older subjects, perhaps reflecting the well-documented age-related decline in GABA. Although much of this research is in progress, it appears that evoked potentials produced by short-duration signals (onset responses such as ABR and CAP) are decreased in amplitude in older subjects, whereas those evoked potentials produced at higher levels in the CNS by long-duration signals may be unaffected or even increased in amplitude. Age-related increases could reflect a number of factors, including efforts by the CNS to compensate for peripheral deficits or age-related changes in the excitatory/inhibitory balance of the auditory CNS.

ALLEVIATION/TREATMENT

Assuming that any problems with the external and middle ear are diagnosed and treated, that other medical issues are under control, and a diagnosis of presbyacusis in the general sense, the best treatment currently available is a hearing aid. The successful use of hearing aids by older persons and the hearing impaired in general is mixed. There is a literature of substantial magnitude that reports on the successful or unsuccessful use of hearing aids by the elderly. Many older persons who would clearly benefit from an aid do not use one. The reasons for not using an aid or being a dissatisfied user include cost, stigma of a hearing handicap and of being old, difficulty in manipulating controls, and too little benefit, particularly in the presence of background noise.

In a large group of older persons (N = 516) who are participants in a longitudinal study of age-related hearing loss, 53% (n = 272) are candidates for a hearing aid. Candidacy by very conservative audiologic criteria is a speech reception threshold greater than 30 dB in the better ear or hearing level greater than 40 dB at 3 and 4 kHz in the better ear. Using these criteria, nearly half (n = 131 of 272) of those who are considered to be excellent candidates have never tried a hearing aid, and only 38% (n = 104 of 272) of candidates were successful hearing aid users. Thirty-seven persons, for one or several reasons, were dissatisfied hearing aid users. Of those who did not meet the conservative criteria for a hearing aid (n = 244), our clinical judgment suggested that at least 40 to 50% of these persons would benefit from an aid. Clearly, the older persons in our longitudinal study were uninformed about hearing aids, poorly served, or underserved. Indeed, it is our opinion that as a result of excellent advances in hearing aid technology in combination with improved fitting techniques, nearly every older hearing-impaired person should be considered a potential candidate for a hearing aid.

ACKNOWLEDGMENT

Preparation of this chapter was supported by NIH (P01 DC00422). Lois J. Matthews provided data on hearing aid use and transient otoacoustic emissions. Nancy Smythe assisted in preparation of the figures and final document.

REFERENCES

1. Toynbee J. On the pathology and treatment of the deafness attendant upon old age. Mon J Med Sci 1849;1–12.
2. Zwaardemaker H. Der Verlust on hohen Torten mit zunehmendern Alter. Ein neues Gesetz. Arch Ohrnas Kehlkopfheil 1891;32:53–6.
3. Ward WD. Effects of noise exposure on auditory sensitivity. In: Neff DW, editor. Handbook of physiology. Section 9: reactions to environmental agents. Bethesda (MD): American Physiology Society; 1977. p. 1–15.
4. Dobie RA. Medical-legal evaluation of hearing loss. Albany (NY): Singular Thomson Learning; 2001.
5. Gates GA, Cooper JC, Kannel WB, Miller NJ. Hearing in the elderly: the Framingham cohort, 1983–85. Ear Hear 1990;11:247–56.
6. Matthews LJ, Lee FS, Mills JH, Schum DJ. Audiometric and subjective assessment of hearing handicap. Arch Otolaryngol Head Neck Surg 1990; 116:1325–30.
7. Stelmachowicz PG, Beauchaine KA, Kalberer A, Jesteadt W. Normative thresholds in the 8- to 20-kHz range as a function of age. J Acoust Soc Am 1989; 86:1384–91.
8. Moller MB. Hearing in 70 and 75 year old people: results from a cross-sectional and longitudinal population study. Am J Otolaryngol 1981;2:22–9.
9. Matthews LJ, Lee FS, Mills JH, Dubno JR. Extended high-frequency thresholds in older adults. J Speech Lang Hear Res 1997;40:208–14.
10. Bredberg G. Cellular pattern and nerve supply of the human organ of corti. Acta Otolaryngol Suppl (Stockh) 1968;236:1–135.
11. Gates GA, Couropmitree NN, Myers RH. Genetic associations in age-related hearing thresholds. Arch Otolaryngol Head Neck Surg 1999;125:654–9.
12. Rosen S, et al. Presbycusis study of a relatively noise-free population in the Sudan. Ann Otol Rhinol Laryngol 1962;71:727–37.
13. Schuknecht HF. Pathology of the ear. Cambridge (MA): Harvard University Press; 1974.
14. Suga F, Lindsay JR. Histopathological observations of presbyacusis. Ann Otol Rhinol Laryngol 1976; 85:169–76.
15. Schuknecht HF, Gacek MR. Cochlear pathology in presbycusis. Ann Otol Rhinol Laryngol 1993;102 Suppl 158:1–16.

16. Mills JH, Schmiedt RA, Kulish LF. Age-related changes in auditory potentials of mongolian gerbil. Hear Res 1990;46:301–10.
17. Willcott JF. Aging and the auditory system. In: Anatomy, physiology, and psychophysics. San Diego: Singular Publishing Group; 1991.
18. Schulte BA, Schmiedt RA. Lateral wall Na, K-ATPase and endocochlear potentials decline with age in quiet-reared gerbils. Hear Res 1992;61:35–46.
19. Gratton MA, Schulte BA. Alterations in microvasculature are associated with atrophy of the stria vascularis in quiet-aged gerbils. Hear Res 1995; 82:44–52.
20. Gratton MA, Schmiedt RA, Schulte BA. Age-related decreases in endocochlear potential are associated with vascular abnormalities in the stria vascularis. Hear Res 1996;102:181–90.
21. Thomopoulos GN, Spicer SS, Gratton MA, Schulte BA. Age-related thickening of basement membrane in stria vascularis capillaries. Hear Res 1997;111: 31–41.
22. Sakaguchi N, Spicer SS, Thomopoulos GN, Schulte BA. Increased laminin deposition in capillaries of the stria vascularis of quiet-aged gerbils. Hear Res 1997;105:44–56.
23. Sakaguchi N, Spicer SS, Thomopoulos GN, Schulte BA. Immunoglobulin deposition in thickened basement membranes of aging strial capillaries. Hear Res 1997;108:83–91.
24. Schmiedt RA, Mills JH, Adams JC. Tuning and suppresion in auditory nerve fibers of aged gerbils raised in quiet or noise. Hear Res 1990;45:221–36.
25. Hamernick R, Qiu W. Correlations among evoked potential thresholds, distortion product otoacoustic emissions and hair cell loss following various noise exposures in the chinchilla. Hear Res 2000;150:245–57.
26. Boettcher FA, Gratton MA, Schmiedt RA. Effects of age and noise on hearing. Occup Med 1995; 10:577–91.
27. Ueda N, Oshima T, Ikeda K, et al. Mitchochondrial DNA deletion is a predisposing cause for sensorineural hearing loss. Laryngoscope 1998;108: 580–4.
28. Seidman MD. Effects of dietary restriction and antioxidants on presbyacusis. Laryngoscope 2000; 110:727–38.
29. Ohlemiller KK, McFadden SL, Ding D-L, et al. Targeted mutation of the gene for cellular glutathione peroxidase (*Gpx1*) increases noise-induced hearing loss in mice. J Assoc Res Otolaryngol 2000;1:243–54.
30. McFadden SL, Ding D, Burkard RF, et al. Cu/Zn SOD deficiency potentiates hearing loss and cochlear pathology in aged 129, CD-1 mice. J Comp Neurol 1999;413:101–12.
31. Li HS. Genetic influences on susceptibility of the auditory system to aging and environmental factors. Scand Audiol 1992;Suppl 36:1–39.
32. Erway LC, Willott JF. Genetic susceptibility to noise-induced hearing loss. In: Axelsson A, editor. A scientific basis of noise-induced hearing loss. New York: Thieme; 1996. p. 56–64.
33. He NJ, Dubno JR, Mills JH. Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects. J Acoust Soc Am 1998;103:553–65.
34. Jerger J. Audiological findings in aging. Adv Otorhinolaryngol 1973;20:115–24.
35. Dubno JR, Lee FS, Matthews LJ, Mills JH. Age-related and gender-related changes in monaural speech recognition. J Speech Lang Hear Res 1997;40:444–52.
36. Dubno J, Dirks D, Morgan D. Effects of age and mild hearing loss on speech recognition in noise. J Acoust Soc Am 1984;76:87–96.
37. Schum DJ, Matthews LJ, Lee FS. Actual and predicted word recognition performance of elderly hearing impaired listeners. J Speech Hear Res 1991;34:636–42.
38. Humes LE, Christopherson L, Cokely C. Central auditory processing disorders in the elderly: fact or fiction? In: Katz J, Stecker N, Henderson D, editors. Central auditory processing. St. Louis: Mosby Year Book; 1992. p. 141–9.

CHAPTER 22

Tinnitus and Hyperacusis

Pawel J. Jastreboff, PhD, ScD, Margaret M. Jastreboff, PhD

TINNITUS

Tinnitus is commonly described as a perception of sound that is not related to an external acoustic source or electrical stimulation.[1] It is an extremely common condition, but only a fraction of those who experience tinnitus are significantly disturbed.[2–4] Moreover, it has been shown that perception of tinnitus can be evoked in 94% of young, healthy subjects by putting them in a sufficiently low level of sound.[5] Although many unanswered questions regarding tinnitus remain and there is a lack of objective measures and no clear agreement as to the most efficient way to help those suffering with tinnitus, significant progress has been made over the last 20 years in the tinnitus field. Still, most patients are left with proverbial advice: "Learn to live with it."

Tinnitus is not a disease. It is a symptom that, similar to pain, headache, or fever, can vary in severity and can affect patients' lives to varying degrees. Sounds described by patients have different spectra and loudness, can change in loudness and type of sound, and can persist or be transient.[1,2] Tinnitus can be annoying to those who experience it and lead to a vicious circle whereby the tinnitus becomes the center of attention in patients' lives. Additionally, tinnitus can exist independently or as part of a complex medical condition.[6] There is no truly tracked etiology, specific site, or molecular/cellular mechanism that has been proven. Tinnitus affects people of all ages.

Frequently, tinnitus is accompanied by decreased sound tolerance and hearing loss.[4,6–8] Decreased sound tolerance can represent hyperacusis and/or misophonia. There is no consensus regarding the testing of loudness discomfort levels (LDLs) to evaluate the problem, and there are no established normative data. The prevalence and epidemiology of hyperacusis are not well documented, and its etiology and mechanisms are poorly understood. Hyperacusis can occur alone or as an adjunct to complex medical conditions. Gradual desensitization can lead to the successful treatment of the problem.

Definitions

Writings about tinnitus can be found in ancient documentation of Babylonian, Egyptian, Greek, Indian, and Assyrian medicine. Throughout the world, a variety of terms have been used to describe a ringing, tinkling (Latin, *tinnire*), buzzing, whistling, or unpleasant sound (French, *acouphnes*) in the ears or in the head, leaving us with the two most commonly used terms, tinnitus (in English) and *acufenos* (in Spanish). The historical review of tinnitus may be found in a number of publications, particularly those by Stephens[9] and Feldmann.[10,11]

There is no precise, short, and distinctive definition of tinnitus. Today, commonly used definitions of tinnitus focus on different characteristics. Patient experience–oriented definition describes tinnitus as ringing, buzzing, the sound of escaping steam, hissing, humming, cricket-like, or noise in the ears.[1,4] Tinnitus as "a phantom auditory perception" represents a physiologic definition of tinnitus pointing out a lack of a physical acoustic stimulus related to tinnitus.[12,13] There is also the definition proposed by the Committee on Hearing, Bioacoustics and Biomechanics of the US National Research Council: "a conscious experience of sound that originates in the head" of its owner.[1]

Etiology and Prevalence

The results of studies conducted in numerous places around the world have shown a significant variability in the estimation of tinnitus prevalence in the general population.[2–4,14–18] Six to 17% of people experience tinnitus lasting for a period of at least 5 minutes. About 3 to 7% of the general population seek help for their tinnitus, and 0.5 to 2.5% report

severe effects of tinnitus in their lives.[1–4,18] The prevalence of tinnitus in adults with hearing problems is very high (59 to 86%), and it is estimated that tinnitus is present in 50% of patients with sudden hearing loss, 70% with presbycusis, and 50 to 90% with noise-induced hearing loss.[19]

The prevalence of tinnitus increases significantly with aging, but people of all ages experience tinnitus.[2,4,8] Although frequently not reported, children are also affected by tinnitus, and the estimated prevalence is similar to that reported in adults. Nodar, in a sample of 2,000 children, reported the average prevalence of 15%, with 13.3% for children with normal hearing and 58.6% with hearing loss.[20] Similar data were reported by others with a general prevalence in the range of 15 to 29% in healthy children and approximately 50% in children with otologic problems or hearing loss.[21,22] A significant proportion of children reported having problems with tinnitus including sleep disturbance (42%), problems with concentration (47%), and sensitivity to sound (33%).[23] In younger healthy children (age 5 to 16 years), 29% have tinnitus, and 9.6% reported their tinnitus as troublesome.[21]

Extensive studies have been performed in an attempt to link various factors with tinnitus prevalence.[1,2,8,14] Hearing loss, specifically the extent of high-frequency impairment in the worse ear, is one of the main predicting factors for tinnitus.[2] Conductive hearing loss seems to be a separate factor,[24] and noise exposure has been correlated with tinnitus as well.[14] Tinnitus is also experienced by those with normal hearing; 18% of tinnitus patients were reported to have normal hearing.[24]

Other epidemiologic factors do not appear to be correlated with tinnitus. There is no clear effect of gender on tinnitus prevalence, although pregnancy has been shown to increase significantly the probability of tinnitus.[25] Neither smoking, coffee, nor alcohol has been shown to increase tinnitus prevalence directly.[26–28]

Mechanisms and Models

Our knowledge of the mechanisms of tinnitus is still limited and based more on theoretical speculations than on strong research data or stringent clinical studies. Past models were focused on peripheral mechanisms,[29–35] whereas recent models tend to involve or even focus on processing information within the central auditory pathways and central nervous system.[12,36–38] The neurophysiological model of tinnitus combines all of the levels and differentiates between the perception of tinnitus versus tinnitus-induced activation of nonauditory structures in the brain.[7,12,37–39] Table 22–1 summarizes the approaches to the mechanisms of tinnitus.[19,29,37,40–44]

Tinnitus as a Symptom of Medically Treatable Diseases

Tinnitus may also be a part of more complex medical conditions. The most common ones are listed in Table 22–2.[37,45]

DECREASED SOUND TOLERANCE

Tinnitus is frequently accompanied by decreased sound tolerance (oversensitivity to sound),[3,6,12,17,46–51] which, in many cases, is a sum of hyperacusis and misophonia.[46,52]

Definitions

There is no generally accepted definition for decreased sound tolerance to suprathreshold sounds, although a variety of terms have been proposed, with hyperacusis used most frequently. According to *Stedman's Medical Dictionary*,[53] hyperacusis is defined as "Abnormal acuteness of hearing due to increased irritability of the sensory neural mechanism. Syn: auditory hyperesthesia," with hyperesthesia defined as "Abnormal acuteness of sensitivity to touch, pain, or other sensory stimuli" or, according to *American Heritage Dictionary*, as "An abnormal or pathological increase in sensitivity to sensory stimuli, as of the skin to touch or the ear to sound."[54] It has been recognized that decreased sound tolerance might reflect physical discomfort or be related to a fear of sound.[6]

We proposed the approach to decreased sound tolerance based on neurophysiology, recognizing the systems that can be involved, the auditory system (both peripheral and central parts) and the limbic and autonomic systems, and consequently propose the following definitions. Hyperacusis can be defined as abnormally strong reactions to sound occurring within the auditory pathways. At the behavioral level, it is manifested by a subject experiencing physical discomfort as a result of exposure to

moderate/weak sound that would not evoke such reaction in the average population. The strength of a reaction is linked to the physical characterization of a sound, for example, its spectrum and intensity. Misophonia and phonophobia can be defined as abnormally strong reactions of the autonomic and limbic systems without abnormally high activation of the auditory system, resulting from enhanced connections between the auditory and limbic systems. At the behavioral level, patients have a generally negative attitude to sound (misophonia, from Greek: *miso* meaning strong dislike, hate) or could

TABLE 22–1. Main Concepts Related to Proposed Mechanisms of Tinnitus

Structures involved
Auditory system
Periphery (cochlea, auditory nerve)
Central auditory pathways
Auditory and central nervous system
Manifestation of tinnitus-related neuronal activity
Increase in spontaneous activity
Modification in temporal patterns of discharges, including bursting activity
Synchronization of the activity between neurons
Proposed mechanisms responsible for the emergence of tinnitus-related neuronal activity
Abnormal coupling between neurons
Local decrease of spontaneous activity enhanced by lateral inhibition
Discordant damage/dysfunction of outer and inner hair cells
Unbalanced activation of Type I and II auditory nerve fibers
Abnormal neurotransmitter release from inner hair cells
Decreased activity of the efferent system
Mechanical displacement within the organ of Corti
Abnormalities in transduction processes
Various aspects of calcium function
Physical/biochemical stress on the auditory nerve
Enhanced sensitivity of the auditory pathways after decreased auditory input
Level of interest
Molecular—ion channels, synapses, cellular membranes
Single neuron—processing information within one cell
Neuronal assemblies—interaction within group of cells
System—interaction between various systems in the brain
Somatosounds
Included as "objective tinnitus"
Separated, with name tinnitus reserved to auditory phantom perception

Adapted from Spoendlin H;[19] Tonndorf J;[29,41] Jastreboff PJ;[37,44] Pujol R;[40] Zenner H-P and Ernst A;[42] LePage EL.[43]

specifically be afraid of sound (phonophobia; phobia, fear). The strength of a reaction will be only partially determined by the physical characterization of a sound and will depend on the patient's previous evaluation of a sound (eg, potential threat, beliefs that it can be harmful), the patient's psychological profile, and the context in which the sound is presented.

Note that neither hyperacusis nor phonophobia has any relation to the threshold of hearing, which can be normal or can reflect hearing loss.

Most frequently, decreased sound tolerance results from a combination of hyperacusis and phonophobia. It is important to assess the presence and the extent of both phenomena in a patient as each needs to be treated using different methods. Terms used in the literature do not differentiate these problems, and since the reported decreased sound tolerance is typically dominated by hyperacusis, we will use the term hyperacusis in describing reports in the literature. As misophonia is a new term, and phonophobia is commonly used, in this chapter, when the word "phonophobia" is used, it actually denotes "misophonia/phonophobia."

Additionally, note that the term recruitment is not related to a decreased sound tolerance or hyperacusis. Recruitment refers to unusually rapid growth of loudness as the level of a tone is increased, occurs in association with hearing loss, and is purely a cochlear phenomenon. It might coexist with hyperacusis, but there is no functional link between these two phenomena.

Prevalence and Epidemiology of Decreased Sound Tolerance

There are limited data available on the prevalence of hyperacusis. Questionnaires provide an assessment of hyperacusis prevalence in the general population. Recent data gathered from 10,349 randomly selected subjects showed that 15.3% reported hyperacusis.[17]

Patients being evaluated for other otologic problems frequently undergo audiologic evaluation, which involves assessment of speech discomfort level and pure-tone LDLs. There are no good normative data. Several studies indicated that in the normal population, LDLs are in the range of 90 to 110 dB SPL, with varied results depending on the specific method used (eg, stimuli: pure tone, warble tone, noise; presentation: free field, insert earphones, headphones; and instructions given to patients).[55–57]

TABLE 22–2. Medical Conditions That May Be Associated with Tinnitus

Conductive hearing losses
Otitis media
Cerumen impaction
Ossicular stiffness/discontinuity
Otosclerosis
Sensorineural hearing losses
Meniere's disease
Presbycusis
Cochlear otosclerosis
Vestibular schwannoma
Sudden hearing loss
Hormonal changes
Pregnancy
Menopause
Thyroid dysfunction
Some medications or withdrawal from them
Somatosounds
Produced by structures adjacent to the ear
Pulsatile
Neoplasm
Arterial
Venous
Beginning of intracranial hypertension
Great vessel bruits
Nonpulsatile
Tensor tympani myoclonus
Tensor veli palatini myoclonus
Patent eustachian tube
Produced by structures in the ear
Spontaneous otoacoustic emissions
Produced by joint abnormalities
Temporomandibular joint disorders

Adapted from Jastreboff PJ;[37] Perry BP and Gantz BJ.[45]

The results tend to cluster within 95 to 110 dB SPL for frequencies from 500 to 8,000 Hz, which correspond to approximately 90 to 100 dB HL.[55,58]

Hyperacusis and tinnitus frequently coexist, and it has been postulated that in some patients, hyperacusis might actually be a pretinnitus state.[12] Approximately 40% of tinnitus patients exhibit decreased sound tolerance, with 27% requiring specific treatment for hyperacusis.[46,59,60] Conversely, study of 100 patients with hypersensitivity to sound showed that 86% of them suffered from tinnitus.[61] Considering the clinical observation that approximately 27% of tinnitus patients required treatment for hyperacusis and 86% of hyperacusis patients reported tinnitus and accepting that about 4 to 5% of the general population have clinically significant tinnitus, it is possible to estimate that significant hyperacusis exists in approximately 1 to 1.5% of the general population. An even larger proportion has some hyperacusis that, although detectable in questionnaires, is not sufficiently strong to initiate intervention.

Evaluation of Tinnitus

We do not have any objective method to detect and measure tinnitus. Therefore, interview and psychoacoustic characterization are typical approaches in clinical practice, sometimes expanding into the direction of physiologic testing. New advances in research offer the possibility to detect tinnitus in an objective manner using imaging techniques[62–65] or magnetoencephalography.[13] These techniques are very promising but cannot yet be implemented into clinical practice owing to their complexity and cost. Table 22–3 lists the methods used for evaluation of tinnitus.

TABLE 22–3. Methods Used for Evaluation of Tinnitus

Interview/questionnaires
Psychoacoustic
Perceptual location
Pitch
Loudness
Maskability
Postmasking effects
Physiologic
Otoacoustic emissions
Efferent-mediated suppression of otoacoustic emissions
Spontaneous auditory nerve activity
Auditory brainstem responses
Late cortical potentials
Positron emission tomography/single photon emission tomography
Functional magnetic resonance imaging
Magnetoencephalography

Adapted from Jastreboff PJ et al;[46,69] Penner MJ;[66] Lind O;[67] McKee GJ and Stephens SD;[68] Jacobson et al;[70,71] Martin et al.[72]

Problems Evoked by Tinnitus

As tinnitus can present as part of a complex medical condition, full physical examination is needed to exclude all medically treatable problems that can be linked to tinnitus. Even though tinnitus is classified as a symptom and not a disease, it does require treatment as it can cause significant emotional and somatic distress and can significantly influence patients' quality of life, particularly if allowed to become a chronic problem. The list of reported associated complaints is long and includes emotional problems such as irritation, annoyance, anxiety, and depression; hearing problems such as difficulty with speech comprehension; and somatic problems such as headache, neck pain, and jaw pain.[1,6,73] Tinnitus can be very intrusive and may cause difficulty with sleep and concentration and a decreased ability to participate in everyday activities and social events; it may also create problems in relationships. A detailed interview, aimed at characterizing the specifics and degree of tinnitus impact on patients, coupled with otolaryngologic evaluation, provides the most thorough assessment and allows the practitioner to address all of the issues that need to be considered, including potential intervention of a psychologist/psychiatrist to accompany the commencement of a specific tinnitus-oriented treatment.

Hyperacusis and Misophonia as a Problem

Decreased sound tolerance can have an extremely strong effect on patients' lives and can be even more debilitating than tinnitus. Whereas tinnitus may affect

attention, sleep, work, and life enjoyment and make social contact less rewarding, hyperacusis can *prevent* people from exposing themselves to louder environments and therefore prevent them from working and interacting socially; it can also control a patient's life. In extreme cases, patients do not leave their homes, and their lives and the lives of their families are controlled by the issue of avoidance of sound. Misophonia can have similar effects, and since it is inevitable in all cases with significant hyperacusis, misophonia further enhances the effects of hyperacusis.

Decreased Sound Tolerance as a Symptom of Medical Conditions

Hyperacusis has been linked to a number of medical conditions (Table 22–4).[46,74–86]

TABLE 22–4. Medical Conditions Linked to Hyperacusis

Tinnitus
Bell's palsy
Lyme disease
Williams syndrome
Ramsay Hunt syndrome
Post-stapedectomy
Perilymphatic fistula
Head injury
Migraine
Depression
Withdrawal from benzodiazepines
Cerebrospinal fluid high pressure
Addison's disease

Adapted from Jastreboff PJ et al;[46] Adour KK and Wingerd J;[74] Fallon BA et al;[75] Nields JA et al;[76] Klein AJ et al;[77] Wayman et al;[78] McCandless GA and Georing DM;[79] Fukaya T and Nomura Y;[80] Waddell PA and Gromwall DMA;[81] Vingen JV et al;[82] Gopal KV et al;[83] Lader M;[84] Oen JM et al;[85] Henkin RI and Daly RL.[86]

Etiology and Potential Mechanisms of Decreased Sound Tolerance

In the majority of cases, the cause of hyperacusis is unknown. Hyperacusis has been linked to sound exposure (particularly short, impulse noise), head injury, stress, and medications. The lack of strong epidemiologic data and of animal models of hyperacusis prevents proving the validity of any potential mechanism responsible for hyperacusis.

At the peripheral level, it is possible to speculate that the abnormal enhancement of vibratory signals within the cochlea by the outer hair cells might result in overstimulation of the inner hair cells and subsequently results in hyperacusis.[12] Indeed, in some cases, it is possible to observe high-amplitude distortion-product otoacoustic emissions and distortion products evoked by low-level primaries.[87] The presence of asymmetric hyperacusis[46] might indicate a peripheral mechanism since central mechanisms would more likely act similarly on both sides.

Laboratory research has shown that damage to the cochlea or a decrease in the auditory input results in a decrease of the threshold of response in a significant proportion of neurons in the ventral cochlear nucleus and inferior colliculus.[88] Studies with evoked potentials indicated abnormal increase of the gain in the auditory pathways after such manipulations.[89] Some of the medical conditions listed in Table 22–4 can be linked to the central processing of signals and modification of the level of neuromodulators as possible factors inducing or enhancing hyperacusis. Moreover, serotonin has been implicated in hyperacusis,[90] and a case report indicated that serotonin reuptake inhibitors might be helpful for hyperacusis.[83]

Mechanisms of misophonia could involve enhancement of the functional links between the auditory and limbic systems, both at the cognitive and subconscious levels.[12] Alternatively, a tonic high level of activation of the limbic and autonomic nervous systems may result in strong behavioral reactions to moderate sounds.[39,91]

Methods of Evaluation of Decreased Sound Tolerance

Whereas there is no one clearly accepted method for the evaluation of decreased sound tolerance, hyperacusis, and misophonia, there is a general agreement that LDLs provide a good estimation of the problem. There are several variants of the protocols of LDL evaluation with various features, for example, continuous, pulsed, or beeps of sound; pure tone;

narrow-band noise; etc.[56,92,93] The approach we are pursuing incorporates modifications of the standard procedure[55] aimed at obtaining results dominated by hyperacusis by decreasing the effects of the misophonic component to a minimum. To achieve this, a situation is created during testing in which patients have the feeling of full control over the maximal sound level to which they will be exposed.[46] A detailed interview is needed with each patient to determine the relative contribution of hyperacusis and misophonia to decreased sound tolerance, reflected in decreased values of LDLs. As normative data are not uniform and there is substantial individual variability (even while using one method) in measuring LDL,[57] it is advisable to pay attention to the potential presence of hyperacusis when average LDL values are lower than 95 to 100 dB HL.

Review of Treatments

The list of approaches and techniques attempted to help tinnitus patients is very long. Table 22–5 lists the most commonly used of these treatments.[45]

TABLE 22–5. Review of Treatments for Tinnitus

Antireassurance
Pharmacology
Surgery
Electrical stimulation
Masking
Psychological approaches
Tinnitus retraining therapy
Other approaches
Biofeedback
Temporomandibular joint treatment
Acupuncture
Hyperbaric oxygen therapy
Homeopathy
Magnets

Adapted from Perry BP and Gantz BJ.[45]

Antireassurance Over the years, the most common advice to tinnitus patients has been the statement, "Nothing can be done—go home and learn to live with it."[7,48] This is actually a very powerful form of negative counseling, sufficient on many occasions to convert a person who just experiences tinnitus to a patient who suffers from it.[12,47,52]

Pharmacology Many pharmacologic agents have been considered for tinnitus treatment (Table 22–6),[6,94] but no single, effective, specific, secure, and reliable drug has yet to be identified. In this respect, strong consideration must be given to the side effects of pharmacologic treatments such as tolerance, dependence, and withdrawal effects. A recent review of randomized clinical trials of drugs for tinnitus has shown that all studied drugs have failed to prove their efficacy, as compared with a placebo.[94] Future double-blind randomized studies, with proper outcome measurements and adequate sample size, might identify some promising pharmacologic agents.

Surgery Surgery can offer help to some patients with somatosound, conductive hearing loss, and Meniere's disease.[6] However, there is no specific surgical procedure shown to be consistently effective for tinnitus that does not have a clear surgically treatable cause. Neither transection nor microvascular decompression of the auditory nerve, promoted in the past,[95,96] has proven to be effective.[45,97]

Electrical Suppression Electrical suppression of tinnitus was first reported in 1855.[98] Over the years, many different approaches have been attempted (Table 22–7),[98,99] and methodologies based on electrical stimulation are still raising significant interest. Recently, two new variants of electrical stimulation for tinnitus have been introduced. The first involves deep brain stimulation, sometimes performed for movement disorder and chronic pain.[100] Anecdotal reports of patients receiving deep brain stimulation initiated an investigation that demonstrated that, indeed, some patients experienced a decrease in their tinnitus with deep brain stimulation.[101] A second approach involves high-frequency electrical stimulation of the cochlea performed by placing an electrode on the promontory or via a cochlear implant (Rubinstein, personal communication, 2001). Both approaches are at a very early stage of investigation.

Only intracochlear (or with the electrode on the promontory) stimulation has shown consistent and positive results in approximately 50% of patients.[99,102,103] Other approaches were less effec-

TABLE 22–6. Drugs Frequently Prescribed for Treatment for Tinnitus

Local anesthetics (lidocaine, procaine, tocainide, flecainide)
Sedatives (diazepam, flurazepam, oxazepam, alprazolam)
Antidepressants (nortriptyline, trimipramine)
Anticonvulsants (carbamazepine, clonazepam, amino-oxyacetic acid, lamotrigine, baclofen)
Vasodilators (niacin)
Calcium channel blockers (nimodipine, nifedipine)
Others (misoprostol, zinc, betahistine, cinnarizine, caroverine, melatonin, furosemide, ginkgo biloba)

Adapted from Jastrelboff PJ et al;[6] Dobie RA.[94]

TABLE 22–7. Conditions Used in Electrical Stimulation for Suppression of Tinnitus

Sites of stimulation
Behind the ear lobe/around the ear
Mastoid
Near cheeks
External auditory canal
Promontory
Tympanic membrane
Round window
Intracochlear
Type of stimulus
Direct current/positive pulses
Alternating current
Amplitude-modulated high-frequency carrier
Electrodes
Acute/chronic

Adapted from Dauman R;[98] Hazell JWP et al.[99]

tive.[98] Although positive direct/pulsed current can provide tinnitus suppression,[104] it has no clinical application as it would damage the cochlea if used for a prolonged period of time.

Masking The use of an external sound to cover tinnitus and thus bring immediate relief to patients, known as masking, was first used in 1825 by Itard.[9] At the end of the 1970s, Vernon and Schleuning revisited this idea and introduced the first commercial masker.[105,106] Initial reports proclaimed high success of maskers,[106,107] but the approach did not withstand the test of time.[94] The problem was the criterion used in evaluating the effectiveness of masking; for example, if the masker was still in use after 6 months, it was counted as a success.[107,108] Obviously, patients may continue using maskers even while not getting relief from their tinnitus. Presently, this method of alleviation of tinnitus is rarely used.

Psychology Psychological management of chronic tinnitus can be helpful for some patients.[109] As tinnitus affects patients' well-being, the application of cognitive behavioral therapies may have a positive impact on the quality of life by improving their ability to cope with tinnitus. Cognitive therapies, behavioral modifications, coping strategies, and minimizing distress are examples of the psychological approach.

Neurophysiological Model of Tinnitus and Tinnitus Retraining Therapy

More than a decade has passed since the introduction of the neurophysiological model of tinnitus and, based on it, treatment, which is known as tinnitus retraining therapy (TRT).[7,110,111] Several observations led to the neurophysiological model of tinnitus and hyperacusis. It is known that tinnitus induces distress in only about 25% of those who perceive it.[2,4,8] There is no correlation among the psychoacoustic characterization of tinnitus, tinnitus-induced distress, and the treatment outcome.[112] The experiment by Heller and Bergman showed that the perception of tinnitus cannot be pathologic since essentially everyone (tinnitus emerged in 94% of people without prior tinnitus when isolated for several minutes in an anechoic chamber) experiences it when put in a sufficiently quiet environment.[5] These observations strongly argue that the auditory system plays a secondary role, and other systems in the brain are dominant in clinically relevant tinnitus (ie, tinnitus that creates discomfort and annoyance and requires intervention).

Analysis of the problems reported by tinnitus patients, who exhibit a strong emotional reaction to its presence, a high level of anxiety, and psychosomatic problems, indicated that the limbic and autonomic nervous systems are crucial in individuals with clinically relevant tinnitus. It was postulated that the sustained activation of the limbic and autonomic nervous systems is essential in creating distress and, therefore, clinically relevant tinnitus.[12]

Tinnitus-related neuronal activity is processed by the parts of the central nervous system involved in memory and attention. It is possible to distinguish several feedback loops, with two major categories: loops involving the conscious perception of tinnitus and those that act at a subconscious level (Figure 22–1), with the subconscious loop dominant in most patients.[7,39] It is further suggested that the activation of the limbic and autonomic nervous systems by tinnitus-related neuronal activity follows the principles of conditioned reflexes.[39,52]

The processing of tinnitus-related neuronal activity occurs in a dynamic balance scenario, with continuous modification of the weights of synaptic connections. Both learning and memory have a physiologic basis in the modification of the strength of synaptic connections.[113] A continuous presence of tinnitus, combined with attention given to it, results in plastic modifications of synaptic connections, yielding the modification of receptive fields corresponding to the tinnitus signal and its subsequent enhancement.[37,39] Recently, this postulate has been proven using magnetoencephalography.[13]

Whereas the initial signal provided by the auditory system is needed to start the cascade of events, its strength is irrelevant, as the extent of activation of the limbic and autonomic nervous systems depends on the strength of negative associations linked to tinnitus and the susceptibility of the feedback loops to modification.[114] It appears that tinnitus-related neuronal activity may result from compensatory processes that occur within the cochlea and the auditory pathways to minor dysfunction at the periphery.[12,37]

Notably, once plastic modifications of neuronal connections occur, the peripheral signal itself may become of little importance, as is similarly observed in chronic pain.[38] Indeed, there are clear similarities between tinnitus and chronic pain, including the phenomenon of prolonged exacerbation of tinnitus as a result of exposure to sound, which is observed in some patients.[52,59]

The neurophysiological model includes several systems of the brain involved in analysis of clinically relevant tinnitus. All levels of the auditory pathways, starting from the cochlea through the subcortical centers and ending at the auditory cortex, are essential in creating the perception of tinnitus.[12] When subjects are not bothered or annoyed by tinnitus, auditory pathways are the only pathways involved, and tinnitus-related neuronal activity is constrained within the auditory system. Therefore, although subjects are perceiving tinnitus, they are not disturbed by it.[2,39,114]

In approximately 20% of those with tinnitus, strong negative emotions are induced, which, in turn, evoke a variety of physiological defense mechanisms of the brain. The limbic and autonomic nervous systems play a crucial role, and improper

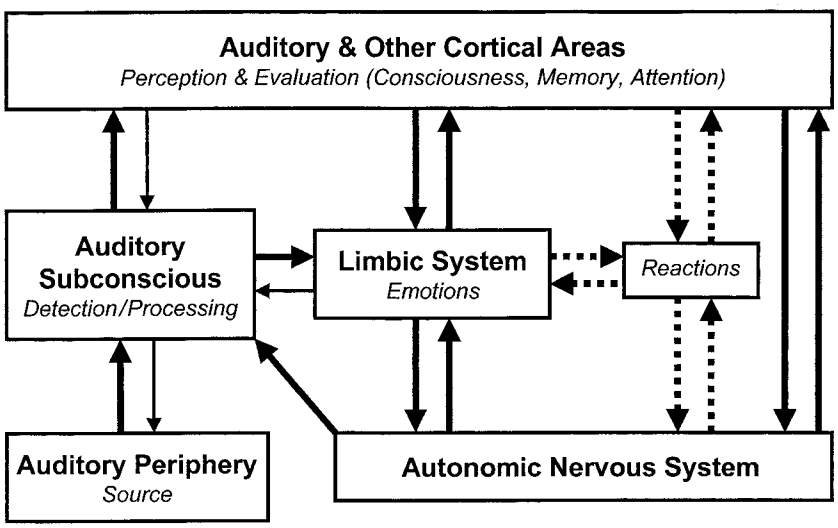


FIGURE 22–1. Neurophysiological model of tinnitus.

activation of these systems by tinnitus-related neuronal activity results at the behavioral level in the problems described by patients. The connections between the auditory, limbic, and autonomic systems with various cortical areas, as proposed in the neurophysiological model of tinnitus, are outlined in Figure 22–1.[6,7,12,37,39,60]

The model points out that the sustained activation of the limbic and autonomic nervous systems is responsible for the distress induced by clinically relevant tinnitus. Activation of both systems can be achieved through two routes. The first includes stimulation of the autonomic and limbic systems from higher-level cortical areas, which are involved in our awareness, verbalization, and beliefs. The second arises from the subconscious and provides stimulation from the lower-level auditory centers. The activation going through these two routes changes the strength of synaptic connections, enhancing the stimulation of the limbic and autonomic nervous systems by the tinnitus-related neuronal activity during the process of development of tinnitus as a clinical problem.

The question of how the neutral signal of tinnitus can evoke persistent strong distress can be answered by the principles of conditioned reflexes.[115] Basically, to create a conditional reflex, the temporal coincidence of sensory stimuli with negative (or positive) reinforcement is sufficient (Figure 22–2).[110,114–116] This initial association can be coincidental, without any real dependence. These types of associations of sensory stimuli are constantly created in normal life.

As long as the sensory stimulus is limited in time and there is no functional link between stimulus and reinforcement, this conditioned reflex will gradually disappear (habituate) owing to passive extinction of the reflex (the sensory stimulus is present but is not accompanied by a reinforcement) (see Figure 22–2). Since the 1930s, habituation has been defined as "The extinction of a conditioned reflex by repetition of the conditioned stimulus, … the method by which the nervous system reduces or inhibits responsiveness during repeated stimulation."[115] Habituation of perception of this stimulus will follow in the same manner as for all unimportant stimuli.

Notably, there are two different types of habituation. The first, *habituation of reaction*, is defined as "disappearance of a reaction to a neutral stimulus due to its repetitive appearance without reinforcement."[117] The second, *habituation of perception*, occurs when awareness of this particular stimulus disappears (Figure 22–3).[7] Habituations of reaction and perception are natural processes. Habituation is a crucial characteristic of brain function necessitated by the brain's inability to perform two tasks requiring complete attention simultaneously. When forced to carry out

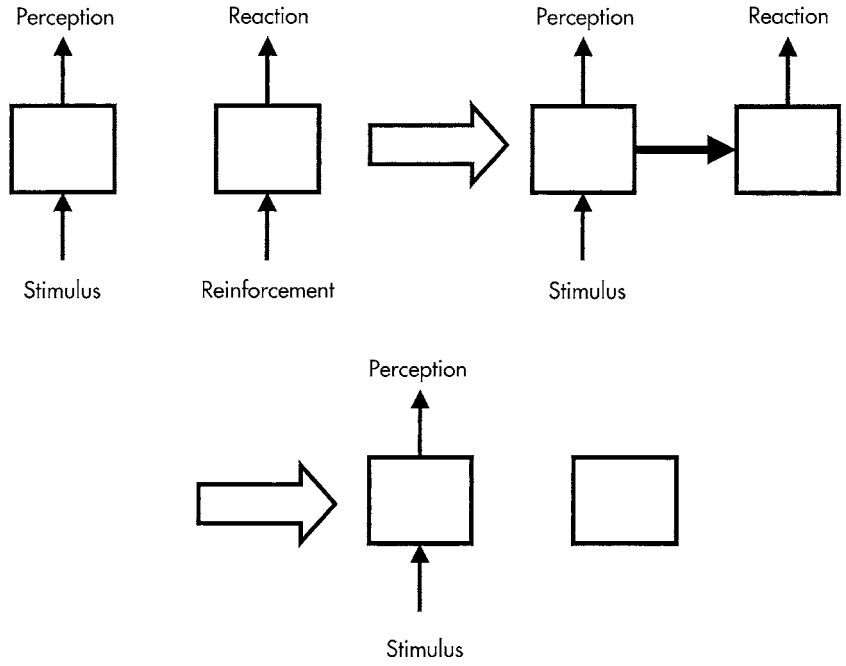


FIGURE 22–2. Principles of establishing conditioned reflexes and their passive extinction.

two tasks concurrently that require full consciousness, the brain uses its function of task switching by actually devoting consciousness to only one task at a time. Recently, brain areas involved in task switching have been indicated by a functional magnetic resonance imaging study.[118] If forced to monitor all of the incoming sensory stimuli, our brain would not be able to perform any tasks, except that of switching perception from one sensory stimulus to the other, ultimately paralyzing us in our actions.

To solve this problem, the central nervous system screens and categorizes all stimuli at the subconscious level. If the stimulus is new and unknown, it is passed to a higher cortical level, where it is perceived and evaluated. However, in the case of a stimulus to which we have previously been exposed, the stimulus is compared with patterns stored in memory. If the stimulus was classified as nonimportant and does not require action, it is blocked at the subconscious level of the auditory pathways and does not produce any reactions or reach the level of awareness. The reaction to this stimulus and its perception is habituated. In everyday life, habituation occurs to the majority of sensory stimuli surrounding us.

However, if a specific stimulus was once classified as important and, on the basis of comparison with the patterns stored in memory, was linked to something unpleasant or dangerous, this stimulus is perceived and attracts attention. Furthermore, the sympathetic part of the autonomic nervous system is activated, inducing reaction to this stimulus (frequently of the "fight or flight" category), which further reinforces memory patterns associated with this stimulus. Consequently, if the previous assessment of the importance of a stimulus has been confirmed, this specific stimulus becomes even more important, and its next appearance will result in faster identification, even in the presence of other competing stimuli, preventing the habituation of this stimulus. In the case of auditory stimuli, our auditory system becomes tuned into recognizing specific patterns of sound that have negative links. Under such conditions, the natural habituation of the tinnitus signal becomes impossible. In everyday life, this results in people having problems with their work, concentration, and sleep.

A simplistic description of the above process can be outlined to a patient as increased concern for tinnitus results in an increase of its significance, which ultimately increases the amount of time a person pays attention to it. This is a classic feedback loop or the "vicious circle" scenario, which increases patients' level of distress up to the limit of mental and physical endurance. At this stage, the patient will move from acute tinnitus, which can be easily relieved by proper counseling, into a chronic stage, which is much more difficult to deal with.

In the case of tinnitus, it is impossible to remove the reactions induced by the excitation of the sympathetic autonomic nervous system or even change them in a substantial manner. The solution to achieve the passive extinction of conditioned reflex, in which both stimulus (tinnitus) and nega-

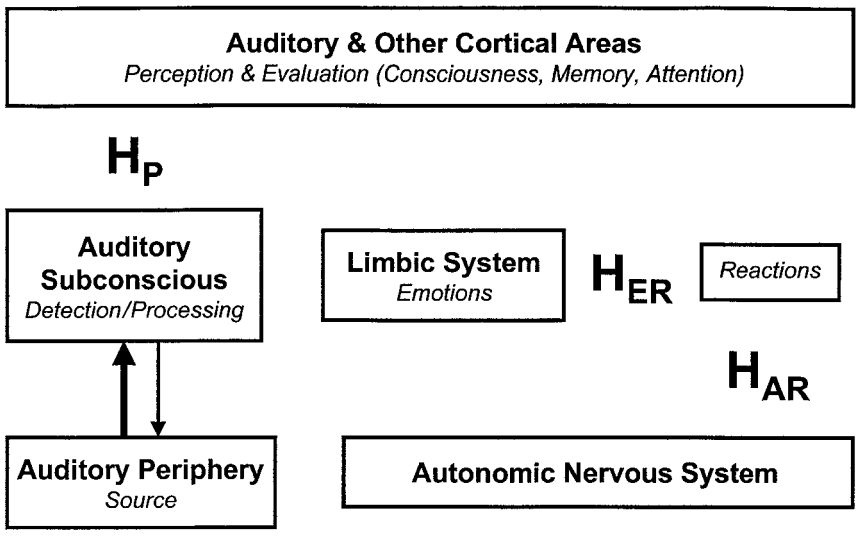


FIGURE 22–3. Habituation of autonomic (H_{AR}) and emotional reactions (H_{ER}) and habituation of perception (H_P).

tive reinforcement are continuously present, is to decrease the magnitude of this negative reinforcement over a period of time. This will result in partial weakening of the reflex but has to be applied consistently to yield positive effects. Moreover, it is fundamental that patients understand these principles so that the enhancement of this reflex by inducing too much verbal thinking and beliefs can be minimalized.

Once activation of the autonomic nervous system is lowered, this decreases negative reinforcement to a signal that is continuously present and gradually decreases the strength of the conditioned reflex. This causes further decreases in the reaction. Once tinnitus has achieved a neutral status, its habituation is inevitable as the brain is continuously habituating to all types of stimuli, assuming that they are not significant.

Consequently, retraining counseling (the first component of TRT) is oriented toward removal of the patient's negative associations with tinnitus and reclassification of tinnitus into a category of neutral stimuli. This is accomplished by educating the patients that their tinnitus results from a normal compensatory mechanism, which occurs in the auditory system, to typically minor changes in the cochlea. As part of counseling, it is also important to demystify the mechanisms through which tinnitus may affect a patient's life. Counseling in TRT is a teaching session aimed at providing patients with a new frame of reference by explaining potential mechanisms of tinnitus generation, neurophysiological mechanisms through which tinnitus is influencing various aspects of their lives, and that by activating a naturally occurring mechanism of brain function (habituation and the plasticity underlying it), it is possible to achieve primarily habituation of the tinnitus-induced reaction of the brain and the body and secondarily habituation of the tinnitus perception. The clear goal of achieving an active and selective block of tinnitus-induced reactions is set for the patients.

In addition to decreasing the strength of the activation of the limbic and autonomic nervous systems, initiated during the counseling session, the second component of TRT is sound therapy.[91,110] All of our senses are working on the principle of differences of the stimuli from the background, and the perceived strength of a signal is not linked directly to the physical strength of a stimulus. At this moment, we cannot easily suppress tinnitus-related neuronal activity, but by increasing background neuronal activity, we are effectively decreasing the strength of this signal, which activates the limbic and autonomic nervous systems and which is being processed in all of the centers involved. There is no simple proportional relationship between the differences in tinnitus and background neuronal activity and the reactions induced by it. Nonetheless, we can achieve a decrease of reactions induced by the tinnitus and, through this, facilitate extinction of the conditioned reflex.

It is important to analyze the theoretical relationships existing between the physical intensity of sound and its effectiveness on tinnitus habituation (Figure 22–4). Five principles influence this relationship[91,110,111,119]: (1) stochastic resonance (enhancement of the signal by adding low-level noise); (2) dependence of the signal's strength on its contrast with the background; (3) total suppression of the signal, preventing any retraining and consequently habituation; (4) partial suppression ("partial masking"), which does not prevent retraining but does make it more difficult as the training is performed on a different stimulus than the original; and (5) activation of limbic and autonomic nervous systems by too loud or unpleasant sounds, yielding an increase in tinnitus and contracting habituation.

It is intuitive that if the sound level is significantly below the threshold level of detection, there is no basis to believe that it asserts any effect on the auditory system and tinnitus. On the other hand, when external sound becomes strong enough to suppress detection of tinnitus-related neuronal activity, by definition, any kind of retraining (including habituation) is prevented because the brain cannot change reactions to a stimulus it cannot detect.

When the sound level is still below but close to the threshold of detection, the phenomenon of stochastic resonance (eg, addition of a low level of noise can decrease the threshold of detection of the stimulus and enhance it when the stimulus is weak and close to the threshold of detection) comes into play.[117] Recently, the presence of stochastic resonance has been shown at the level of hair cells and the auditory nerve, and preliminary data indicate its effect on the loudness of tinnitus.[119–121] The effective sound level inducing stochastic resonance covers a range of about 15 to 20 dB, beginning from about –5 dB below the threshold of detection of the origi-

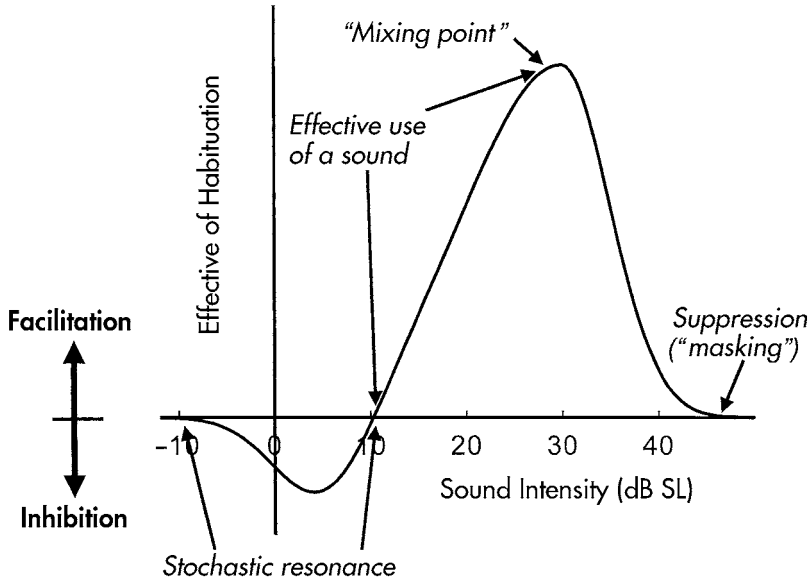


FIGURE 22–4. Functional dependence of habituation effectiveness on physical intensity of a sound. Notice the need to avoid sound levels close to threshold of hearing and those inducing partial or total suppression.

nal signal. Thus, owing to stochastic resonance, by adding low-level external noise (eg, by using sound generators set at the threshold of hearing), paradoxical enhancement of the tinnitus signal occurs. This will, in turn, make habituation more difficult. The results of a study in which a comparison was performed among groups with counseling only (including advice on using environmental sounds), counseling combined with sound generators set at the threshold of hearing, and counseling combined with sound generators set close to the "mixing" point fully support the prediction of the importance of stochastic resonance. The group that performed the worst was the one with the sound level set close to the threshold of hearing, whereas the group that performed the best was the one with the sound level set at the "mixing point," with the counseling-only group in the middle.[122]

When sound level is increased further, the mechanism involving the decreased difference between the tinnitus signals and background neuronal activity becomes the dominant factor. As with all perception, the difference between signal and background plays a dominant role. By decreasing the difference between the tinnitus-related neuronal activity and the background ongoing neuronal activity, the effective strength of the tinnitus signal decreases, and this weaker signal is passed to the higher-level cortical areas and, most importantly, to the limbic and autonomic nervous systems. This helps in initiating and sustaining the process of passive extinction of conditioned reflexes that link tinnitus to negative reactions.[39,110] As the background activity is the sum of spontaneous and evoked activities, a decrease in the difference between the tinnitus signal and the background neuronal activity can be achieved by exposing patients to additional external sound.

This principle, if working alone, would imply that we should use a sound that is as intense as possible. Two other factors (3 and 4), however, become dominant. First, once the tinnitus signal is suppressed, habituation will not occur by definition owing to the lack of a signal to habituate (principle 3). Second, when the sound level surpasses the threshold of partial tinnitus suppression ("partial masking"), it will modify not only the intensity but also the quality of the tinnitus signal. Then retraining of neuronal networks will occur to the modified tinnitus signal and not to the original one. Owing to the generalization principle (eg, reaction can be induced to stimuli similar to the original, with the strength depending on the difference between the original and the modified signals), some habituation may still occur. The higher the external sound is above the threshold of partial masking, the smaller its contribution to habituation. Finally, once a level of total suppression is reached, the effectiveness of habituation is decreased to zero as the brain is unable to retrain to an undetectable signal.

The optimal setting of the sound level is different when hyperacusis is the dominant or only problem. In this situation, the effect of stochastic resonance is of secondary importance to the primary need not to overstimulate the auditory system. Patients start with a sound level closer to their threshold but as high as their sound sensitivity allows, with the aim to be above the range of stochastic resonance. Once a partial reversal of hyperacusis is achieved, the sound level can be increased rapidly to address tinnitus directly. At this point, the rules previously outlined for patients with tinnitus should be followed.

The need to preserve stimulation in the low-frequency range yields a strong recommendation for people who have relatively normal low-frequency hearing to be provided with devices or hearing aids with fittings as open as possible. It is not sufficiently appreciated that in the normal acoustic environment there is a high proportion of low-frequency sounds, below 200 Hz, which provide constant stimulation of the auditory pathways. Since the majority of patients have relatively normal hearing in this frequency range, they benefit from this stimulation. Consequently, blocking the ear canal with closed ear molds decreases the auditory input, and many patients experience enhancement of tinnitus when their ears are blocked.

Note that even the best hearing aids act as earplugs in low frequencies when they are the in-the-canal type or are fitted with a closed mold as they are unable to reproduce frequencies below 200 to 250 Hz owing to restriction based on the physics of sound generation by a loudspeaker. Note that hearing aids for patients with tinnitus are used primarily as a part of sound therapy to provide extra amplification of background sounds and only secondarily for communicative purposes. Other tools may also be used to enrich the auditory background such as nature sounds, neutral music, or tabletop sound generators.

Both counseling and sound use are dependent on patient categorization, and issues related to sound are summarized in Table 22–8. This categorization provides general guidance for treatment with TRT.[59,111] During the process of treatment, patients may move from one category to another (eg, hyperacusis can be totally eliminated and consequently the patient may move from category 3 to 1), and recommendations regarding sound use should be modified. This is one of the reasons why follow-up contacts are important; first, to continue counseling, second, to check patient status; and third, to modify protocol if necessary.

Typically, the first effects of TRT are seen in about 3 months, with clear improvement in about 6 months, and many patients achieve a high level of control of their tinnitus by about 12 months.[47,52,59,125] Patients are advised to follow the TRT protocol for approximately 18 months to prevent a relapse. Improvement in hyperacusis is typically faster, and the success rate is higher than that of tinnitus without hyperacusis. The results from other centers and ours using TRT show satisfactory results in over 80% of patients.[122,124–127]

Treatments for Hyperacusis

Treatments for hyperacusis go into two contrary directions. First, the most common approach is to advise patients to avoid sound and use ear protection. This is based on reasoning that because patients became sensitive to sound, this may indicate that they are more susceptible to sound exposure and consequently need extra protection. Patients easily embrace this philosophy and start to protect their ears, even to the extent of using earplugs in quiet environments. Unfortunately, this approach actually makes the auditory system even more sensitive and further exacerbates hyperacusis.

The second approach involves the desensitization of patients by exposure to a variety of sounds. The desensitization approach has been promoted for some time with a variety of protocols and types of sounds used, such as the recommendation of using sound with certain frequencies removed, short exposures to moderately loud sound, or prolonged exposures to low-level sounds.[50,52,128] According to principles of the neurophysiological model of tinnitus, the latter approach is recommended and is used as part of TRT. Note, that the misophonic component cannot be removed by desensitization, and a separate approach needs to be implemented.

Tinnitus Retraining Therapy for Decreased Sound Tolerance Tinnitus retraining therapy can help patients with both tinnitus and hyperacusis, and the presence of hyperacusis is one of the crucial factors in categorization of the patients (see Table 22–8) and determining the optimal protocol for

TABLE 22–8. Categories of Patients with Tinnitus and Hyperacusis

Category	*Impact on Life*	*Tinnitus*	*Subjective Hearing Loss*	*Hyperacusis*	*Prolonged Sound-Induced Exacerbation*	*Treatment (always involves counseling and use of enriched background sound)*
0	Low	Present	—	—	—	Abbreviated counseling
1	High	Present	—	—	—	Sound generators set at mixing point
2	High	Present	Present	—	—	Hearing aid with stress on enrichment of the auditory background
3	High	Not relevant	Not relevant	Present	—	Sound generators set above threshold of hearing
4	High	Not relevant	Not relevant	Present	Present	Sound generators set at the threshold; very slow increase of sound level

Hyperacusis is significant sensitivity to environmental sounds typically associated with loudness discomfort levels below 100 dB HL; Prolonged Sound-Induced Exacerbation of tinnitus/hyperacusis is when the effects persist to the following day; Subjective Hearing Loss is perceived subjectively by patients as having a significant impact on their life; Impact on Life is the extent of impact of tinnitus and/or hyperacusis on the patient's life. Common treatment for each category involves counseling and the use of an enriched auditory background.

their treatment. It is recommended that if hyperacusis is present, it has to be treated first. Although TRT offers only a treatment for tinnitus (not a cure), it can, in some patients, totally remove hyperacusis and misophonia, thus providing a cure for these conditions.[59,60,111]

In some patients, tinnitus and hyperacusis are two manifestations of the same internal mechanisms of increased gain within the auditory pathways, and the improvement in hyperacusis results in the improvement in tinnitus as well. Moreover, the removal of hyperacusis yields a decrease in general anxiety and stress, which, in combination with proper counseling, significantly facilitates tinnitus habituation.

A few parameters of the TRT protocol are of specific importance in treating patients with hyperacusis: avoidance of silence and continued exposure to sound are even more important than for patients with only tinnitus. The level of sound should be better controlled during the treatment, which necessitates the use of wearable sound generators. The sound used should never induce discomfort or annoyance.

The setting of sound generators for hyperacusis is more complex than for tinnitus treatment. There might be a need for a low setting of initial sound level, which might bring patients to the range of effects of stochastic resonance. The use of real-ear measurements, as a guide in setting and checking sound level for all patients with instrumentation during initial and follow-up visits, is crucial in the case of hyperacusis patients.

The method of desensitization works on the auditory system and consequently will not affect misophonia, which needs to be addressed using active extinction of conditioned reflexes between the auditory and limbic systems. This is achieved in practice by instructing patients to engage systematically in activities involving sound as a fundamental component and that are pleasant (such as actively listening to music).

CONCLUSION

Tinnitus and hyperacusis are still challenging topics to study and symptoms to treat. Many questions remain unanswered. Mechanisms of tinnitus and hyperacusis are speculative and not yet proven. We do not yet have objective methods for detection and evaluation of tinnitus. We believe that the neurophysiological model of tinnitus and TRT provide a

promising approach that may ultimately result in a better understanding of tinnitus and in providing greater help to patients with tinnitus and hyperacusis.

REFERENCES

1. McFadden D. Tinnitus: facts, theories, and treatments. Washington (DC): National Academy Press; 1982.
2. Coles RRA. Epidemiology, aetiology and classification. In: Reich GE, Vernon JA, editors. Proceedings of the Fifth International Tinnitus Seminar, 1995, Portland, OR, U.S.A. Portland (OR): American Tinnitus Association; 1996. p. 25–30.
3. Pilgramm M, Rychlick R, Lebisch H, et al. Tinnitus in the Federal Republic of Germany: a representative epidemiological study. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 64–7.
4. Davis A, El Refaie A. Epidemiology of tinnitus. In: Tyler R, editor. Tinnitus handbook. San Diego: Singular, Thomson Learning; 2000. p. 1–23.
5. Heller MF, Bergman M. Tinnitus in normally hearing persons. Ann Otol Rhinol Laryngol 1953; 62:73–93.
6. Jastreboff PJ, Gray WC, Mattox DE. Tinnitus and hyperacusis. In: Cummings CW, Fredrickson JM, Harker LA, et al, editors. Otolaryngology head and neck surgery. St. Louis: Mosby; 1998. p. 3198–222.
7. Jastreboff PJ. Tinnitus habituation therapy (THT) and tinnitus retraining therapy (TRT). In: Tyler R, editor. Tinnitus handbook. San Diego: Singular, Thomson Learning; 2000. p. 357–76.
8. Davis A. The aetiology of tinnitus: risk factors for tinnitus in the UK population—a possible role for conductive pathologies? In: Reich GE, Vernon JA, editors. Proceedings of the Fifth International Tinnitus Seminar, 1995, Portland, OR, U.S.A. Portland (OR): American Tinnitus Association; 1996. p. 38–45.
9. Stephens D. A history of tinnitus. In: Tyler RS, editor. Tinnitus handbook. San Diego: Singular, Thomson Learning; 2000. p. 437–48.
10. Feldmann H. History of tinnitus research. In: Shulman A, editor. Tinnitus diagnosis and treatment. Philadelphia: Lea & Febiger; 1991. p. 3–37.
11. Feldmann H. Pathophysiology of tinnitus. In: Kitahara M, editor. Tinnitus: pathophysiology and management. Tokyo: Igaku-Shoin; 1988. p. 7–35.
12. Jastreboff PJ. Phantom auditory perception (tinnitus): mechanisms of generation and perception. Neurosci Res 1990;8:221–54.
13. Muhlnickel W, Elbert T, Taub E, Flor H. Reorganization of auditory cortex in tinnitus. Proc Natl Acad Sci U S A 1998;95:10340–3.
14. Axelsson A, Ringdahl A. Tinnitus—a study of its prevalence and characteristics. Br J Audiol 1989; 23:53–62.
15. Leske MC. Prevalence estimates of communicative disorders in the U.S.A.: language, learning and vestibular disorders. ASHA 1981;23:229–37.
16. George RN, Kemp S. A survey of New Zealanders with tinnitus. Br J Audiol 1991;25:331–6.
17. Fabijanska A, Rogowski M, Bartnik G, Skarzynski H. Epidemiology of tinnitus and hyperacusis in Poland. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 569–71.
18. Siwiec H. Epidemiology of tinnitus in Lublin district. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 572–3.
19. Spoendlin H. Inner ear pathology and tinnitus. In: Feldmann H, editor. Proceedings III International Tinnitus Seminar, Muenster 1987. Karlsruhe, Germany: Harsch Verlag; 1987. p. 42–51.
20. Nodar RH. Tinnitus aurium in school age children: a survey. J Audiol Res 1972;12:133–5.
21. Mills RP, Albert DM, Brain CE. Tinnitus in childhood. Clin Otolaryngol 1986;11:431–4.
22. Mills RP, Cherry JR. Subjective tinnitus in children with otological disorders. Int J Pediatr Otorhinolaryngol 1984;7:21–7.
23. Gabriels P. Children with tinnitus. In: Reich GE, Vernon JA, editors. Proceedings of the Fifth International Tinnitus Seminar, 1995, Portland, OR, U.S.A. Portland (OR): American Tinnitus Association; 1996. p. 270–5.
24. Stouffer JL, Tyler RS. Characterization of tinnitus by tinnitus patients. J Speech Hear Disord 1990;55: 439–53.
25. Gurr P, Owen G, Reid A, Canter R. Tinnitus in pregnancy. Clin Otolaryngol 1993;18:294–7.
26. Kemp S, George RN. Diaries of tinnitus sufferers. Br J Audiol 1992;26:381–6.
27. Zelman S. Correlation of smoking history with hearing loss. JAMA 1973;223:920.

28. Ronis ML. Alcohol and dietary influences on tinnitus. J Laryngol Otol 1984;98 Suppl 9:242–6.
29. Tonndorf J. The analogy between tinnitus and pain: a suggestion for a physiological basis of chronic tinnitus. Hear Res 1987;28:271–5.
30. Tonndorf J. Stereociliary dysfunction, a case of sensory hearing loss, recruitment, poor speech discrimination and tinnitus. Acta Otolaryngol (Stockh) 1981;91:469–79.
31. Kiang NYS, Moxon EC, Levine RA. Auditory-nerve activity in cats with normal and abnormal cochleas. In: Wolstenholme GEW, Knight J, editors. Ciba Foundation Symposium on Sensorineural Hearing Loss. London: Churchill; 1970. p. 241–73.
32. Salvi RJ, Ahroon WA. Tinnitus and neural activity. J Speech Hear Res 1983;26:629–32.
33. Penner MJ. Two-tone forward masking patterns and tinnitus. J Speech Hear Res 1980;23:779–86.
34. Møller AR. Pathophysiology of tinnitus. Ann Otol Rhinol Laryngol 1984;93:39–44.
35. Eggermont JJ. On the pathophysiology of tinnitus: a review and a peripheral model. Hear Res 1990; 48:111–24.
36. Langner G, Wallhdusser-Franke E. Computer simulation of a tinnitus model based on labelling of tinnitus activity in the auditory cortex. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 20–5.
37. Jastreboff PJ. Tinnitus as a phantom perception: theories and clinical implications. In: Vernon J, Møller AR, editors. Mechanisms of tinnitus. Boston: Allyn & Bacon; 1995. p. 73–94.
38. Møller AR. Pathophysiology of severe tinnitus and chronic pain. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 26–31.
39. Jastreboff PJ. The neurophysiological model of tinnitus and hyperacusis. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 32–8.
40. Pujol R. Neuropharmacology of the cochlea and tinnitus. In: Aran J-M, Dauman R, editors. Tinnitus 91. Proceedings IV International Tinnitus Seminar, Bordeaux, France, 1991. Amsterdam: Kugler Publications; 1992. p. 103–7.
41. Tonndorf J. Stereociliary dysfunction, a cause of sensory hearing loss, recruitment, poor speech discrimination and tinnitus. Acta Otolaryngol (Stockh) 1981;91:469–79.
42. Zenner H-P, Ernst A. Cochlear-motor, transduction and signal-transfer tinnitus. Eur Arch Otolaryngol 1993;249:447–54.
43. LePage EL. A model for cochlear origin of subjective tinnitus: excitatory drift in the operating point of inner hair cells. In: Vernon J, Møller AR, editors. Mechanisms of tinnitus. Boston: Allyn & Bacon; 1995. p. 115–47.
44. Jastreboff PJ. Processing of the tinnitus signal within the brain. In: Reich GE, Vernon JA, editors. Proceedings of the Fifth International Tinnitus Seminar, 1995, Portland, OR, U.S.A. Portland (OR): American Tinnitus Association; 1996. p. 58–67.
45. Perry BP, Gantz BJ. Medical and surgical evaluation and management of tinnitus. In: Tyler R, editor. Tinnitus handbook. San Diego: Singular, Thomson Learning; 2000. p. 221–41.
46. Jastreboff PJ, Jastreboff MM, Sheldrake JB. Audiometrical characterization of hyperacusis patients before and during TRT. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 495–8.
47. Jastreboff PJ, Gray WC, Gold SL. Neurophysiological approach to tinnitus patients. Am J Otol 1996;17: 236–40.
48. Jastreboff PJ, Hazell JWP. A neurophysiological approach to tinnitus: clinical implications. Br J Audiol 1993;27:1–11.
49. Wolk C, Seefeld B. The effects of managing hyperacusis with maskers (noise generators). In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 512–4.
50. Hazell JWP, Sheldrake JB. Hyperacusis and tinnitus. In: Aran J-M, Dauman R, editors. Tinnitus 91. Proceedings IV International Tinnitus Seminar, Bordeaux, France, 1991. Amsterdam: Kugler Publications; 1992. p. 245–8.
51. Bartnik G, Fubijunska A, Ragowski M. Effects of tinnitus retraining therapy (TRT) for patients with tinnitus and subjective hearing loss versus tinnitus only. Scand Audiol 2001;Suppl:206–8.
52. Jastreboff PJ, Jastreboff MM. Tinnitus retraining therapy (TRT) as a method for treatment of tinnitus and hyperacusis patients. J Am Acad Audiol 2000; 11:156–61.
53. Stedman's medical dictionary. 26th ed. Baltimore: Williams & Wilkins; 1995. Hyperacusis.

54. American heritage dictionary. 3rd ed. Boston: SoftKey International; 1994. Hyperacusis.
55. Hood JD, Poole JP. Tolerable limit to loudness: its clinical and physiological significance. J Acoust Soc Am 1966;40:47–53.
56. Cox RM, Alexander GC, Taylor IM, Gray GA. The contour test of loudness perception. Ear Hear 1997;18:388–400.
57. Byrne D, Dirks D. Effects of acclimatization and deprivation on non-speech auditory abilities. Ear Hear 1996;17 Suppl:29S–37S.
58. Stephens SD, Anderson C. Experimental studies on the uncomfortable loudness level. J Speech Hear Res 1971;14:262–70.
59. Jastreboff PJ. Categories of the patients and the treatment outcome. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 394–8.
60. Jastreboff PJ. Tinnitus. In: Gates GA, editor. Current therapy in otolaryngology head and neck surgery. St. Louis: Mosby, 1998. p. 90–5.
61. Anari M, Axelsson A, Elies W, Magnusson L. Hypersensitivity to sound—questionnaire data, audiometry and classification. Scand Audiol 1999;28:219–30.
62. Lockwood AH, Salvi RJ, Coad ML, et al. The functional neuroanatomy of tinnitus: evidence for limbic system links and neural plasticity. Neurology 1998; 50:114–20.
63. Melcher JR, Sigalovski I, Levine RA. Tinnitus-related fMRI activation patterns in human auditory nuclei. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 166–70.
64. Lockwood AH, Wack DS, Burkard RF, et al. The functional anatomy of gaze-evoked tinnitus and sustained lateral gaze. Neurology 2001;56:472–80.
65. Andersson G, Lyttkens L, Hirvela C, et al. Regional cerebral blood flow during tinnitus: a PET case study with lidocaine and auditory stimulation. Acta Otolaryngol (Stockh) 2000;120:967–72.
66. Penner MJ. Linking spontaneous otoacoustic emissions and tinnitus. Br J Audiol 1992;26:115–23.
67. Lind O. Transient-evoked otoacoustic emissions and contralateral suppression in patients with unilateral tinnitus. Scand Audiol 1996;25:167–72.
68. McKee GJ, Stephens SD. An investigation of normally hearing subjects with tinnitus. Audiology 1992;31:313–7.
69. Jastreboff PJ, Ikner CL, Hassen A. An approach to objective evaluation of tinnitus in humans. In Aran J-M, Dauman R, editors. Tinnitus 91. Proceedings IV International Tinnitus Seminar, Bordeaux, France, 1991. Amsterdam: Kugler Publications; 1992. p. 331–40.
70. Jacobson GP, Ahmad BK, Morgan J, et al. Auditory evoked cortical magnetic field (M100-M200) measurements in tinnitus and normal groups. Hear Res 1991;56:44–52.
71. Jacobson GP, Calder JA, Newman CW, et al. Electrophysiological indices of selective auditory attention in subjects with and without tinnitus. Hear Res 1996;97:66–74.
72. Martin WH, Schwegler JW, Shi J, et al. Developing an objective measurement tool for evaluating tinnitus: spectral averaging. In: Vernon JA, Reich G, editors. Proceedings of the Fifth International Tinnitus Seminar, 1995, Portland, OR, U.S.A. Portland (OR): American Tinnitus Association; 1996. p. 127–34.
73. Hallberg LR-M, Erlandsson SI. Tinnitus characteristics in tinnitus complainers and noncomplainers. Br J Audiol 1993;27:19–27.
74. Adour KK, Wingerd J. Idiopathic facial paralysis (Bell's palsy): factors affecting severity and outcome in 446 patients. Neurology 1974;24:1112–6.
75. Fallon BA, Nields JA, Burrascano JJ, et al. The neuropsychiatric manifestation of Lyme borreliosis. Psychiatr Q 1992;63:95–117.
76. Nields JA, Fallon BA, Jastreboff PJ. Carbamazepine in the treatment of Lyme disease-induced hyperacusis. J Neuropsychiatr Clin Neurosci 1999;11:97–9.
77. Klein AJ, Armstrong BL, Greer MK, Brown FR. Hyperacusis and otitis media in individuals with Williams syndrome. J Speech Hear Disord 1990; 55:339–44.
78. Wayman DM, Pham HN, Byl FM, Adour KK. Audiological manifestations of Ramsay Hunt syndrome. J Laryngol Otol 1990;104:104–8.
79. McCandless GA, Goering DM. Changes in loudness after stapedectomy. Arch Otolaryngol 1974;100: 344–50.
80. Fukaya T, Nomura Y. Audiological aspects of idiopathic perilymphatic fistula. Acta Otolaryngol Suppl (Stockh) 1988;456:68–73.
81. Waddell PA, Gronwall DMA. Sensitivity to light and sound following minor head injury. Acta Neurol Scand 1984;69:270–6.
82. Vingen JV, Pareja JA, Storen O, et al. Phonophobia in migraine. Cephalalgia 1998;18:243–9.

83. Gopal KV, Daly DM, Daniloff RG, Pennartz L. Effects of selective serotonin reuptake inhibitors on auditory processing: case study. J Am Acad Audiol 2000;11:454–63.
84. Lader M. Anxiolytic drugs: dependence, addiction and abuse. Eur Neuropsychopharmacol 1994;4:85–91.
85. Oen JM, Begeer JH, Staal-Schreinemachers A, Tijmstra T. Hyperacusis in children with spina bifida; a pilot-study. Eur J Pediatr Surg 1997;7 Suppl 1:46.
86. Henkin RI, Daly RL. Auditory detection and perception in normal man and in patients with adrenal cortical insufficiency: effect of adrenal cortical steroids. J Clin Invest 1968;47:1269–80.
87. Jastreboff PJ, Mattox DE. Treatment of hyperacusis by aspirin. In: Abstracts of the 21st Midwinter Research Meeting, Association for Research in Otolaryngology, St. Petersburg Beach, FL, 1998. www.aro.org.
88. Boettcher FA, Salvi RJ. Functional changes in the ventral cochlear nucleus following acute acoustic overstimulation. J Acoust Soc Am 1993;94:2123–34.
89. Gerken GM. Alteration of central auditory processing of brief stimuli: a review and a neural model. J Acoust Soc Am 1993;93:2038–49.
90. Marriage J, Barnes NM. Is central hyperacusis a symptom of 5-hydroxytryptamine (5-HT) dysfunction? J Laryngol Otol 1995;109:915–21.
91. Jastreboff PJ. Optimal sound use in TRT—theory and practice. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 491–4.
92. Ricketts TA, Bentler RA. The effect of test signal type and bandwidth on the categorical scaling of loudness. J Acoust Soc Am 1996;99:2281–7.
93. Hawkins DB, Walden BE, Montgomery A, Prosek RA. Description and validation of an LDL procedure designed to select SSPL90. Ear Hear 1987;8:162–9.
94. Dobie RA. A review of randomized clinical trials in tinnitus. Laryngoscope 1999;109:1202–11.
95. Pulec JL. Tinnitus: surgical therapy. Am J Otol 1984;5:479–80.
96. Møller MB, Møller AR, Jannetta PJ, Jho HD. Vascular decompression surgery for severe tinnitus: selection criteria and results. Laryngoscope 1993;103: 421–7.
97. Berliner KI, Shelton C, Hitselberger WE, Luxford WM. Acoustic tumors: effect of surgical removal on tinnitus. Am J Otol 1992;13:13–7.
98. Dauman R. Electrical stimulation for tinnitus supression. In: Tyler R, editor. Tinnitus handbook. San Diego: Singular, Thomson Learning; 2000. p. 377–98.
99. Hazell JWP, Jastreboff PJ, Meerton LE, Conway MJ. Electrical tinnitus suppression: frequency dependence of effects. Audiology 1993;32:68–77.
100. Olanow CW, Brin MF, Obeso JA. The role of deep brain stimulation as a surgical treatment for Parkinson's disease. Neurology 2000;55(12 Suppl 6):60–6.
101. Martin WH, Shi Y-B, Burchiel KJ, Anderson VC. Deep brain stimulation effects on hearing function and tinnitus. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 68–72.
102. Tyler RS. Tinnitus in the profoundly hearing-impaired and the effects of cochlear implants. Ann Otol Rhinol Laryngol Suppl 1995;165:25–30.
103. Tyler RS, Kelsay D. Advantages and disadvantages reported by some of the better cochlear-implant patients. Am J Otol 1990;11:282–9.
104. Portmann M, Cazals Y, Negrevergne M, Aran J-M. Temporary tinnitus suppression in man through electrical stimulation of the cochlea. Acta Otolaryngol (Stockh) 1979;87:294–9.
105. Vernon J. Attemps to relieve tinnitus. J Am Audiol Soc 1977;2:124–31.
106. Vernon J, Schleuning A. Tinnitus: a new management. Laryngoscope 1978;88:413–9.
107. Schleuning AJ, Johnson RM, Vernon JA. Evaluation of a tinnitus masking program: a follow-up study of 598 patients. Ear Hear 1980;1:71–4.
108. Johnson RM. The masking of tinnitus. In: Vernon JA, editor. Tinnitus treatment and relief. Boston: Allyn and Bacon; 1998. p. 164–86.
109. Wilson PH, Henry JL. Psychological management of tinnitus. In: Tyler R, editor. Tinnitus handbook. San Diego: Singular, Thomson Learning; 2000. p. 263–79.
110. Jastreboff PJ. How TRT derives from the neurophysiological model. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 87–91.
111. Jastreboff PJ, Jastreboff MM. Tinnitus retraining therapy. In: Baguley D, editor. Perspectives in tinnitus management. New York: Thieme; 2001. p. 51–63.
112. Jastreboff PJ, Hazell JWP, Graham RL. Neurophysiological model of tinnitus: dependence of the

minimal masking level on treatment outcome. Hear Res 1994;80:216–32.
113. Albus JS. A theory of cerebellar function. Math Biosci 1971;10:25–61.
114. Jastreboff PJ, Jastreboff MM. The neurophysiological model of tinnitus and its practical implementation: current status. In: Myers EN, Bluestone CD, Brackman DE, Krause CJ, editors. Advances in otolaryngology-head and neck surgery. Vol 15. St. Louis: Mosby; 2001. p. 135–47.
115. Konorski J. Conditioned reflexes and neuronal organization. Cambridge (UK): Cambridge University Press; 1948.
116. Konorski J. Integrative activity of the brain. Chicago: University of Chicago Press; 1967.
117. Thompson RF, Donegan NH. Learning and memory. In: Adelman G, editor. Encyclopedia of neuroscience. Boston: Birkhauser; 1987. p. 571–4.
118. Koechlin E, Basso G, Pietrini P, et al. The role of the anterior prefrontal cortex in human cognition. Nature 1999;399:148–51.
119. Jastreboff PJ, Jastreboff MM. Potential impact of stochastic resonance on tinnitus and its treatment. In: Abstracts of the 23rd Midwinter Meeting. Association for Research in Otolaryngology, St. Petersburg Beach, FL, 2000. www.aro.org.
120. Jaramillo F, Wiesenfeld K. Mechanoelectrical transduction assisted by Brownian motion: a role for noise in the auditory system. Nat Neurosci 1998; 1:384–8.
121. Morse RP, Evans EF. Enhancement of vowel coding for cochlear implants by addition of noise. Nat Med 1996;2:928–32.
122. McKinney CJ. An evaluation of the TRT method. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 99–105.
123. Jastreboff PJ, Jastreboff MM, Mattox DE. Statistical analysis of the progress of tinnitus treatment during Tinnitus Retraining Therapy (TRT). In: Abstracts of the 24th Midwinter Research Meeting. Association for Research in Otolaryngology, St. Petersburg Beach, FL, 2001. www.aro.org.
124. Sheldrake JB, Hazell JWP, Graham RL. Results of tinnitus retraining therapy. In Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 292–6.
125. Bartnik G, Fabijanska A, Rogowski M. Our experience in treatment of patients with tinnitus and/or hyperacusis using the habituation method. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 415–7.
126. Heitzmann T, Rubio L, Cardenas MR, Zofio E. The importance of continuity in TRT patients: results at 18 months. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 509–11.
127. Herraiz C, Hernandez FJ, Machado A, et al. Tinnitus retraining therapy: our experience. In: Hazell JWP, editor. Proceedings of the Sixth International Tinnitus Seminar, 1999, Cambridge, UK. London: Tinnitus and Hyperacusis Center; 1999. p. 483–4.
128. Vernon J, Press L. Treatment for hyperacusis. In: Vernon JA, editor. Tinnitus treatment and relief. Boston: Allyn and Bacon; 1998. p. 223–7.

CHAPTER 23

Cochlear Implants

Richard T. Miyamoto, MD, Karen Iler Kirk, PhD

Hearing loss poses a monumental obstacle to the acquisition and maintenance of effective communication skills. The perception and the production of speech are highly dependent on the ability to process auditory information. Early identification of hearing loss is an important first step in managing the effects of hearing impairment. Once identified, the level of residual hearing, if any, must be determined and an appropriate sensory aid recommended. Conventional amplification is usually the initial procedure of choice. If little or no benefit is realized with hearing aids, cochlear implants become therapeutic options. Communication skills and needs must be assessed and a communication mode selected. A sophisticated multidisciplinary team approach that addresses the varied needs of the deaf recipient is required. Essential components of the aural/oral (re)habilitation program include listening skill development, speech therapy, speech-reading training, and language instruction.

Djourno and Eyries first described direct electrical excitation of the auditory nerve in 1957.[1] Since then, increasingly more sophisticated cochlear implants have been developed. Cochlear implants seek to replace a nonfunctional inner ear hair cell transducer system by converting mechanical sound energy into electrical signals that can be delivered to the cochlear nerve in profoundly deaf patients. The essential components of a cochlear implant system are a microphone, which picks up acoustic information and converts it to electrical signals; an externally worn speech processor that processes the signal according to a predefined strategy; and a surgically implanted electrode array that is in the cochlea near the auditory nerve.

The processed signal is amplified and compressed to match the narrow electrical dynamic range of the ear. The typical response range of a deaf ear to electrical stimulation is on the order of only 10 to 20 dB, even less in the high frequencies. Transmission of the electrical signal across the skin from the external unit to the implanted electrode array is most commonly accomplished by the use of electromagnetic induction or radiofrequency transmission. The critical residual neural elements stimulated appear to be the spiral ganglion cells or axons. Damaged or missing hair cells of the cochlea are bypassed.

COCHLEAR IMPLANT SYSTEMS

Multichannel, multielectrode cochlear implant systems are designed to take advantage of the tonotopic organization of the cochlea. The incoming speech signal is filtered into a number of frequency bands, each corresponding to a given electrode in the array. Thus, multichannel cochlear implant systems use place coding to transfer spectral information in the speech signal as well as encode the durational and intensity cues of speech.

Nucleus Cochlear Implant Systems

The Nucleus 22-channel cochlear implant manufactured by Cochlear Ltd. of Australia was the first multichannel cochlear implant to receive US Food and Drug Administration (FDA) approval for use in adults and children, and it has been used in more patients than any other cochlear implant system worldwide.[2] The Nucleus CI24M cochlear implant received FDA approval for adults and children in 1998.

Early speech processing strategies (F0F2 and F0F1F2) for the Nucleus 22-channel cochlear implant used feature extraction strategies that conveyed information about key speech features such as the amplitude and frequency of vowel formants and the fundamental frequency of voiced sounds. The third-generation speech processing strategy, MPEAK, encoded additional high-frequency information by stimulating two of three more basal fixed

electrodes; the goal was to provide additional information that would yield improved consonant recognition scores.

Three processing strategies are currently available for use with the Nucleus cochlear implants. Two of the strategies use the *n-of-m* approach, in which the speech signal is filtered into *m* bandpass channels and the *n* highest envelope signals are selected for each cycle of stimulation.[3] The spectral peak (SPEAK) strategy is the most widely used with the Nucleus 22-channel cochlear implant and is available to users of either the Nucleus 22-channel or the Nucleus CI24M system. This strategy filters the incoming speech signal into 20 frequency bands; on each stimulation cycle, six electrodes (on average) are stimulated at a rate that varies adaptively from 180 to 300 pulses per second. An *n-of-m* strategy using much higher rates of stimulation, known as the Advanced Combined Encoder (ACE) strategy, can be implemented in the new Nucleus CI24M device. The third processing strategy available with the Nucleus CI24M system is the Continuous Interleaved Sampling (CIS) strategy.[4] The CIS strategy filters the speech signal into a fixed number of bands, obtains the speech envelope, and then compresses the signal for each channel. On each cycle of stimulation, a series of interleaved digital pulses rapidly stimulates consecutive electrodes in the array. The CIS strategy is designed to preserve fine temporal details in the speech signal by using high-rate, pulsatile stimuli.

Two different speech processors are available for new Nucleus cochlear implant recipients. The body-worn SPRINT processor (Cochlear Ltd., Australia) can implement any of the three current speech processing strategies. The ear-level ESPRIT speech processor (Cochlear Ltd., Australia) can currently implement only the SPEAK processing strategy.

Clarion Cochlear Implant System

The Clarion multichannel cochlear implant system is manufactured by Advanced Bionics Corporation (Sylmar, Calif.).[5–7] This device has been approved by the FDA for use in adults (1996) and children (1997). The Clarion multichannel cochlear implant has an eight-channel electrode array. Two processing strategies can be implemented through a body-worn processor. The first is CIS, described above, which is used to stimulate monopolar electrodes. The second strategy, Simultaneous Analog Stimulation (SAS), filters and then compresses the incoming speech signal for simultaneous presentation to the corresponding enhanced bipolar electrodes. The relative amplitudes of information in each channel and the temporal details of the waveforms in each channel convey speech information.

Medical Electronic (Med-El) Cochlear Implant System

The Combi 40+ cochlear implant system manufactured by the Med-El Corporation in Innsbruck, Austria, is currently undergoing clinical trials in the United States. The Med-El cochlear implant has 12 electrode pairs and has the capability of deep electrode insertion into the apical regions of the cochlea.[8] This device uses the CIS processing strategy and has the capacity to provide the most rapid stimulation rate of any of the cochlear implant systems currently available. Both body-worn and ear-level speech processors (the CIS Pro+ and Tempo+, respectively) are available for the Med-El cochlear implant.

New Developments in Cochlear Implant Electrode Design

New designs of the internal electrode array have recently been introduced for the Nucleus and Clarion cochlear implants. The Nucleus Contour electrode array is a curved electrode that is straightened by a stylet for insertion purposes. After surgical placement into the scala tympani, the stylet is withdrawn. The electrode then assumes its preformed shape, more closely approximating the modiolar wall of the cochlea (Figure 23–1). The Clarion HiFocus electrode is positioned closer to the modiolar wall by inserting a separate positioner into the scala tympani. Because the spiral ganglion cells are thought to be the sites stimulated by cochlear implants, directing the electrodes toward the modiolus and further positioning the array may improve spatial specificity of stimulation and reduce the current need to drive the electrodes.[9]

PATIENT SELECTION

The selection of cochlear implant candidates is a complex process that requires careful consideration of many factors. Current selection considerations are as follows.

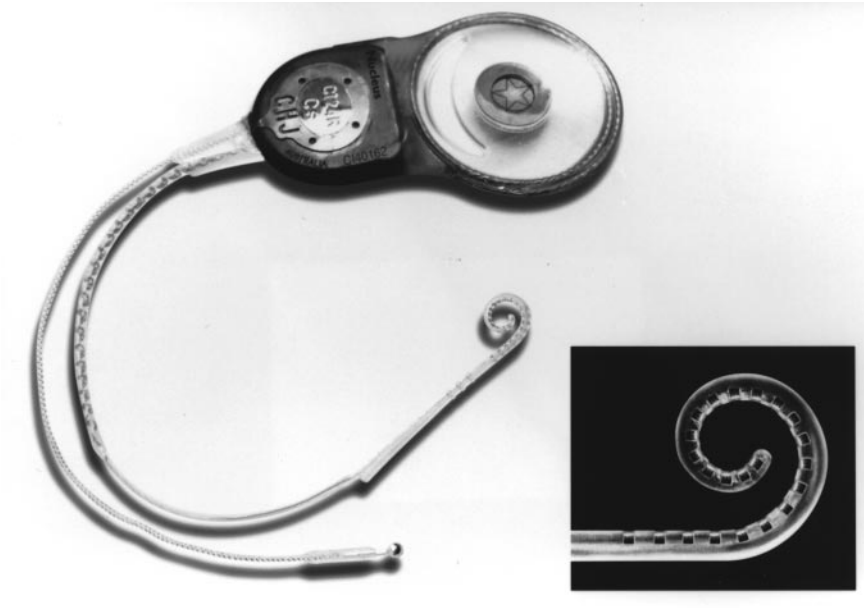


FIGURE 23–1. New electrode design that hugs the modiolus after the insertion stylus is withdrawn. Courtesy of Cochlear Ltd.

Adults

Cochlear implantation was initially limited to postlingually deafened adults who received no benefit from hearing aids and had no possibility of worsening residual hearing. This population, particularly those with a recent onset of deafness, has been the most readily identifiable beneficiary of cochlear implants. A period of auditory experience adequate to develop normal speech perception, speech production, and language skills before the onset of deafness is an invaluable prerequisite. Experience gained with this initial cochlear implant population served to establish expected performance limits.[10]

Table 23–1 presents the current cochlear implant candidacy criteria for adults. Adult candidacy criteria are based primarily on aided speech recognition abilities. No upper age limit is used in the selection process as long as the patient's health will permit an elective surgical procedure under general anesthesia. In a survey of Nucleus 22-channel cochlear implant recipients over the age of 65 years, Horn et al showed that elderly cochlear implant patients obtained benefits that were similar to those obtained by younger adult patients with the same device.[11]

Adult candidacy criteria recently have been broadened to include adults with severe-to-profound hearing loss who derive some limited benefit from conventional hearing aids. Implantation of an ear with any residual, aidable hearing carries the risk that the implanted ear could be made worse than that ear with a hearing aid. Current investigations are testing the hypothesis that an ear with some residual hearing may have a better neuronal population, increasing the likelihood of superior performance with a cochlear implant, especially with more complex multichannel stimulation.

Adults with prelingual hearing loss are generally not considered good candidates for cochlear implantation.[12] However, prelinguistically deafened adults who have followed an aural/oral educational approach may receive significant benefit.

TABLE 23–1. Adult Candidacy Criteria for Cochlear Implantation

Criteria
≥ 18 years of age
Bilateral severe-to-profound hearing loss
Minimal benefit from conventional hearing aids (typically defined as sentence recognition scores < 50 to 60% correct in the best aided condition)
No medical contraindications

Children

Cochlear implant technology complicates the already challenging management of the deaf child. The general selection guidelines applied to adults are applicable to children; however, the selection of pediatric cochlear implant candidates is a far more complex and ever-evolving process that requires careful consideration of many factors. Pediatric candidacy criteria are presented in Table 23–2. In contrast to adults, both pre- and postlingually deafened children are candidates for cochlear implantation.

A trend toward earlier cochlear implantation in children has emerged in an attempt to ameliorate the devastating effects of early auditory deprivation. Electrical stimulation appears to be capable of preventing at least some of the degenerative changes in the central auditory pathways.[13]

Implanting very young children remains controversial because the audiologic assessment, surgical intervention, and postimplant management in this population are challenging. Profound deafness must be substantiated and the inability to benefit from conventional hearing aids demonstrated. This can be difficult to determine in young children with limited language abilities. For very young children, parental questionnaires are commonly used to assess amplification benefit.

Until recently, the youngest age at which children could be implanted under FDA clinical trials was 2 years. In 1998, the age limit was dropped to 18 months, and the most recently initiated clinical trials of new cochlear implant systems permit implantation of children as young as 12 months of age. Because the development of speech perception, speech production, and language competence normally begins at a very early age, implantation in very young congenitally or neonatally deafened children may have substantial advantages. Early implantation may be particularly important when the etiology of deafness is meningitis as progressive intracochlear ossification can occur and preclude standard electrode insertion. A relatively short window of time exists during which this advancing process can be circumvented.

Special consideration must be given to the small dimensions of the temporal bone and to potential problems from postoperative temporal bone growth. In addition, the high incidence of otitis media in children under the age of 2 years might compromise the biosafety of cochlear implants. None the less, extension of implant candidacy to the 6- to 12-month age group is feasible on an anatomic basis. The cochlea is adult size at birth, and by age 1 year, the facial recess and mastoid antrum, which provide access to the middle ear for electrode placement, are adequately developed.[14]

CLASSIFICATION OF COCHLEAR IMPLANT RECIPIENTS

Cochlear implant recipients can be divided into three main categories. Significantly different performance outcomes can be anticipated:

1. *Postlingually deafened adults and children.* Patients who become deaf at or after age 5 years are generally classified as postlingually deafened. These patients have developed many or all aspects of spoken language before the onset of their deafness. However, once they lose access to auditory input and feedback, they frequently

TABLE 23–2. Pediatric Candidacy Criteria for Cochlear Implantation

Children Aged 12 to 23 Mo	*Children Aged 24 Mo to 17 Yr, 11 Mo*
Bilateral profound hearing loss	Bilateral severe-to-profound hearing loss
Lack of auditory skills development and minimal hearing aid benefit (documented by parent questionnaire)	Lack of auditory skills development and minimal hearing aid benefit (eg, word recognition scores < 30% correct)
No medical contraindications	No medical contraindications
Enrolment in a therapy of education program emphasizing auditory development	Enrolment in a therapy of education program emphasizing auditory development

demonstrate rapid deterioration in the intelligibility of their speech. Implantation soon after the onset of deafness can potentially ameliorate this rapid deterioration in speech production and perception abilities. A postlingual onset of deafness is an infrequent occurrence in the pediatric population. If this were to be the only category for which cochlear implants positively impacted deaf children, there would be limited applicability for this technology in children.

2. *Congenitally or early deafened children.* Congenital or early acquired deafness is the most frequently encountered type of profound sensorineural hearing loss in children. The acquisition of oral communication skills can be a difficult process for these children. However, with early implantation and appropriate habilitation, many children in this category are developing spoken language. Although considerable outcome variability exists, the most successful pediatric cochlear implant recipients demonstrate age-appropriate speech and language skills.
3. *Congenitally or early deafened adolescents and adults.* When cochlear implantation is considered in adolescence or young adulthood for a patient who has had little or no experience with sound because of congenital or early-onset deafness, caution must be exercised because this group has not demonstrated high levels of success with electrical stimulation of the auditory system.

AUDIOLOGIC ASSESSMENT

The audiologic evaluation is the primary means of determining suitability for cochlear implantation. Audiologic evaluations should be conducted in both an unaided condition and with appropriately fit conventional amplification. Thus, all potential candidates must have completed a period of experience with a properly fit hearing aid, preferably coupled with training in an appropriate aural re(habilitation) program. The audiologic evaluation includes measurement of pure-tone thresholds along with word and sentence recognition testing. Aided speech recognition scores are the primary audiologic determinant of cochlear implant candidacy. For very young children or those with limited language abilities, parent questionnaires are used to determine hearing aid benefit.

MEDICAL ASSESSMENT

The medical assessment includes the otologic history and physical examination. Radiologic evaluation of the cochlea is performed to determine whether the cochlea is present and patent and to identify congenital deformities of the cochlea. High-resolution, thin-section computed tomographic (CT) scanning of the cochlea remains the imaging technique of choice.[15] Intracochlear bone formation resulting from labyrinthitis ossificans can usually be demonstrated by CT scanning; however, when soft tissue obliteration occurs following sclerosing labyrinthitis, CT may not image the obstruction. In these cases, T_2-weighted magnetic resonance imaging is an effective adjunctive procedure providing additional information regarding cochlear patency. The endolymph/perilymph signal may be lost in sclerosing labyrinthitis. Intracochlear ossification is not a contraindication to cochlear implantation but can limit the type and insertion depth of the electrode array that can be introduced into the cochlea. Congenital malformations of the cochlea are likewise not contraindications to cochlear implantation. Cochlear dysplasia has been reported to occur in approximately 20% of children with congenital sensorineural hearing loss.[16] Several reports of successful implantations in children with inner ear malformations have been published.[17–21] A thin cribriform area between the modiolus and a widened internal auditory canal is often observed[22] and is believed to be the route of egress of cerebrospinal fluid (CSF) when it occurs during surgery or postoperatively. A CSF gusher was reported in several patients. Temporal bone dysplasia also may be associated with an anomalous facial nerve, which may increase the surgical risk.

The precise etiology for the deafness cannot always be determined but is identified whenever possible; however, stimulable auditory neural elements are nearly always present regardless of cause of deafness.[23] Two exceptions are the *Michel deformity*, in which there is congenital agenesis of the cochlea, and the *small internal auditory canal syndrome*, in which the cochlear nerve may be congenitally absent.

Otoscopic evaluation of the tympanic membrane is performed. An otologically stable condition should be present prior to considering implantation. The ear proposed for cochlear implantation must be

free of infection, and the tympanic membrane should be intact. If these conditions are not met, medical or surgical treatment before implantation is required. The management of middle ear effusions in children who are under consideration for cochlear implantation or who already have a cochlear implant deserves special consideration. Conventional antibiotic treatment usually accomplishes this goal, but when it does not, treatment by myringotomy and insertion of tympanostomy tubes may be required. Removal of the tube several weeks before cochlear implantation usually results in a healed, intact tympanic membrane. When an effusion occurs in an ear with a cochlear implant, no treatment is required as long as the effusion remains uninfected. Chronic otitis media, with or without cholesteatoma, must be resolved before implantation; this is accomplished with conventional otologic treatments. Prior ear surgery that has resulted in a mastoid cavity does not contraindicate cochlear implantation, but this situation may require mastoid obliteration with closure of the external auditory canal or reconstruction of the posterior bony ear canal.

PSYCHOLOGICAL ASSESSMENT

Psychological testing is performed for exclusionary reasons to identify subjects who have organic brain dysfunction, mental retardation, undetected psychosis, or unrealistic expectations. Valuable information related to the family dynamics and other factors in the patient's milieu that may affect implant acceptance and performance are assessed.

SURGICAL IMPLANTATION

Cochlear implantation in both children and adults requires meticulous attention to the delicate tissues and small dimensions. Skin incisions are designed to provide access to the mastoid process and coverage of the external portion of the implant package while preserving the blood supply of the postauricular skin. The incision used at the Indiana University Medical Center has eliminated the need to develop a large postauricular flap. The inferior extent of the incision is made well posterior to the mastoid tip to preserve the branches of the postauricular artery. From here the incision is directed posterior-superiorly and then superiorly, without a superior-anterior limb. In children, the incision incorporates the temporalis muscle to give added thickness. A subperiosteal pocket is created for positioning the implant induction coil (Figure 23–2). A bone pocket well tailored to the device being implanted is created, and the induction coil is fixed to the cortex with a fixation suture or periosteal flaps.

Following development of the skin incision, a mastoidectomy is performed. The horizontal semicircular canal is identified in the depths of the mastoid antrum, and the short process of the incus is identified in the fossa incudis. The facial recess is opened using the fossa incudis as an initial landmark. The facial recess is a triangular area bound by (a) the fossa incudis superiorly, (b) the chorda tympani nerve laterally and anteriorly, and (c) the facial nerve medially and posteriorly (Figure 23–3). The facial nerve can usually be visualized through the bone without exposing it. The round window niche is visualized through the facial recess about 2 mm inferior to the stapes. Occasionally, the round window niche is posteriorly positioned and is not well visualized through the facial recess or is obscured by ossification. Particularly in these situations, it is important not to be misdirected by hypotympanic air cells. Entry into the scala tympani is accomplished best through a cochleostomy created anterior and inferior to the annulus of the round window membrane. A small fenestra slightly larger than the electrode to be implanted (usually 0.5 mm) is developed. A small diamond bur is used to "blue line" the endosteum of the scala tympani, and the endosteal membrane is removed by using small picks. This approach bypasses the hook area of the scala tympani, allowing direct insertion of the active electrode array (Figure 23–4). After insertion of the active electrode array, the cochleostomy area is sealed with small pieces of fascia.

Special Surgical Considerations

Cochlear Dysplasia In cases of cochlear dysplasia, a CSF gusher may be encountered on fenestrating the cochlea while performing the cochleostomy. The flow of CSF has been successfully controlled by entering the cochlea through a small fenestra, allowing the CSF reservoir to drain off, inserting the electrode into the cochleostomy, and tightly packing the electrode at the cochleostomy with fascia. It is postulated that the source of the leak is through the lateral end of the internal auditory canal. Supplementally, a lumbar drain can be placed to reduce the

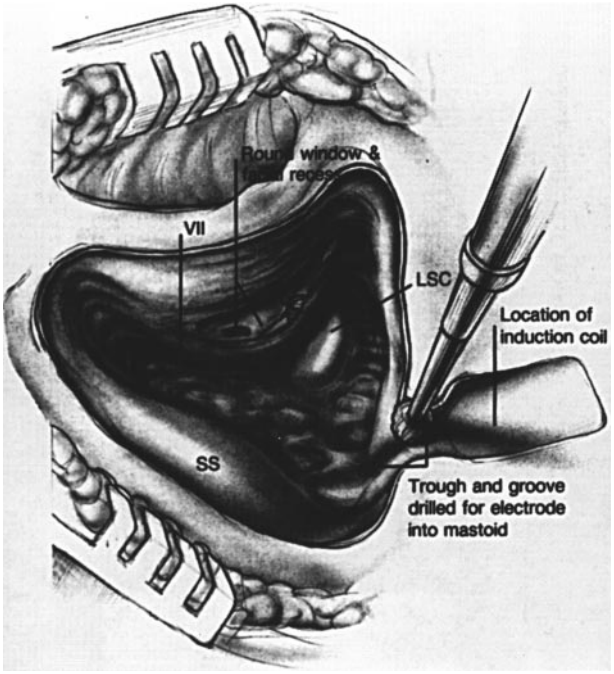


FIGURE 23–2. Mastoidectomy performed to gain access to facial recess. Location of induction coil and trough and groove drilled for electrode implantation into the mastoid. VII = facial nerve VII; LSC = lateral semicircular canal; SS = sigmoid sinus.

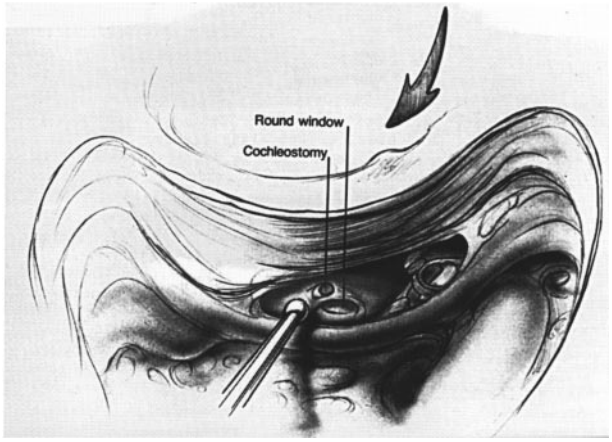


FIGURE 23–3. The facial recess is bounded by the fossa incudis superiorly, the chorda tympani nerve laterally and anteriorly, and the facial nerve medially and posteriorly. A cochleostomy is created anterior and inferior to the round window.

spinal fluid reservoir until a satisfactory tissue seal has occurred. In severe dysplasia cases with a common cavity deformity, the electrode array may be inserted directly by a transmastoid labyrinthotomy approach. The otic capsule is opened posterosuperior to the second genu of the facial nerve, and the common cavity is entered. Several patients have been treated in this way with no vestibular side effects.[24]

Aberrant Facial Nerve In patients who have malformations of the labyrinth, and occasionally in patients with otherwise normal anatomy, the facial nerve may follow an aberrant course. Although not all aberrant facial nerves impact cochlear implant surgery, those that do must be recognized and dealt with effectively. Two anomalous courses of the facial nerve that place it at risk are the laterally and anteriorly displaced vertical portion of the facial nerve and a facial nerve that courses over the promontory over or anterior to the round window.[25]

Intracochlear Ossification Ossification at the round window is common in patients after meningitis and has been encountered in approximately one half of the children whose cause of deafness was meningitis who have received a cochlear implant at our center. In these patients, a cochleostomy is developed anterior to the round window and the new bone is drilled until an open scala is entered. A full electrode insertion can then be accomplished. Less frequently, labyrinthitis ossificans with extensive intracochlear bone formation may occur with complete obliteration of the scala tympani. In these cases, our preference has been to drill open the basal turn of the cochlea and create a tunnel approximately 6 mm deep and partially insert a Nucleus electrode. This allows implantation of 10 to 12 active electrodes, which has yielded very satisfactory results. More recently, a specially designed split electrode developed by the Med-El Corporation has been used wherein one branch of the electrode array is placed in the tunnel described above and the second active electrode is inserted into an additional cochleostomy developed just anterior to the oval window. Gantz and his colleagues described an extensive drill-out procedure to gain access to the upper basal turn.[26] Steenerson et al described insertion of the active electrode into the scala vestibuli in cases of cochlear ossification.[27] Although this procedure has merit, the scala vestibuli is frequently ossified when the scala tympani is completely obliterated.

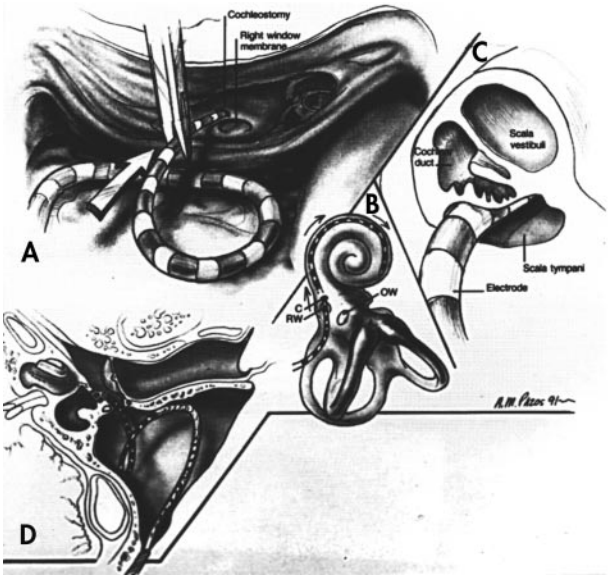


FIGURE 23–4. A, Electrode introduced into the basal turn of the cochlea. *B*, Electrode in the scala tympani (*C*, cross-section of the basal turn on the cochlea). *D*, Receiver-stimulator with a redundant loop of electrode in the mastoid.

Complications Complications have been infrequent with cochlear implant surgery and can largely be avoided by careful preoperative planning and meticulous surgical technique. Among the most commonly encountered problems are those associated with the incision and postauricular flap and facial nerve injury.[28] Using the incision we describe, we have experienced only one flap breakdown in our pediatric cochlear implant population. (This occurred several years postoperatively after head trauma.) We experienced one transient delayed facial paresis and one CSF gusher in a child with a Mondini deformity.[29] Several additional patients with the large vestibular aqueduct syndrome have also had gushers.

Because children are more susceptible to otitis media than adults, justifiable concern has been expressed that a middle ear infection could cause an implanted device to become an infected foreign body, requiring its removal. Two children in our series experienced a delayed mastoiditis (several years after the implant surgery), resulting in a postauricular abscess. These patients were treated by incision and drainage and intravenous antibiotics without the need to remove the implant. Of even greater concern is that infection might extend along the electrode into the inner ear, resulting in a serious otogenic complication, such as meningitis or further degeneration of the central auditory system. To date, although the incidence of otitis media in children who have received cochlear implants parallels that seen in the general pediatric population, no serious complications related to otitis media have occurred in our patients.

CLINICAL RESULTS

Cochlear implants are an established therapeutic option for selected deaf adults and children. However, there remains a wide range of performance with current implant systems. Some cochlear implant recipients can communicate without the benefit of speech reading and are able to communicate on the telephone without a telephone code, whereas others use their implants primarily to reestablish environmental contact and enhance their speech-reading abilities. This variation in performance levels is thought to relate to biologic and cognitive factors. It would be expected that poor auditory nerve survival or atrophic central auditory systems would correlate with poor performance, whereas a more intact auditory nervous system should permit better results, given a well-designed and fitted cochlear prosthesis.[30]

Performance Results in Adults

Early Nucleus cochlear implant systems using feature extraction speech processing strategies yielded moderate vowel and consonant recognition[31] but only limited auditory-only word recognition.[32] However, with each successive generation of Nucleus speech processing strategy, substantial gains were achieved in open-set word and sentence recognition and were reported with mean word and sentence recognition scores of 36 and 74%, respectively, for adults with the SPEAK strategy.[33–38] Similar performance levels have been reported for adults who use either the Clarion or the Med-El cochlear implant system.[39,40] Compared with the results obtained with previous generations of cochlear implants, adults who use the current devices achieve higher word recognition skills and acquire those skills at a faster rate. Many adults now demonstrate substantial speech understanding as early as 3 months following cochlear implantation.[41] On average, multichannel cochlear implant systems provide

moderate to good levels of auditory-only speech understanding to the majority of adult users. However, a great deal of variability in performance remains. Some adults are unable to understand any speech through listening alone, whereas others can communicate successfully on the telephone. As Wilson and his colleagues pointed out, a number of within-subject factors also contribute to successful cochlear implant use.[42] Two such factors are age at implantation and duration of deafness.[41,43–47] Specifically, patients who are implanted at a young age and have a shorter period of auditory deprivation are more likely to achieve good outcomes. The findings regarding other predictive factors have been less conclusive. For example, Gantz et al found that measures of cognitive ability were not associated with patient performance,[32] whereas Cohen et al reported that measures of IQ were significantly associated with good speech perception skills.[44] Other factors that have been found to significantly correlate with adult outcomes include speech-reading ability[44,45] and degree of residual hearing.[45,48]

Performance Results in Children

Postlingually deafened adults and children use the information transmitted by a cochlear implant to make comparisons to previously stored representations of spoken language. However, the majority of children who receive cochlear implants have congenital or prelingually acquired hearing loss. These children must use the sound provided by a cochlear implant to acquire speech perception, speech production, and spoken language skills. Furthermore, because young children have limited linguistic skills and attention spans, the assessment of performance in this population can be quite challenging. To evaluate the communication benefits of cochlear implant use in children effectively, a battery of tests that are developmentally and linguistically appropriate should be employed.[49,50]

Speech Perception Outcomes In early investigations, children who used the Nucleus cochlear implant with a feature extraction strategy demonstrated significant improvement in closed-set word identification (ie, the ability to identify words from a limited set of alternatives) but very limited open-set word recognition.[51,52] The introduction of newer processing strategies yielded greater speech perception benefits in children, just as in adults. Many children with current cochlear implant devices achieve at least moderate levels of open-set word recognition. For example, Cohen et al reported word recognition scores for a group of 19 children that ranged from 4 to 76% words correct with a mean of 44%.[53] Similarly, Osberger et al reported average scores of approximately 30% correct on a more difficult measure of isolated word recognition in children with the Clarion cochlear implant.[54] The development rate of postimplant auditory skills seems to be increasing as cochlear implant technology improves and as children are implanted at a younger age.[53–55] Furthermore, comparison studies have shown that the speech perception abilities of pediatric cochlear implant recipients meet or exceed those of their peers with unaided pure-tone average thresholds ≥ 90 dB HL who use hearing aids.[56,57]

A number of demographic factors have been shown to influence performance results in children with cochlear implants. Early results suggested better speech perception performance in children deafened at an older age, with a corresponding shorter period of deafness.[52,58,59] However, when only children with prelingual deafness (ie, < 3 years) were considered, age at onset of hearing loss was no longer a significant factor.[60] It is evident that earlier implantation yields superior cochlear implant performance in children.[61–66] Although the critical period for implantation of congenitally or prelingually deafened children has not been determined,[67] preliminary evidence suggests that implantation prior to age 3 years may yield improved results.[68–70] Finally, the variables of communication mode and/or unaided residual hearing also influence speech perception performance.[71–74] Oral children, and those who have more residual hearing prior to implantation, typically demonstrate superior speech understanding. This has led to some controversy regarding whether to implant the better- or the poorer-hearing ear.[48,72]

Speech Intelligibility and Language Improvements in speech perception are the most direct benefit of cochlear implantation. However, if children with cochlear implants are to succeed in the hearing world, they must also acquire intelligible speech and their surrounding linguistic system. The speech intelligibility and language abilities of children with cochlear implants improve significantly over time[67,75–78] and, on average, exceed those of their

age- and hearing-matched peers with hearing aids.[77,78] Speech intelligibility and spoken language acquisition are significantly correlated with the development of auditory skills.[76,79] Although a great deal of variability exists, the best pediatric cochlear implant users demonstrate highly intelligible speech and age-appropriate language skills. These superior performers are usually implanted at a young age and are educated in an oral/aural modality.[76]

CONCLUSION

Cochlear implants are an appropriate sensory aid for selected deaf children and adults who receive minimal benefit from conventional amplification. Improvements in technology and refinements in candidacy criteria have secured a permanent role for cochlear implantation. With improved postoperative performance, a clear justification for implanting not only patients with bilateral profound sensorineural hearing loss but also patients with severe sensorineural hearing loss has been established. Patients as young as 12 months of age may be implanted under current FDA guidelines for clinical trials, and experience with even younger children is accumulating.

Wide intersubject performance variability continues to exist. However, most postlingually deafened adults with current cochlear implants achieve auditory-only word recognition and communicate very effectively when auditory cues are combined with speech reading. The best adult recipients can converse fluently without speech-reading cues. Children using cochlear implants have acquired speaking and listening skills and have developed a spoken language system that is beyond what previously could be achieved with hearing aids. Children who are implanted at a young age and use oral communication have the best prognosis for developing intelligible speech and age-appropriate language abilities.

Challenges remain in effectively assessing peripheral auditory neuronal survival and matching electrically transmitted signals to the future potential of the central auditory system in deaf subjects.

ACKNOWLEDGMENT

This work was supported in part by National Institutes of Health National Institute on Deafness and Other Communication Disorders grants RO1 DC00064, RO1 DC00423, and K23 DC00126 and by Psi Iota Xi.

REFERENCES

1. Djourno A, Eyries C. Prothese auditive par excitation electrique a distance du nerf sensoriel a l'aide d'un bobinage inclus a demeure. Presse Med 1957;35:14–7.
2. Clark GM, Tong YC, Dowell RC, et al. A multiple-channel cochlear implant: an evaluation using nonsense syllables. Ann Otol Rhinol Laryngol 1981; 90:227–30.
3. Wilson BS. Strategies for representing speech information with cochlear implants. In: Niparko JK, Kirk KI, Robbins AM, et al, editors. Cochlear implants: principles and practices. Philadelphia: Lippincott Williams & Wilkins; 2000. p. 129–70.
4. Wilson BS, Lawson DT, Finley CC, et al. Coding strategies for multichannel cochlear prostheses. Am J Otol 1991;12 Suppl 1:56–61.
5. Schindler RA, Kessler DK. The UCSF/Storz cochlear implant: patient performance. Am J Otol 1987; 8:247–55.
6. Schindler RA, Kessler DK. Clarion cochlear implant: phase I investigational results. Am J Otol 1993; 14:263–72.
7. Kessler DK, Schindler RA. Progress with a multistrategy cochlear system: the Clarion. In: Hochmair-Desoyer IJ, Hochmair ES, editors. Advances in cochlear implants. Wein, Austria: Manz; 1994. p. 354–62.
8. Gstoettner WK, Baumgartner WD, Franz P, et al. Cochlear implant deep insertion surgery. Laryngoscope 1997;107:544–6.
9. Wilson BS. Cochlear implant technology. In: Niparko JK, Kirk KI, Robbins AM, et al, editors. Cochlear implants: principles and practices. Philadelphia: Lippincott Williams & Wilkins; 2000. p. 109–27.
10. Berliner KI. Selection of cochlear implant patients. In: Schindler RA, Merzenich MM, editors. Cochlear implants. New York: Raven Press; 1985. p. 395–402.
11. Horn KL, McMahon NB, McMahon DC, et al. Functional use of the Nucleus 22-channel cochlear implant in the elderly. Laryngoscope 1991;101: 284–8.
12. Waltzman SB, Cohen NL. Implantation of patients with prelingual long-term deafness. Ann Otol Rhinol Laryngol 1999;108:84–7.

13. Matsushima JI, Shepard RK, Seldon HL, et al. Electrical stimulation of the auditory nerve in deaf kittens: effects on cochlear nucleus morphology. Hear Res 1991;56:133–42.
14. Lenarz T. Cochlear implantations in children under the age of two years. Adv Otorhinolaryngol 1997; 52:204–10.
15. Yune HY, Miyamoto RT. Medical imaging in cochlear implant candidates. Am J Otol 1991;12 Suppl:11–7.
16. Jensen J. Tomography of the inner ear in deaf children. Radiological demonstration of two cases with the Mondini malformation. J Laryngol Otol 1967; 81:27–35.
17. Mangabeira-Albernaz PL. The Mondini dysplasia: from early diagnosis to cochlear implant. Acta Otolaryngol (Stockh) 1983;95:627–31.
18. Miyamoto RT, McConkey AJ, Myres WA, et al. Cochlear implantation in the Mondini inner ear malformation. Am J Otol 1986;7:258–61.
19. Jackler RK, Luxford WM, House WF. Sound detection with the cochlear implant in five ears of four children with congenital malformations of the cochlea. Laryngoscope 1987;97 Suppl 40:15–7.
20. Silverstein H, Smouha E, Morgan N. Multichannel cochlear implantation in a patient with bilateral Mondini deformities. Am J Otol 1988;9:451–5.
21. Tucci DL, Telian SA. Cochlear implantation in patients with cochlear malformations. Arch Otolaryngol Head Neck Surg 1995;121:833–8.
22. Schuknecht HF. Mondini dysplasia: a clinical and pathological study. Ann Otol Rhinol Laryngol 1980;89 Suppl 65:1–23.
23. Hinojosa R, Marion M. Histopathology of profound sensorineural deafness. Ann N Y Acad Sci 1983; 405:459–84.
24. McElveen JT, Carrasco VN, Miyamoto RT, Linthicum FH. Cochlear implantation in common cavity malformations using a transmastoid labyrinthotomy approach. Laryngoscope 1997;107:1032–6.
25. Miyamoto RT, Kaiser AR. Facial nerve anomalies in cochlear implantation. Adv Otorhinolaryngol 2000;57:131–3.
26. Gantz BJ, McCabe BF, Tyler RS. Use of multichannel cochlear implants in obstructed and obliterated cochleas. Otolaryngol Head Neck Surg 1988;98: 72–81.
27. Steenerson RL, Gary LB, Wynens MS. Scala vestibuli cochlear implantations for labyrinthine ossification. Am J Otol 1990;11:360–3.
28. Hoffman RA, Cohen NL. Complications of cochlear implant surgery. Ann Otol Rhinol Laryngol 1995;104 Suppl 166:420–2.
29. Miyamoto RT, Young M, Myres WA, et al. Complications of pediatric cochlear implantation. Eur Arch Otorhinolaryngol 1996;253:1–4.
30. Kessler DK, Loeb GE, Barker MJ. Distribution of speech recognition results with the Clarion cochlear prosthesis. Ann Otol Rhinol Laryngol 1995;104 Suppl 166:283–5.
31. Blamey PJ, Dowell RC, Brown AM, et al. Vowel and consonant recognition of cochlear implant patients using formant-estimating speech processors. J Acoust Soc Am 1987;82:48–57.
32. Gantz B, Tyler RS, Abbas PJ, et al. Evaluation of five different cochlear implant designs: audiologic assessment and predictors of performance. Laryngoscope 1988;98:1100–6.
33. Skinner MW, Holden LK, Holden TA, et al. Performance of postlinguistically deaf adults with the Wearable Speech Processor (WSP III) and Mini Speech Processor (MSP) of the Nucleus multi-electrode cochlear implant. Ear Hear 1991;12:3–22.
34. Hollow RD, Dowell RC, Cowan RS, et al. Continuing improvements in speech processing for adult cochlear implant patients. Ann Otol Rhinol Laryngol Suppl 1995;166:292–4.
35. Parkinson AJ, Tyler RS, Woodworth GG, et al. A within-subject comparison of adult patients using the Nucleus F0F1F2 and F0F1F2B3B4B5 speech processing strategies. J Speech Hear Res 1996;39:261–77.
36. Holden LK, Skinner MW, Holden TA. Speech recognition with the MPEAK and SPEAK speech-coding strategies of the Nucleus cochlear implant. Otolaryngol Head Neck Surg 1997;166:163–7.
37. Staller S, Menapace C, Domico E, et al. Speech perception abilities of adult and pediatric Nucleus implant recipients using the Spectral Peak (SPEAK) coding strategy. Otolaryngol Head Neck Surg 1997;117(3 Pt 1):236–42.
38. Fraysee B, Diller N, Klenzer T, et al. Cochlear implants for adults obtaining marginal benefit from acoustic amplification. Am J Otol 1998;19:591–7.
39. Schindler RA, Kessler DK, Barker M. Clarion patient performance: an update on the clinical trials. Ann Otol Rhinol Laryngol 1995;104 Suppl 166:269–72.
40. Helms J, Muller J, Schon F, et al. Evaluation of performance with the COMBI 40 cochlear implant in adults: a multicentric clinical study. ORL J Otorhinolaryngol Relat Spec 1997;59:23–35.

41. Geier LL, Fisher LM, Barker EJ, et al. The effect of long-term deafness on speech recognition in postlingually deafened adult Clarion cochlear implant users. Ann Otol Rhinol Laryngol 1999; 108:80–3.
42. Wilson BS, Lawson DS, Finley CC, et al. Importance of patient and processor variables in determining outcomes with cochlear implants. J Speech Hear Res 1993;36:373–9.
43. Blamey PJ, Pyman BC, Clark GM, et al. Factors predicting postoperative sentence scores in postlingually deaf adult cochlear implant patients. Ann Otol Rhinol Laryngol 1992;101:342–8.
44. Cohen NL, Waltzman SB, Fisher SG, et al. A prospective, randomized study of cochlear implants. The Department of Veterans Affairs Cochlear Implant Study Group. N Engl J Med 1993;328:233–7.
45. Gantz B, Woodworth G, Knutson JF, et al. Multivariate predictors of audiological success with multichannel cochlear implants. Ann Otol Rhinol Laryngol 1993;102:909–16.
46. Battmer RD, Gupta SP, Allum-Mecklenburg DJ, et al. Factors influencing cochlear implant perceptual performance in 132 adults. Ann Otol Rhinol Laryngol Suppl 1995;166:185–7.
47. Shipp D, Nedzelski J, Chen J, Hanusaik L. Prognostic indicators of speech recognition performance in postlinguistically deafened adult cochlear implant users. Adv Otorhinolaryngol 1997;52:74–7.
48. Rubinstein JT, Miller CA. How do cochlear prostheses work? Curr Opin Neurobiol 1999;9:399–404.
49. Kirk KI, Diefendorf AO, Pisoni DB, et al. Assessing speech perception in children. In: Mendel L, Danhauer J, editors. Audiological evaluation and management and speech perception training. San Diego: Singular; 1997. p. 101–32.
50. Kirk KI. Challenges in the clinical investigation of cochlear implant outcomes. In: Niparko JK, Kirk KI, Mellon NK, et al, editors. Cochlear implants: principles and practices. Philadelphia: Lippincott-Raven; 2000. p. 225–59.
51. Miyamoto RT, Osberger MJ, Robbins AJ, et al. Comparison of sensory aids in deaf children. Ann Otol Rhinol Laryngol 1989;98(Suppl 8 Pt 2):2–7.
52. Staller SJ, Beiter AL, Brimacombe J, et al. Pediatric performance with the Nucleus 22-channel cochlear implant system. Am J Otol 1991;12 Suppl:126–36.
53. Cohen NL, Waltzman SB, Roland JT, et al. Early results using the Nucleus C124M in children. Am J Otol 1999;20:198–204.
54. Osberger MJ, Zimmerman-Phillips S, Barker M, Geier L. Clinical trials of the CLARION cochlear implant in children. Ann Otol Rhinol Laryngol 1999;177 Suppl 108:88–92.
55. Young NM, Carrasco VN, Grohne KM, Brown CJ. Speech perception of young children using Nucleus 22-channel or Clarion cochlear implants. Ann Otol Rhinol Laryngol 1999;108:99–103.
56. Meyer TA, Svirsky MA, Kirk KI, Miyamoto RT. Improvements in speech perception by children with profound prelingual hearing loss: effects of device, communication mode, and chronological age. J Speech Lang Hear Res 1998;41:846–58.
57. Svirsky MA, Meyer TA. Comparison of speech perception in pediatric Clarion cochlear implant and hearing aid users. Ann Otol Rhinol Laryngol 1999;108 Suppl 177:104–9.
58. Osberger MJ, Todd SL, Berry SW, et al. Effect of age at onset of deafness on children's speech perception abilities with a cochlear implant. Ann Otol Rhinol Laryngol 1991;100:883–8.
59. Fryauf-Bertschy H, Tyler RS, Kelsay DM, et al. Performance over time of congenitally deaf and postlingually deafened children using a multichannel cochlear implant. J Speech Hear Res 1992;35:913–20.
60. Miyamoto RT, Osberger MJ, Robbins AM, et al. Prelingually deafened children's performance with the Nucleus multichannel cochlear implant. Am J Otol 1993;14:437–45.
61. Fryauf-Bertschy H, Tyler RS, Kelsay DM, et al. Cochlear implant use by prelingually deafened children: the influences of age at implant and length of device use. J Speech Lang Hear Res 1997;40:183–99.
62. Miyamoto RT, Kirk KI, Robbins AM, et al. Speech perception and speech intelligibility in children with multichannel cochlear implants. Adv Otorhinolaryngol 1997;52:198–203.
63. O'Donoghue GM, Nikolopoulos TP. Speech perception in children after cochlear implantation. Am J Otol 1999;19:762–7.
64. Nikolopoulos TP, O'Donoghue GM, Archbold SM. Age at implantation: its importance in pediatric cochlear implantation. Laryngoscope 1999;109: 595–9.
65. Lenarz T, Illg A, Lesinki-Schiedat A, et al. Cochlear implantation in children under the age of two: the MHH experience with the Clarion cochlear implant. Ann Otol Rhinol Laryngol 1999;108:44–9.
66. Illg A, Lesinki-Schiedat A, von der Haar-Heise S, et al. Speech perception results for children implanted

with the Clarion cochlear implant at the Medical University of Hannover. Ann Otol Rhinol Laryngol 1999;108:93–8.
67. Brackett D, Zara CV. Communication outcomes related to early implantation. Am J Otol 1998; 19:453–9.
68. Waltzman S, Cohen N, Shapiro W. Effects of cochlear implantation on the young deaf child. Adv Otorhinolaryngol 1995;50:125–8.
69. Waltzman S, Cohen NL, Gomolin R, et al. Perception and production results in children implanted between two and five years of age. Adv Otorhinolaryngol 1997;52:177–80.
70. Waltzman SB, Cohen NL. Cochlear implantation in children younger than 2 years old. Am J Otol 1998;19:158–62.
71. Cowan RS, DelDot J, Barker EJ, et al. Speech perception results for children with implants with different levels of preoperative residual hearing. Am J Otol 1997;18 Suppl:125–6.
72. Zwolan TA, Zimmerman-Phillips S, Ashbaugh CJ, et al. Cochlear implantation of children with minimal open-set speech recognition skills. Ear Hear 1997; 18:240–51.
73. Osberger MJ, Fisher LM. Preoperative predictors of postoperative implant performance in children. Ann Otol Rhinol Laryngol Suppl 2000;185:44–6.
74. Hodges AV, Ash MD, Balkany TJ, et al. Speech perception results in children with cochlear implants: contributing factors. Otolaryngol Head Neck Surg 1999;121:31–4.
75. Allen MC, Nikolopoulos TP, O'Donoghue GM. Speech intelligibility in children after cochlear implantation. Am J Otol 1998;19:742–6.
76. Moog JS, Geers A. Speech and language acquisition in young children after cochlear implantation. Otolaryngol Clin North Am 1999;32:1127–41.
77. Svirsky MA. Speech intelligibility of pediatric cochlear implant users and hearing aid users. In: Waltzman SB, Cohen NL, editors. Cochlear implants. New York: Thieme; 2000. p. 312–4.
78. Svirsky MA, Robbins AM, Kirk KI, et al. Language development in profoundly deaf children with cochlear implants. Psychol Sci 2000;11:153–8.
79. Pisoni DB, Svirsky MA, Kirk KI, et al. Looking at the "Stars": a first report on the intercorrelations among measures of speech perception, intelligibility and language development in pediatric cochlear implant users. Bloomington (IN): Indiana University; 1997.

CHAPTER 24

Facial Paralysis

Phillip A. Wackym, MD, John S. Rhee, MD

Facial nerve abnormalities represent a broad spectrum of lesions, including numerous congenital and acquired causes.[1] The patient who suffers with facial paralysis experiences not only functional consequences but also the psychological impact of a change in self-image and impaired communicative ability. In fact, a 1991 poll revealed that the level of discomfort that Americans felt on meeting those with facial abnormalities was second only to that associated with interacting with the mentally ill, and this discomfort far exceeded anxiety about encountering the senile, mentally retarded, deaf, blind, and those confined to a wheelchair.[2]

FACIAL NERVE ABNORMALITIES

Congenital

Möbius' Syndrome (Congenital Facial Diplegia) Möbius' syndrome is a rare congenital disorder, which usually includes bilateral seventh nerve paralysis and unilateral or bilateral sixth nerve paralysis. Since the disorder was described, many authors have studied families with the syndrome.[3] It is considered to have an autosomal dominant inheritance pattern with variable expressivity. The inheritance pattern is thought to be no higher than 1 in 50 in families in whom myopathies or other extremity anomalies such as clubfoot, arthrogryposis, or digital anomalies are not present.

The etiology of Möbius' syndrome is unclear. Neuropathologic studies have noted that the nuclei of cranial nerves (CNs) VI, VII, and XII are abnormal, with lesser abnormalities being found in the nuclei of CNs III and XI.[4] Other authors have reported that the facial nerves are smaller or absent at autopsy.[5] Pitner et al advanced yet another hypothesis based on their observation that normal facial nuclei were present on postmortem analysis; they suggested that there was a primary failure of facial muscle development.[6]

The clinical observation of congenital extraocular muscle paralysis and facial paralysis is the typical presentation of this disorder (Figure 24–1). No mass lesions will be found on magnetic resonance imaging (MRI). Ophthalmologic consultation and management are mandatory. Reinnervation procedures such as crossfacial grafts or hypoglossal-facial nerve anastomosis yield poor results, either owing to the paucity of motor end plates or the atrophic seventh nerves. Significant improvements of resting tone and voluntary animation can result from temporalis muscle transposition, which brings in a new neuromuscular system.

Hemifacial Microsomia The term hemifacial microsomia refers to patients with unilateral microtia, macrostomia, and mandibular hypoplasia. Goldenhar's syndrome (oculoauriculovertebral dysplasia) is considered to be a variant of this complex and is characterized by vertebral anomalies and epibulbar dermoids. Although approximately 25% of patients with hemifacial microsomia have facial nerve weakness, one patient with Goldenhar's syndrome has been reported to have aplasia of the facial nerve.[7]

Osteopetrosis Osteopetrosis is a generalized dysplasia of bone that may have an autosomal dominant or recessive inheritance pattern. The recessive form is more rapidly progressive and causes hepatosplenomegaly and severe neural atrophy secondary to bony overgrowth at neural foramina. Optic atrophy, facial paralysis, sensorineural hearing loss, and mental retardation are common in the recessive form, and death usually occurs by the second decade. However, in these severe cases of osteopetrosis, which were previously fatal, bone marrow transplantation has been reported to be of value.[8]

The dominant form causes progressive enlargements of the cranium and mandible and clubbing of

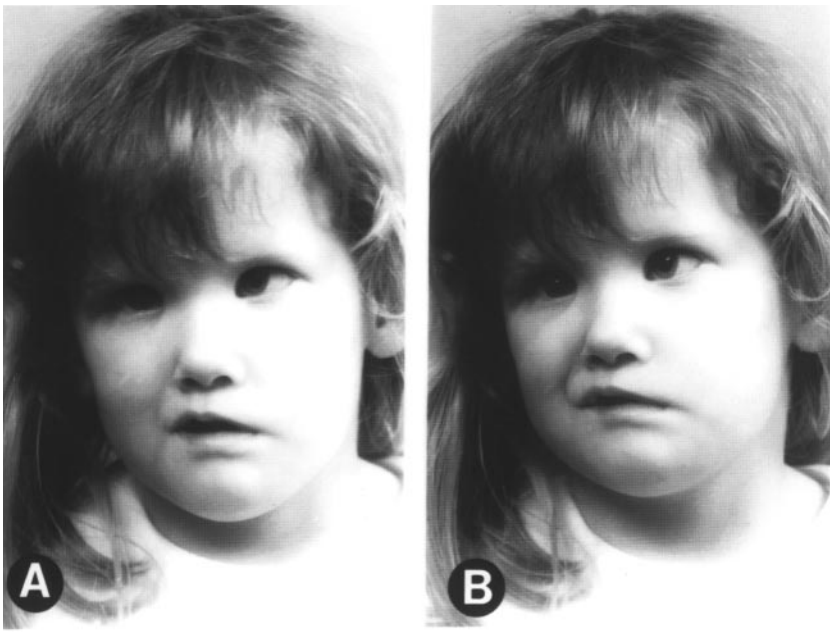


FIGURE 24–1. Möbius' syndrome. *A*, Left facial paralysis and right facial paresis photographed at rest. *B*, With smiling and rightward gaze, the right facial paresis and right sixth nerve paralysis are apparent. Reproduced with permission; copyright © 1992 P. A. Wackym.

the long bones. Increased bone density is seen radiographically. Progressive optic atrophy, trigeminal hypesthesia, recurrent facial paralysis, and sensorineural hearing loss are common. Complete decompression of the intratemporal facial nerve should be performed in patients with recurrent facial paralysis and radiographic evidence of osteopetrosis.

Acquired

Trauma Approximately 90% of all congenital peripheral facial nerve paralysis spontaneously improves, and most can be attributed to difficult deliveries, cephalopelvic disproportion, high forceps delivery, or intrauterine trauma. These types of congenital facial paralysis are often unilateral and partial, especially involving the lower division of the facial nerve. Since these causes involve extratemporal compression, surgical exploration or bony decompression is not indicated.[9]

Blunt trauma resulting in temporal bone fracture is best evaluated with high-resolution temporal bone computed tomographic (CT) scans (Figures 24–2 and 24–3). Temporal and parietal blows to the head may occur anywhere along a coronal arc, from the vertex to the cranial base. When the vector of force is directed toward the base, it classically passes toward the external auditory canal, deflects off the otic capsule, and extends anteromedially along the anterior edge of the petrous bone to the foramen lacerum and foramen ovale. The resulting fracture is described as a longitudinal temporal bone fracture. This is the most common type of temporal bone fracture (approximately 90%) and is also the most common type of fracture associated with facial nerve injury. The geniculate ganglion region of the facial nerve is most frequently injured. The indications for facial nerve decompression and exploration are the same as those discussed in detail under the Bell's palsy section of this chapter.

Frontal and particularly occipital blows to the head tend to result in transverse fractures of the temporal bone. More severe head injury is usually required to cause these fractures.[10] Since they often extend through the internal auditory canal or across the otic capsule, hearing loss and vertigo commonly result. Although only 10 to 20% of temporal bone fractures are transverse in orientation, they cause facial nerve injury in approximately 50% of patients. The anatomic region of the facial nerve most commonly injured is the labyrinthine segment.

Penetrating injuries to the extratemporal facial nerve should be explored urgently to facilitate identification of the transected distal branches using a facial nerve stimulator. If primary repair is not possible, the principles of facial nerve repair using cable grafts, described later in this chapter, should be followed. In infected wounds, urgent exploration and tagging of identified distal branches should precede

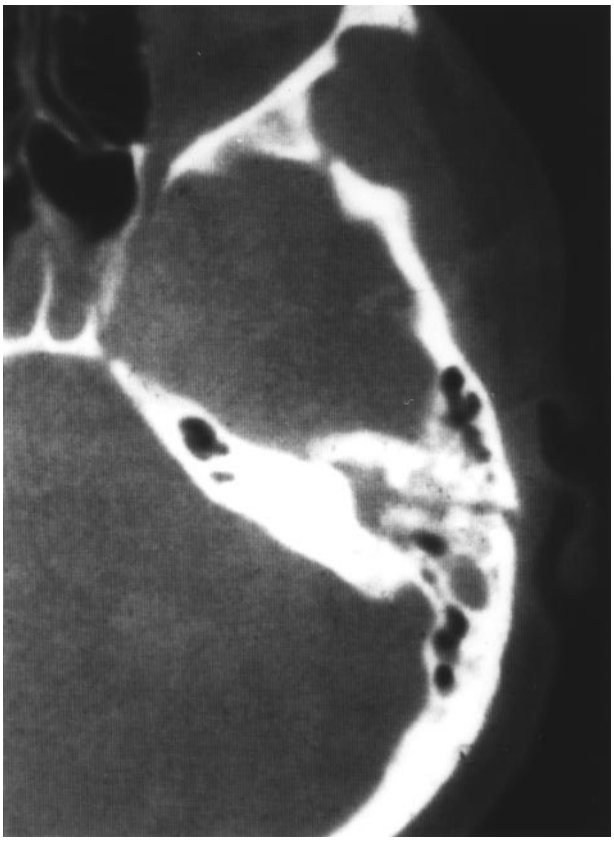


FIGURE 24–2. Axial computed tomographic scan with bone windows demonstrates a left longitudinal temporal bone fracture. Reproduced with permission; copyright © 1998 P. A. Wackym.

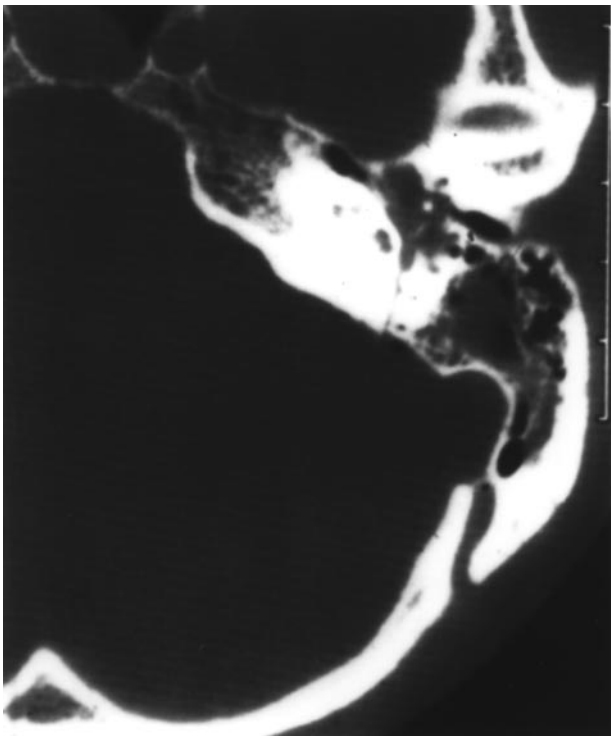


FIGURE 24–3. Axial computed tomographic scan with bone windows demonstrates a left transverse temporal bone fracture extending through the otic capsule. Reproduced with permission; copyright © 1998 P. A. Wackym.

control of the infection and granulation. Subsequent repair usually requires the use of cable grafts.

Facial nerve injury may occur with otologic surgery. The risk of injury of the facial nerve is particularly high in children with congenital ear malformations.[11,12] Additional groups at higher risk for injury to the facial nerve include infants who are undergoing mastoid surgery since the mastoid tip has not become pneumatized and the facial nerve exits the stylomastoid foramen laterally. In these young children, a semihorizontal, curvilinear skin incision should be used, and, as is the case with all otologic surgery, a facial nerve monitoring system should be used.

Infection *Bacterial.* Facial paralysis as a complication of otitis media has become rare in children owing to the ready access to medical care and antibiotics. Takahashi et al published their series of over 1,600 patients with facial paralysis and found that only 11 of these patients were younger than 20 years old and had facial paralysis owing to otitis media (0.67%).[13] They described the facial paralysis in this group of patients to have a slower progression and less complete paralysis than that seen in Bell's palsy. Temporal bone CT should be performed in all patients to eliminate the diagnosis of coalescent mastoiditis. Intravenous antibiotics in combination with myringotomy and tympanostomy tube placement remain our initial management algorithms for bacterial otitis media complicated by facial paralysis. Bacterial cultures should always be obtained at the time of myringotomy, and antibiotic selection should be tailored to the culture results.

Facial paralysis complicating mastoiditis or cholesteatoma is also rare. In Sheehy's series of over 180 children undergoing surgery for cholesteatoma, only 1 patient (0.5%) had facial nerve weakness.[14] The surgical management of these patients includes mastoidectomy, excision of the cholesteatoma, and appropriate antibiotic therapy.

Infection with the spirochete *Borrelia burgdorferi* (Lyme disease) can result in facial paralysis. This tick-borne infection is endemic to the northeastern United States and is named for the town of Lyme, Connecticut. Widespread infections have been reported from the west coast, Midwest, and east coast, as well as throughout Europe and Australia. As is the case with other spirochete infections, the clinical manifestations of Lyme disease are protean. Facial diplegia has been reported in Lyme disease[15] and should be considered in children presenting with facial paralysis. Serologic diagnosis should be followed by antibiotic therapy. Tetracycline is considered to be the agent of choice; however, erythromycin and penicillin have been successfully used.

Viral. Ramsay Hunt Syndrome (Herpes Zoster Oticus). James Ramsay Hunt (1872–1937), an American neurologist, published his seminal article associating the clinical syndrome that now bears his name with herpetic inflammation of the geniculate ganglion.[16] Critical to the development of this hypothesis was the publication of the pathologic studies of Head and Campbell in 1900, which advanced a new hypothesis regarding the etiology of herpes zoster.[17] This work inspired Hunt to first postulate that the etiology of herpes zoster oticus was recrudescence of herpes varicella-zoster virus (VZV) in the geniculate ganglion.[16] Clinically, Ramsay Hunt's syndrome may have a variety of manifestations. However, the original four classifications of the disease by Hunt included (1) disease affecting the sensory portion of the CN VII, (2) disease affecting the sensory and motor divisions of the CN VII, (3) disease affecting the sensory and motor divisions of the CN VII with auditory symptoms, and (4) disease affecting the sensory and motor divisions of the CN VII with both auditory and vestibular symptoms.

Herpes zoster oticus is the cause of 2 to 10% of all cases of facial paralysis, including 3 to 12% of adults and approximately 5% of children.[1,18–23] Patients may experience paresis or complete paralysis, with the poorest prognosis for recovery in the latter group. Approximately half of patients with Ramsay Hunt syndrome retain some facial motor disturbance; only a few maintain a complete paralysis.[18–21]

Based on the sensory distributions reflected by the VZV recrudescence observed by Hunt over his career,[16,24,25] the most common site of vesicular eruption is in the concha of the auricle (Figure 24–4). In addition, he described three other areas where vesicles can be found during herpes zoster oticus: a small strip of skin on the posteromesial surface of the auricle, the mucosa on the palate, and the anterior two-thirds of the tongue.[24] In his final publication, he detailed the sensory distributions of the facial nerve associated with the geniculate ganglion

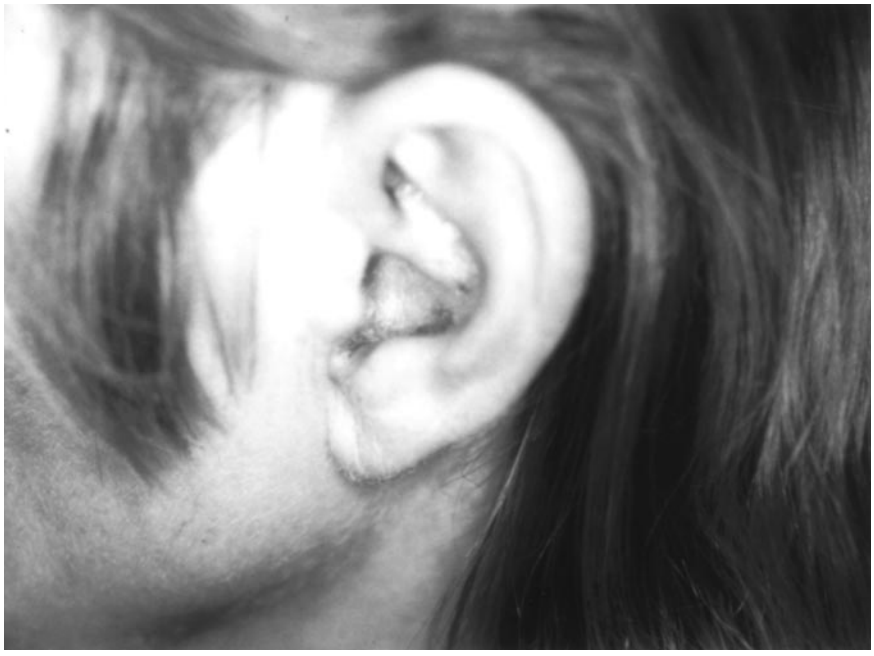


FIGURE 24–4. Vesicles in the concha of a patient with Ramsay Hunt syndrome (herpes zoster oticus) of the left facial nerve. Reproduced with permission; copyright © 2000 P. A. Wackym.

that subserve the axonal transport of the recrudescent VZV to form the vesicles visible during Ramsay Hunt syndrome.

With the advent of gadolinium diethylenetriamine pentaacetic acid (Gd-DTPA)-enhanced magnetic resonance imaging (MRI), acute imaging of the facial nerve is possible throughout the clinical course of Ramsay Hunt syndrome.[26] Despite histopathologic reports of diffuse inflammation along the entire intratemporal facial nerve in the disorder,[18,27–30] Korzec et al found that in three of the six patients with Ramsay Hunt syndrome who were studied with Gd-DTPA MRI, enhancement of the involved nerve was localized to the geniculate ganglion, labyrinthine segment, and the premeatal regions of the affected facial nerve, whereas there was no enhancement at all in the remaining three patients.[26] Likewise, in none of their six patients did they find enhancement of the vertical or tympanic segments of the facial nerve. These findings suggest that in the early stages of the clinical syndrome, the majority of the inflammation is found near the geniculate ganglion, whereas in the later stages, as examined postmortem, the inflammation has extended throughout the intratemporal facial nerve. However, the senior author has seen diffuse enhancement of the facial nerve on MRI examination of patients with Ramsay Hunt syndrome (Figure 24–5).

Blackley et al reviewed the histopathology associated with Ramsay Hunt syndrome in their one case and five others.[27] Aleksic et al added another histologic case.[31] Payten and Dawes reviewed the pathologic studies of herpes zoster oticus with facial paralysis for a 60-year interval.[32] Wackym et al used a molecular approach to demonstrate the presence of VZV deoxyribonucleic acid (DNA) combined with traditional histopathologic techniques to the study of the temporal bones from two patients with Ramsay Hunt syndrome.[30,33] The most consistent observation was diffuse inflammatory infiltration throughout the involved facial nerve; in addition, several investigators have shown lymphocytic infiltration of the geniculate ganglion, alteration of the geniculate ganglion somata, or both conditions.[18,27,29–31] Based on the models of VZV and herpes simplex virus (HSV) latency and reactivation,[34] these findings are expected. Unlike HSV, which remains dormant directly within sensory neuronal somata, VZV remains latent within the non-neuronal satellite cells surrounding each sensory neuronal somata. With recrudescence, the replication VZV virions are released from satellite cells into the extracellular matrix, where some are taken up by the sensory cell body and transported via the axons back to the skin or mucosa. This release of VZV into the extracellular matrix within the geniculate ganglion would cause an immune response that could result in inflammatory infiltrates throughout the intratemporal facial nerve. Therefore, the diffuse lymphocytic infiltration of the entire facial nerve remains consistent with Hunt's hypothesis.[29,30,32] Furuta et al reported the distribution of VZV DNA in 11 (79%) of 14 trigeminal ganglia and in 9 (69%) of 13 geniculate ganglia collected at autopsy of adults.[35] These data suggest that latent VZV in the geniculate ganglion is a common biologic phenomenon; however, the small sample size and the lack of clinical history

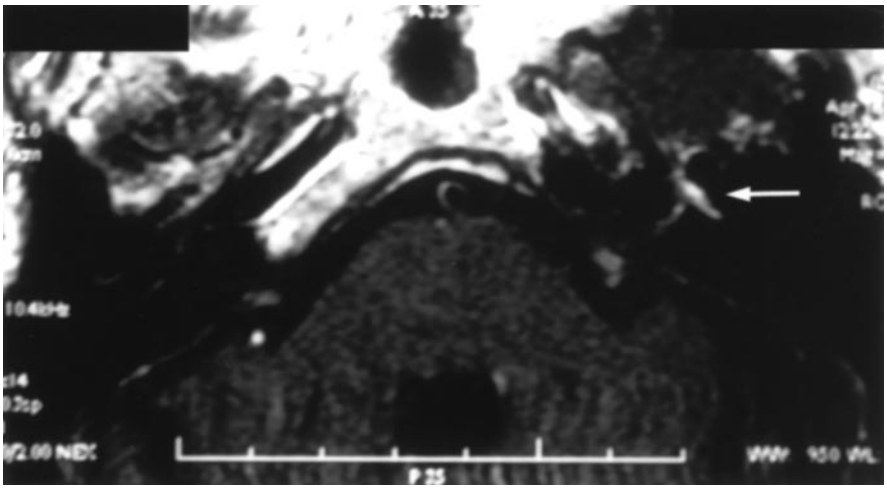


FIGURE 24–5. Magnetic resonance image (MRI) of a facial nerve in Ramsay Hunt syndrome (herpes zoster oticus). Gadolinium-enhanced axial MRI shows inflammation of the facial nerve within the internal auditory canal, labyrinthine segment, geniculate ganglion, tympanic segment, and greater superficial petrosal nerve (*arrow*). Reproduced with permission; copyright © 2000 P. A. Wackym.

regarding whether the patients experienced herpes zoster oticus during their lifetime necessitate confirmation of these observations with a much larger sample size of well-characterized patients.

Other authors have suggested that Ramsay Hunt syndrome represents a cranial polyneuropathy.[28,36,37] Care must be taken in interpreting the histopathologic findings in each case as some may represent cephalic zoster with neuritis of multiple cranial nerves.[28,37] Severe central neurologic deficits or multiple cranial motor neuropathies are well recognized in cephalic zoster.[38,39]

Surgical management with decompression of the facial nerve in Ramsay Hunt syndrome has been advocated by some authors (reviewed by Crabtree).[19] Although the advent of facial electroneuronography (ENoG) has resulted in a better idea about which patients with facial paralysis have severe facial nerve injuries requiring surgical decompression,[40] the diffuse inflammatory edema common in Ramsay Hunt syndrome has led most clinicians to avoid completion of facial nerve decompression.

Experience with the use of intravenous acyclovir (Zovirax) during the acute presentation of Ramsay Hunt syndrome suggests that antiviral medications may facilitate recovery and minimize the morbidity associated with facial paralysis.[23,41–43] This intravenous route has more inherent expenses than an oral route of administration. However, tissue levels of acyclovir delived by an oral route are not high enough to treat varicella-zoster infections. Alternate antiviral agents such as valacyclovir (Valtrex) (1 g orally three times a day for 10 to 14 days) or famciclovir (Famvir) (500 mg orally three times a day for 10 days), which achieve adequate levels by an oral route, are now available as an alternative to intravenous acyclovir for the treatment of patients with Ramsay Hunt syndrome.[44] Because the oral route is much more cost effective, this route is preferred. Likewise, oral corticosteroids have been advocated in patients with Ramsay Hunt syndrome.[45]

BELL'S PALSY. Bell's palsy is responsible for 60 to 75% of all cases of facial paralysis. In the past, Bell's palsy was defined as an "idiopathic facial paralysis" or as a mononeuropathy of undetermined origin. Recent observations have linked the cause to HSV 1.[46,47] Burgess et al, using polymerase chain reaction amplification, identified HSV 1 DNA in paraffin-embedded sections of the geniculate ganglion from a patient who died 6 days after the onset of Bell's palsy.[46] Furuta et al found HSV DNA in 71% of geniculate ganglia and in 81% of trigeminal ganglia in eight random autopsy specimens taken from adult cadavers.[48] These data suggest that latent HSV in the geniculate ganglion may be a common biologic phenomenon; however, the small sample size and the lack of clinical history regarding whether the patients experienced Bell's palsy during their lifetime necessitate confirmation of these observations with a much larger sample size of well-characterized patients. However, genetic, immunologic, vascular, entrapment, and other infectious causes have all been advanced in the etiology of Bell's palsy.[49]

Bell's palsy is an acute, unilateral paresis or paralysis of the facial nerve in a pattern consistent with peripheral nerve dysfunction (Figure 24–6). The onset and evolution are typically rapid, less than 48 hours, and the onset of paralysis may be preceded by a viral prodrome. The symptoms during the early phase of facial paralysis include facial numbness, epiphora, pain, dysgeusia, hyperacusis (dysacusis), and decreased tearing. The pain is usually retroauricular and sometimes radiates to the face, pharynx, or shoulder. Physical findings of this subtle polyneuritis include hypesthesia or dysesthesia of the CNs V and IX and of the second cervical nerve.[50] Motor paralysis of branches of the CN X is seen as a unilateral shift of the palate or vocal cord paresis/paralysis.

Recurrence of Bell's palsy occurs in 7.1[51] to 12%[1] of patients. In the series of 140 patients with recurrent Bell's palsy reported by Pitts et al, ipsilateral recurrences were as common as development of contralateral Bell's palsy.[51] Also of note in this series was the observation that the incidence of diabetes mellitus was 2.5-fold greater than nonrecurrent cases.

Gadolinium-enhanced MRI has been advocated as a diagnostic tool in assessing Bell's palsy. Gadolinium enhancement of the normal facial nerve does not occur. Therefore, enhancement of this structure would be owing to increased extracellular fluid from edema, inflammation, or neoplasm. Our observations with gadolinium-enhanced MRI in Bell's palsy, as well as those of others,[26] are supportive of Fisch's hypothesis of axoplasmic damming at the meatal segment with subsequent edema and nerve conduction impairment (Figure 24–7).[52] However, one study demonstrated that there was no

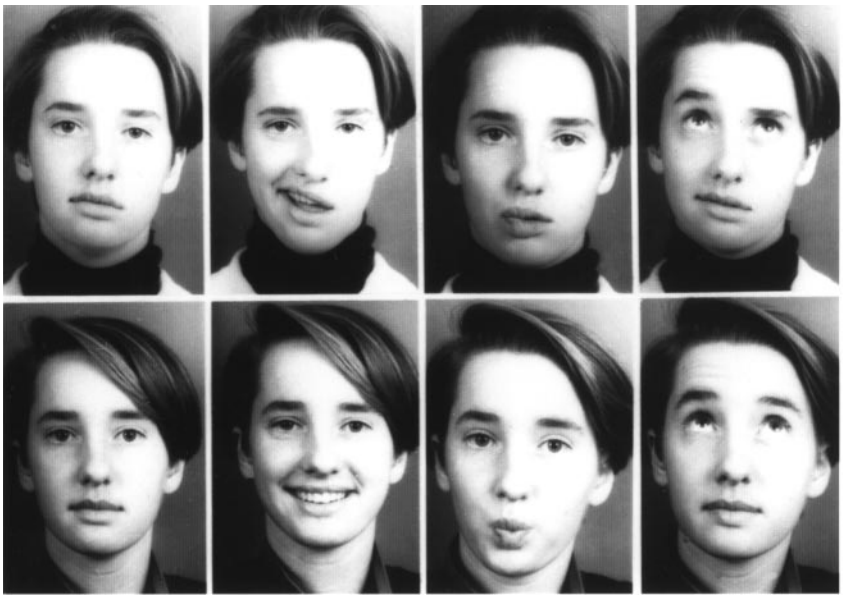


FIGURE 24–6. Preoperative facial photographs of a child with complete facial paralysis owing to Bell's palsy, *top row*. House-Brackmann grade I result 2 months after middle cranial fossa decompression of the labyrinthine segment and geniculate ganglion of the facial nerve, *bottom row*. Reproduced with permission; copyright © 1992 P. A. Wackym.

prognostic significance of gadolinium enhancement of the facial nerve on MRI in patients with Bell's palsy.[26] Therefore, gadolinium-enhanced MRI is not indicated in every patient with facial paralysis. In patients suspected of having a tumor from clinical or electrodiagnostic data, gadolinium-enhanced MRI, along with high-resolution CT of the internal auditory canal (IAC), fallopian canal, skull base, and parotid, should be performed.

Adour and colleagues examined the outcome of treating patients with Bell's palsy with both acyclovir and corticosteroids in a prospective, randomized, double-blind trial.[53] Half of the patients received a 10-day course of oral acyclovir (400 mg five times per day) and prednisone. The control group received prednisone and a placebo. All study patients began treatment within 3 days of the onset of facial paralysis. The placebo-prednisone group had lower facial function recovery scores (House-Brackmann grade III or IV in 23%) and was almost three times as likely to have an unsatisfactory result as the acyclovir-prednisone group (only 7% of the acyclovir-prednisone group had a House-Brackmann grade III or IV).

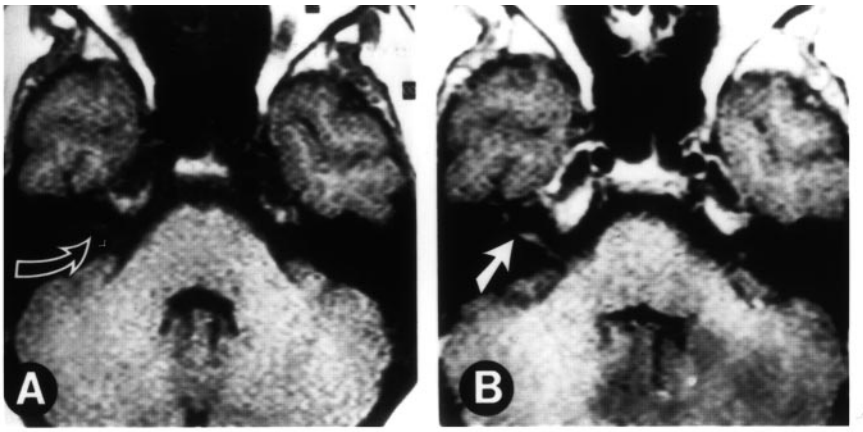


FIGURE 24–7. Magnetic resonance image (MRI) of facial nerve in Bell's palsy. *A*, Non–gadolinium-enhanced axial MRI shows normal-appearing seventh and eighth nerves within the right internal auditory canal (*open arrow*). *B*, Gadolinium-enhanced axial MRI shows axoplasmic damming of the right facial nerve at the meatal foramen (*solid arrow*). This MRI finding was confirmed at surgical decompression. Reproduced with permission; copyright © 1992 P. A. Wackym.

Electroneuronography provides a quantitative assessment of facial nerve function and allows a relative comparison between the normal and affected sides and will be discussed in more detail later in this chapter. Our criteria for surgical decompression include ENoG degeneration greater than 90% relative to the unaffected side, no voluntary facial nerve electromyographic (EMG) activity on the affected side, and the operation within 14 days of onset.[40,54,55] Decompression is limited to the meatal and labyrinthine segments through a middle cranial fossa approach.[55–57] Fisch and Esslen, in 1972, were the first to propose that the most likely site for neural compression and conduction block in Bell's palsy was at the entrance to the meatal foramen, the narrowest bony point through which the facial nerve passes.[58] Interestingly, intraoperative evoked EMG documented the conduction block at this area in 94% of decompressed cases, and marked swelling proximal to this point was the typical observation.[58,59] Gantz et al published a prospective study in which a well-defined surgical decompression of the facial nerve was performed in a population of patients with Bell's palsy who exhibited the electrophysiologic features associated with poor outcomes (ENoG degeneration greater than 90% relative to the unaffected side, no voluntary facial nerve EMG activity on the affected side).[55] Subjects who did not reach 90% degeneration on ENoG within 14 days of paralysis all returned to House-Brackmann grade I (n = 48) or II (n = 6) at 7 months after onset of the paralysis. Control subjects self-selecting not to undergo surgical decompression when > 90% degeneration on ENoG and no motor unit potentials on EMG were identified had a 58% chance of having a poor outcome at 7 months after onset of paralysis (House-Brackmann grade III or IV [n = 19]). A group with similar ENoG and EMG findings undergoing middle fossa facial nerve decompression exhibited House-Brackmann grade I (n = 14) or II (n = 17) in 91% of the cases. It is recommended that these criteria and this surgical algorithm be followed.

Other Viral Infections. Other viral infections such as primary chickenpox, mononucleosis, mumps, and poliomyelitis can result in facial paralysis that may or may not resolve spontaneously. For these specific viral infections, immunization, when available, is the most effective preventive measure, and supportive care is required during the active infection. Facial reanimation procedures are sometimes required after adequate follow-up suggests that spontaneous recovery will not occur.

Benign or Malignant Neoplasms Tumor involvement of the facial nerve should be considered in facial paralysis if one or more of the following clinical features are present: facial paralysis that progresses slowly over 3 weeks, recurrent ipsilateral facial paralysis, facial weakness associated with muscle twitching, long-standing facial paralysis (greater than 6 months), facial paralysis associated with other CN deficits, or evidence of malignancy elsewhere in the body.

Several benign and malignant tumors can involve the facial nerve along its intracranial, intratemporal, or extracranial course (Table 24–1). Schwannoma is the most common primary tumor of the facial nerve. It is benign and usually involves the labyrinthine, tympanic, and mastoid segments of the facial nerve. Nerve resection and interpositional nerve grafting may initially be necessary for restoration of continuity[60,61]; however, decompression will often give patients many years of facial nerve function before resection and grafting must be completed.

The use of radiographic imaging is indicated if the characteristics of the facial paralysis are suggestive of a neoplasm. Radiographic studies should include visualization of the entire course of the facial nerve, from the brainstem to the facial musculature. Gadolinium-enhanced MRI (Figure 24–8) is extremely useful in imaging solid tumors involving the facial nerve, and high-resolution CT scans are useful in identifying bony erosion of the fallopian canal.

Tumors may arise in the vicinity of the facial nerve and cause facial weakness either by compression or direct invasion. When the tumor is benign, the continuity of the facial nerve should be preserved at all costs by sharp dissection and mobilization techniques. This is appropriate management, whether the nerve is compromised in the IAC by an acoustic neuroma or in the parotid gland by a pleomorphic adenoma. A malignant process with direct invasion of the nerve usually mandates resection of the involved portion of the nerve with immediate interpositional nerve grafting. If the management of the disorder involves chemotherapy or radiation therapy rather than surgical intervention, facial

TABLE 24–1. Causes of Facial Paralysis

Birth and Congenital
- Aplasia of the facial nerve
- Molding
- Forceps delivery
- Myotonic dystrophy
- Möbius' syndrome

Trauma
- Barotrauma
- Brainstem injuries
- Cortical injuries
- Facial (soft tissue injuries)
- Penetrating injury to middle ear
- Temporal bone fractures

Neurogenic
- Millard-Gubler syndrome (abducens palsy with contralateral hemiplegia owing to lesion in base of pons involving corticospinal tract)
- Opercular syndrome (cortical lesion in facial motor area)

Infection
- Bell's palsy
- Acute or chronic otitis media
- Cholesteatoma, acquired and congenital
- Herpes zoster oticus (Ramsay Hunt syndrome)
- Mastoiditis
- Malignant otitis externa
- Meningitis
- Parotitis
- Chickenpox
- Encephalitis
- Poliomyelitis
- Mumps
- Mononucleosis
- Leprosy
- HIV (human immunodeficiency virus)
- Influenza
- Coxsackie virus
- Malaria
- Syphilis
- Tuberculosis
- Botulism
- Mucormycosis
- Lyme disease

Vascular
- Endovascular embolization (external carotid artery branches)
- Intratemporal aneurysm of internal carotid artery
- Anomalous sigmoid sinus

Neoplastic
- Facial neuroma
- Acoustic neuroma (rapid expansion or post resection)
- Paraganglioma (glomus jugulare)
- Leukemia
- Meningioma
- Hemangiopericytoma
- Hemangioma
- Hemangioblastoma
- Pontine glioma
- Sarcoma
- Hydradenoma (external auditory canal)
- Teratoma
- Fibrous dysplasia
- von Recklinghausen's disease
- Carcinomatous encephalitis
- Cholesterol granuloma
- Carcinoma (invasive or metastatic)

Genetic and Metabolic
- Diabetes mellitus
- Hyperthyroidism
- Pregnancy
- Hypertension
- Alcoholic neuropathy
- Sickle cell disease
- Bulbopontine paralysis
- Oculopharyngeal muscular dystrophy
- Camurati-Engelmann disease (hereditary diaphyseal dysplasia)

Toxic
- Thalidomide (cranial nerves VI and VII with atretic external ears [Miehlke syndrome])
- Tetanus
- Diphtheria
- Carbon monoxide
- Lead intoxication

Iatrogenic
- Anesthesia, local (mandibular block, face, mastoid)
- Tetanus vaccination
- Vaccine treatment for rabies exposure
- Otologic and neurotologic skull base surgery, parotid surgery
- Interventional neuroradiology (embolization)

Continued

TABLE 24–1. Causes of Facial Paralysis—Continued

Idiopathic	Sarcoidosis
Bell's palsy	Wegener's granulomatosis
Melkersson-Rosenthal syndrome	Eosinophilic granuloma
Hereditary hypertrophic neuropathy	Histiocytosis X
Autoimmune syndromes	Amyloidosis
Thrombotic thrombocytopenia purpura	Paget's disease
Guillain-Barré syndrome	Osteopetrosis
Multiple sclerosis	Kawasaki's disease
Myasthenia gravis	

reanimation procedures may be indicated if there is persistent facial nerve dysfunction.

Hemifacial Spasm Hemifacial spasm is typically a disorder of the fourth and fifth decades of life and occurs twice as often in women as it does in men. Electrophysiologic and surgical observations indicate that the facial nerve hyperactivity in hemifacial spasm is caused by vascular compression of the facial nerve.[62] Trigeminal neuralgia, hemifacial spasm, glossopharyngeal neuralgia, tinnitus, and disabling positional vertigo have all been associated with vascular compression.[63] Microvascular decompression operations involve separating the compressive vessel from its point of contact with the CN root entry or exit zone and interposition of a prosthesis (usually Teflon felt) to prevent further nerve compression.[62] Drawing from the experience with the more common microvascular compression syndrome of the trigeminal nerve, large series have reported that 62 to 64% of trigeminal neuralgia patients have a compressive artery, 12 to 24% have a compressive vein, 13 to 14% have both an artery and a vein, and 8% have either a tumor or vascular malformation pressing on the trigeminal nerve.[64] The initial failure rate for microvascular decompression for trigeminal neuralgia is 2 to 7%, with a 3.5% per year incidence of a major recurrence.[65–67] In the series of 1,185 patients who underwent microvascular decompression for trigeminal neuralgia reported by Barker et al in 1996, there was a mean follow-up of 6.2 years, and 30% of patients experienced a major recurrence, with 11% requiring a second microvascular decompression.[68] The absence of a clear site of arterial compression has been associated with high recurrence rates.[68,69] Patients found to have only venous compression and no arterial compression are more likely to suffer a recurrence.[68] Kureshi and Wilkins reported their surgical experience with 31 posterior fossa re-explorations for recurrent or persistent trigeminal neuralgia and hemifacial spasm.[70] They discovered 3 (10%) cases in which there was new or previously unrealized arterial compression of neural structures. Similarly, a series of

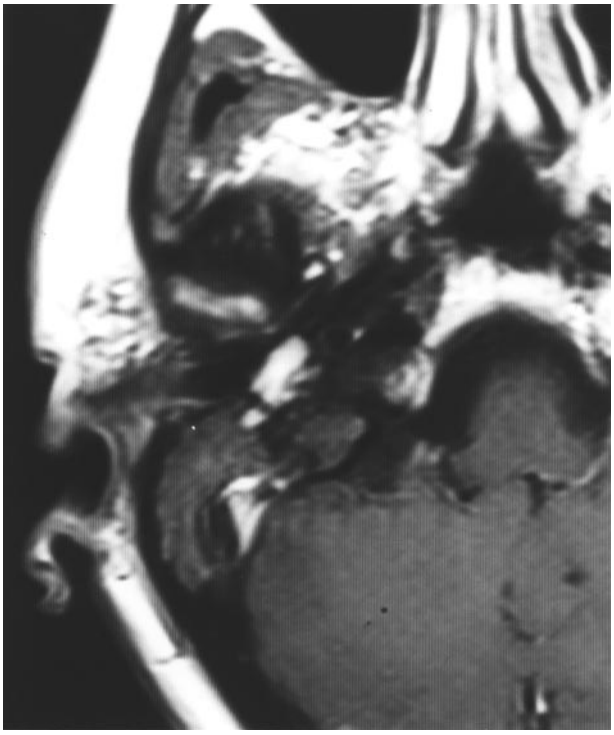


FIGURE 24–8. Gadolinium-enhanced axial magnetic resonance image shows an extensive facial neuroma with extension into the cochlea and internal auditory canal. Reproduced with permission; copyright © 2000 P. A. Wackym.

116 patients with microvascular compression syndrome reoperated on after failure reported the identification of previously unseen arterial compression in 65.5%.[71] Liao et al discovered persistent vascular compression in three of five patients undergoing repeat microvascular decompression.[72]

Although some have doubted that microvascular compression of a CN can represent the cause of hemifacial spasm, the literature indicates that the identification of an arterial vessel compressing the facial nerve in hemifacial spasm and subsequent decompression results in higher cure rates and decreased recurrences than when a specific artery is not identified. We believe that some microvascular decompression procedures fail because the offending vessel has not been identified at the primary operation.[62] Some of these failures occur because the microscope provides incomplete information about the anatomic relationship between the nerves and vessels. The zero-degree endoscope provides a panoramic view of the cerebellopontine angle, and with angled endoscopes, provision for "looking around corners" is made. Magnan et al used an endoscope in 60 patients with hemifacial spasm and demonstrated that the operating microscope was able to visualize the offending vessel in 28% of patients, whereas the endoscope was effective in 93% of the same patients.[73] We have advocated the adjunctive use of the endoscope in microvascular decompression surgery and believe that this will improve surgical outcomes.[62]

Miscellaneous Disorders The onset of simultaneous bilateral facial paralysis suggests Guillain-Barré syndrome, sarcoidosis, sickle cell disease, or some other systemic disorder. Guillain-Barré syndrome is a relatively common neurologic disorder and is an acute inflammatory polyradiculoneuropathy that progresses to varying degrees of paralysis. The etiology remains unknown; however, autoimmune or viral mechanisms have been considered. Classic histopathologic features of the syndrome include a lymphocytic cellular infiltration of peripheral nerves and destruction of myelin. The facial paralysis is typically bilateral and often resolves spontaneously after a prolonged course of paralysis. Although there is no role for surgical decompression of the facial nerve in this disorder, reanimation is only considered late in the course of the disease.

Melkersson-Rosenthal Syndrome Melkersson-Rosenthal syndrome is a neuromucocutaneous disease with a classic triad of recurrent facial (labial) edema and recurrent facial paralysis associated with a fissured tongue. Patients with Melkersson-Rosenthal syndrome may not present with the complete triad, and although facial paralysis is the most commonly recognized neurologic symptom, it is not mandatory for the diagnosis. Headache, granular cheilitis, trigeminal neuralgiform attacks, dysphagia, laryngospasm, and a variety of CN and cervical autonomic dysfunctions may also occur. The patient with Melkersson-Rosenthal syndrome may present at any age and with any variety of classic and associated features, which may wax and wane. Approximately one-third of the patients have recurrent facial paralysis as part of their syndrome. The underlying etiologic factor has been thought to be a neurotropic edema causing compression and paralysis of the facial nerve as it passes through the fallopian canal. Since the anatomically most constricted area of the fallopian canal is the meatal foramen and because most prior reports observed recurrence after transmastoid decompression, Graham and Kemink elected to decompress the proximal segment in addition to the mastoid segment of the facial nerve in all such cases by performing a combined transmastoid and middle cranial fossa facial nerve decompression and neurolysis of the nerve sheath.[74] The preliminary data presented by Graham and Kemink suggest that edematous involvement of the facial nerve in recurrent facial paralysis does occur intratemporally and that the recurrent paralysis can be prevented by transmastoid and middle cranial fossa total facial nerve decompression with neurolysis of the facial nerve sheath.[74] Recurrent paralysis over a prolonged period of time usually results in increasing residual dysfunction. If evidence of residual paresis exists, facial nerve decompression of the labyrinthine segment and geniculate ganglion through a middle cranial fossa exposure is recommended at the time of the next episode of paralysis.

SURGICAL ANATOMY OF THE FACIAL NERVE

Detailed knowledge and familiarity with the complex course of the facial nerve and its anatomic relationship to other vital structures are essential to the surgeon who plans to operate in this area. The facial

nerve (CN VII) exits the brainstem at the pontomedullary junction approximately 1.5 mm anterior to the vestibulocochlear nerve (CN VIII). The facial nerve is smaller in diameter (approximately 1.8 mm) than the oval CN VIII (approximately 3 mm in the largest diameter). A third smaller nerve, the nervus intermedius, emerges between CN VII and CN VIII and eventually becomes incorporated within the sheath of CN VII. After leaving the brainstem, CN VII follows a rostrolateral course through the cerebellopontine cistern for 15 to 17 mm, entering the porus of the IAC of the temporal bone (Figure 24–9). Other important structures in the cerebellopontine cistern include the anterior inferior cerebellar artery (AICA) and the veins of the middle cerebellar peduncle. The AICA passes near or between CN VII and CN VIII; the veins are more variable in position and number. On entering the IAC, the facial nerve occupies the anterosuperior quadrant of this channel for 8 to 10 mm. Then it enters the fallopian canal at the fundus of the IAC. The IAC is anterior to the plane of the superior semicircular canal (SSC). Superiorly, the bone overlying the IAC is within a 60-degree angle, whose vertex is the SSC ampulla. At the entrance of the fallopian canal (meatal foramen), CN VII narrows to its smallest diameter, 0.61 to 0.68 mm.[75,76] Only the pia and arachnoid membranes form a sheath around the nerve at this point since the dural investment terminates at the fundus of the IAC. Many authors believe that the small diameter of the meatal foramen is an important factor contributing to the etiology of facial paralysis in certain diseases such as Bell's palsy and Ramsay Hunt syndrome.[52,59,75,77]

The intratemporal course of the facial nerve has three distinct anatomic regions: the labyrinthine, tympanic, and mastoid segments. The labyrinthine segment is shortest (approximately 4 mm), extending from the meatal foramen to the geniculate ganglion. This segment travels anterior, superior, and lateral, forming an anteromedial angle of 120 degrees with the IAC portion. The basal turn of the cochlea is closely related to the fallopian canal and lays anteroinferior to the labyrinthine segment of the facial nerve. At the lateral end of the labyrinthine segment, the geniculate ganglion is found, and the nerve makes an abrupt posterior change in direction, forming an acute angle of approximately 75 degrees. Anterior to the geniculate ganglion, the greater superficial petrosal nerve exits the temporal bone through the hiatus of the facial canal. The hiatus of the facial canal is quite variable in its distance from the geniculate ganglion. The hiatus of the facial canal also contains the vascular supply to the geniculate ganglion region. The tympanic, or horizontal, segment of the nerve is approximately 11 mm long, running between the lateral semicircular canal superiorly and the stapes inferiorly, forming the superior margin of the fossa ovale. Between the tympanic and mastoid segments, the nerve gently curves inferiorly for about 2 to 3 mm. The mastoid, or vertical, segment is the longest intratemporal portion of the nerve, measuring approximately 13 mm. As the nerve exits the stylomastoid foramen at the anterior margin of the

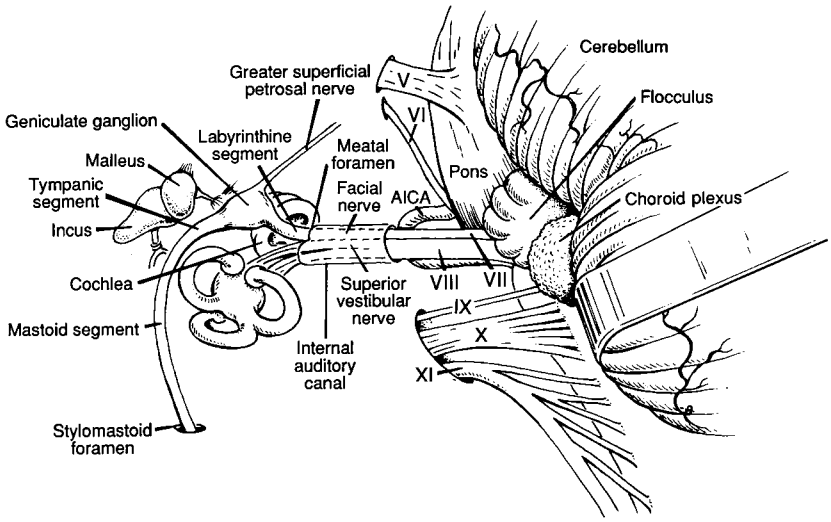


FIGURE 24–9. Course and relationships of the left facial nerve from the pontomedullary junction to the stylomastoid foramen. AICA = anterior inferior cerebellar artery. Reproduced with permission ©2001 P. A. Wackym.

digastric groove, an adherent fibrous sheath of dense vascularized connective tissue surrounds it. The stylomastoid artery and veins are within this dense sheath.

NERVE INJURY AND NEURODIAGNOSTIC TESTS

Facial Nerve Histology

A connective tissue layer, the endoneurium, surrounds each myelinated axon (Figure 24–10). This layer is closely adherent to the Schwann cell layer of each axon. The importance of the endoneurial sheath, in the context of nerve injury and repair, is that it provides a continuous tube through which a regenerating axon can grow. The second layer of the nerve sheath is the perineurium. The perineurium provides tensile strength to the nerve. The perineurium is also the primary barrier to the spread of infection. The outermost layer is the epineurium. This layer contains the vasa nervorum, providing the blood supply as well as the lymphatic vessels.

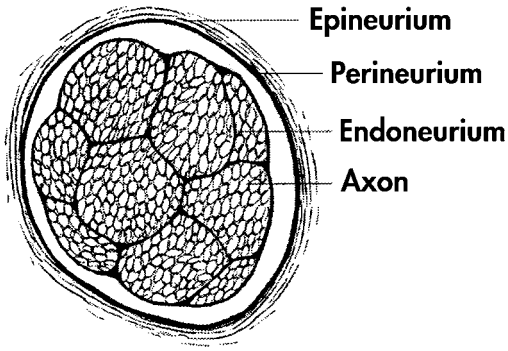


FIGURE 24–10. Schematic illustration of a cross-section of the facial nerve demonstrating the relationships among endoneurium, perineurium, and epineurium. Reproduced with permission; copyright © 2001 P. A. Wackym.

Facial Nerve Injury

It is necessary to review the types of nerve damage to understand better electrodiagnostic testing of the facial nerve, prognosis for recovery, and the development of synkinesis, as well as the rationale for facial nerve decompression. Table 24–2 summarizes the Sunderland and Seddon classifications of nerve injuries.[78,79] A facial nerve grading system has been established by the American Academy of Otolaryngology-Head and Neck Surgery (Table 24–3).[80] The House-Brackmann scoring system was developed as a means of reporting facial recovery after facial nerve paralysis caused by idiopathic facial nerve paralysis or recovery after removal of an acoustic neuroma when an injury to the facial nerve had been sustained.[80] Although this scale is useful when measuring facial nerve function in patients who have suffered an incomplete nerve injury (Sunderland degree I to IV [see Table 24–2]), it is inappropriate for assessing patients who have suffered Sunderland degree V injury and have undergone facial nerve repair or grafting. Gidley et al identified three reasons why it is difficult to apply the House-Brackmann grading system to a repaired facial nerve: (1) all repairs result in mass movement, (2) most patients regain complete eye closure and competent oral sphincter function, and (3) almost none have the ability to raise their forehead.[81] Consequently, they have proposed a grading system that more accurately grades the outcome from facial nerve grafting or repair procedure (Table 24–4). A letter

TABLE 24–2. Sunderland and Seddon Classifications of Nerve Injuries

Pathology	*Sunderland*[78]	*Seddon*[79]
Conduction block, damming of axoplasm	First degree	Neurapraxia
Transection of the axon with intact endoneurium	Second degree	Axonotmesis
Transection of nerve fiber (axon and endoneurium) inside intact perineurium	Third degree	Neurotmesis
Above plus disruption of perineurium (epineurium remains intact)	Fourth degree	Neurotmesis
	Fifth degree	Neurotmesis

TABLE 24–3. House-Brackmann Facial Nerve Grading System

Grade	*Description*	*Characteristics*
I	Normal	Normal facial function in all areas
II	Mild dysfunction	*Gross:* slight weakness noticeable on close inspection; may have very slight synkinesis
		At rest: normal symmetry and tone
		Motion: forehead—moderate to good function; eye—complete closure with minimal effort; mouth—slight asymmetry
III	Moderate dysfunction	*Gross:* obvious but not disfiguring difference between the two sides; noticeable but not severe synkinesis; contracture and/or hemifacial spasm
		At rest: normal symmetry and tone
		Motion: forehead—slight to moderate movement; eye—complete closure with effort; mouth—slightly weak with maximal effort
IV	Moderately severe dysfunction	*Gross:* obvious weakness and/or disfiguring asymmetry
		At rest: normal symmetry and tone
		Motion: forehead—none; eye—incomplete closure; mouth—asymmetric with maximum effort
V	Severe dysfunction	*Gross:* only barely perceptible motion
		At rest: asymmetry
		Motion: forehead—none; eye—incomplete closure; mouth—slight movement
VI	Total paralysis	No movement

Adapted from House JW and Brackmann DE.[80]

grading system was chosen to avoid confusion with the House-Brackmann classification.

Electroneuronography and Electromyography
The two most useful objective electrodiagnostic tests of facial nerve function are ENoG and EMG.

Facial Electroneuronography. Electroneuronography uses supramaximal electrical stimulation of the facial nerve at the level of the stylomastoid foramen to produce a compound muscle action potential. This evoked electromyogenic response is recorded with surface electrodes placed over the perioral (nasolabial) muscles since a large representative population of facial nerve fibers would be sampled by recording the evoked response from this group of muscles. Needle electrodes are not used because intramuscular needle electrodes would not sample a sufficient number of motor units to yield the representative maximal amplitude. A supramaximal bipolar stimulation (galvanic) is provided to saturate the nerve and produce a complete and synchronous depolarization. The galvanic stimulation is typically delivered as rectangular pulses, with a duration of 200 μs and an interpulse interval of 1 s. The amplitude of the evoked response is plotted as a function of time after stimulation. Both the normal and affected sides are tested, and the amplitude of the responses is compared. The percentage of degenerated fibers is calculated arithmetically, as follows:

$$\text{Percentage of degenerated fibers} = 100 - \left(\frac{\text{Amplitude of evoked response [in } \mu\text{V] Affected side}}{\text{Amplitude of evoked response [in } \mu\text{V] Normal side}} \times 100 \right)$$

TABLE 24–4. Repaired Facial Nerve Recovery Scale

Score	*Function*
A	Normal facial function
B	Independent movement of eyelids and mouth, slight mass motion, slight movement of forehead
C	Strong closure of eyelids and oral sphincter, some mass motion, no forehead movement
D	Incomplete closure of eyelids, significant mass motion, good tone
E	Minimal movement in any branch, poor tone
F	No movement

Adapted from Gidley PW et al.[81]

This electrodiagnostic test depends on the physiologic premise of neural injury proposed by both Seddon and Sunderland (see Table 24–2). Injuries that are limited to producing a conduction block within the nerve (neurapraxia) do not disrupt axoplasmic continuity and will continue to conduct a neural discharge if the electrical stimulus is presented distal to the conduction block. With more severe injuries, axoplasmic disruption (axonotmesis) or neural tubule disruption (neurotmesis) will result in wallerian degeneration distal to the site of injury. Nerve fibers that undergo wallerian degeneration cannot propagate electrically evoked potentials distal to the injury. Axonotmesis, in contrast to neurotmesis, has a better prognostic outcome. With resolution of the neural injury, in a nerve that has undergone axonotmesis, the axon will regenerate through the intact neural tubule, potentially allowing complete return of motor function to the muscle fiber innervated by that nerve fiber. The more severely disrupted neural tubule injury of neurotmesis has the potential to regenerate in an unsuccessful manner and can thereby result in misdirection of fibers, clinically causing synkinesis and incomplete return of motor function. Electroneuronography can be used to differentiate nerve fibers that have minor conduction blocks (neurapraxia) from those that have undergone wallerian degeneration; however, ENoG cannot differentiate the type of wallerian degeneration (axonotmesis versus neurotmesis). The severity of the injury can be inferred from the rate of degeneration after injury. More rapid wallerian degeneration is associated with neurotmesis, whereas nerves that degenerate more slowly are more likely to exhibit axonotmesis.[40,54,55]

The timing for performing ENoG should take into consideration the time course of wallerian degeneration. With a known complete transection of the facial nerve (eg, traumatic injury), 100% wallerian degeneration occurs over 3 to 5 days as the distal axon slowly degenerates. Therefore, early testing, within 3 days of paralysis, may not be representative of the degree of injury, and as outlined above, the time course of degeneration may reflect the degree of injury. An important technical detail to be attentive to is the need to stimulate the nerve at the stylomastoid foramen 10 to 20 times before making an amplitude measurement. The initial stimulation will improve the synchronization within the nerve and therefore improve the reliability of the test.

As discussed earlier in this chapter, our criteria for surgical decompression include ENoG degeneration greater than 90% relative to the unaffected side, no voluntary facial nerve EMG activity on the affected side, and the operation within 14 days of onset.[40,54,55] Decompression is limited to the meatal and labyrinthine segments through a middle cranial fossa approach.[55–57] Gantz et al published a prospective study in which a well-defined surgical decompression of the facial nerve was performed in a population of patients with Bell's palsy who exhibited the electrophysiologic features associated with poor outcomes (ENoG degeneration greater than 90% relative to the unaffected side, no voluntary facial nerve EMG activity on the affected side).[55] Subjects who did not reach 90% degeneration on ENoG within 14 days of paralysis all returned to House-Brackmann grade I (n = 48) or II (n = 6) at 7 months after onset of the paralysis. Control subjects self-selecting not to undergo surgical decom-

pression when > 90% degeneration on ENoG and no motor unit potentials on EMG were identified had a 58% chance of having a poor outcome at 7 months after onset of paralysis (House-Brackmann grade III or IV [n = 19]). A group with similar ENoG and EMG findings undergoing middle fossa facial nerve decompression exhibited House-Brackmann grade I (n = 14) or II (n = 17) in 91% of the cases. It is recommended that these criteria and this surgical algorithm be followed.

Electromyography. Facial nerve EMG is important to use as an adjunctive tool when making decisions regarding surgery. Early in the time course of recovery, regenerating nerve fibers conduct at differing rates, producing dyssynchrony, and, therefore, overestimate the degree of wallerian degeneration based on ENoG testing. In fact, it is possible to record "100% degeneration" with ENoG in patients with early recovery from Bell's palsy while voluntary movement is observed. It is for this reason that a voluntary EMG is performed when the ENoG shows greater than 90% degeneration within 14 days of injury and surgical decompression is being considered. Needle electrodes are placed into the orbicularis oculi and orbicularis oris muscles, and the patient is asked to make voluntary contractions. If voluntary contractions occur during the first 2 weeks after the onset of paralysis, early deblocking of the neural conduction block has taken place, and a good recovery of facial function will most likely follow. It is also important to keep in mind that ENoG is useful only during the acute phase of the injury, between days 3 and 21, and after complete loss of voluntary function. Electromyography is the more useful single diagnostic study after 3 weeks of facial paralysis. As will be discussed later in the chapter, EMG testing is also important when deciding whether to perform nerve substitution procedures and other reamination procedures.

GENERAL PRINCIPLES IN FACIAL NERVE SURGERY

Whenever the facial nerve is to be surgically exposed, several technical points must be observed. First, a system for monitoring facial nerve function during the operation should be employed.[82] Historically, visual observation during critical stages of the operation was performed. However, the standard practice for most otologists is to use intraoperative facial nerve EMG with placement of needle electrodes into the orbicularis oculi and orbicularis oris muscles.

Instrumentation is crucial to successful exposure of the facial nerve. The largest diamond bur that the operative site can safely accommodate should be used when the surgeon is near the fallopian canal. Cutting burs have the potential to catch and jump unexpectedly and can consequently cause severe injury to the nerve. Continuous suction-irrigation keeps the burs clean and also dissipates heat, which can induce neural damage.

Blunt elevators, such as the Fisch raspatory (Leibinger, Dallas, Texas), should be used to remove the final layer of bone over the nerve. These instruments are thin but strong enough to remove a thin layer of bone. Stapes curettes are usually too large and can cause compression injury to the nerve. If a neurolysis is to be performed, disposable microblades are available (Beaver No. 59-10). Sharp dissection is less traumatic than blunt elevation when the nerve must be lifted out of the fallopian canal. The medial surface of the nerve usually adheres to the bone and contains a rich vascular supply. Cauterization near the nerve should be performed only with an irrigating bipolar electrocautery, low current, and insulated microforceps.

Middle Cranial Fossa (Transtemporal) Approach: Internal Auditory Canal Porus to Tympanic Segment

The middle cranial fossa exposure is used to expose the IAC and labyrinthine segment of the facial nerve when preserving existing auditory function is desirable.[57,83] The geniculate ganglion and tympanic portion of the nerve can also be decompressed from this approach.

Technique The patient is placed supine on the operating table with the head turned so that the involved temporal bone is upward (Figure 24–11, A). The hair is shaved 6 to 8 cm above and anterior to the ear and 2 cm posterior to it. The surgeon is seated at the head of the table with the instrument nurse at the anterior side of the patient's head. A 6 × 8 cm posteriorly based trapdoor incision, or a preauricular incision, is marked in the hairline above the ear (Figure 24–11, B). If exposure of the mastoid is necessary, the inferior limb of the incision can be carried postauricularly (Figure 24–11, B, dashed

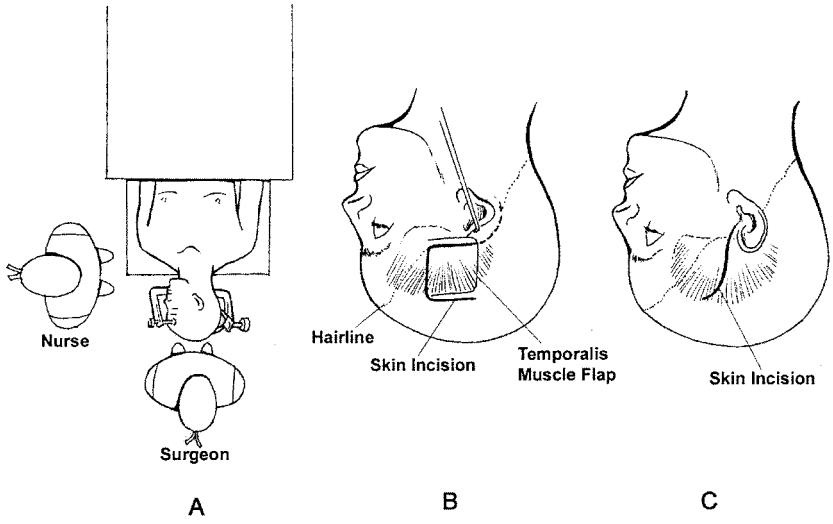


FIGURE 24–11. Patient positioning and incision design for the middle cranial fossa approach. *A*, Patient is in the supine position with the operated ear upward. Surgeon is seated at the vertex of the head. *B*, Posteriorly based trapdoor scalp incision (*bold line*). Surgical position illustrating the skin incision (*solid line*) for the middle cranial fossa approach. The *dashed line* shows the extension of the scalp incision that is required to reach the mastoid area for total facial nerve exposure. Design of anteriorly based temporalis muscle, fascia, and periosteal flap (*thin line*). *C*, Alternative preauricular incision for the middle cranial fossa approach if mastoid exposure is not necessary. Reproduced with permission; copyright © 2001 P. A. Wackym.

line). The skin flap is elevated to expose the temporalis muscle and fascia. A 4 × 4 cm temporalis fascia graft is harvested for use during closure of the IAC dural defect. Alternatively, Alloderm (LifeCell Corp., Woodlands, Texas) can be used. An anteriorly based trapdoor incision is used to elevate the temporalis muscle and periosteum (see Figure 24–11, B, thin line). Staggering the levels of the muscle and skin incisions provides for a double-layer, watertight closure at the completion of the procedure. If exposure of the mastoid is not necessary, a preauricular incision is often used (Figure 24–11, C).

The temporal root of the zygoma is exposed during elevation of the temporalis muscle. This landmark represents the level of the floor of the middle fossa. Dural fishhooks are placed in the skin and temporalis muscle flaps for retraction. A 3 × 5 cm bone flap for facial nerve decompression, or a 4 × 5 cm bone flap for tumor excisions, centered above the temporal root of the zygoma is fashioned with a medium-cutting bur (3 mm). It is important to keep the anterior and posterior margins of the craniotomy parallel to facilitate placement of the self-retaining retractor.

Branches of the middle meningeal artery are occasionally embedded within the inner table of the skull; therefore, elevation of the bone flap must be performed in a controlled manner. Bipolar coagulation and bone wax may be necessary to control bleeding. Elevation of the dura from the floor of the middle fossa can be one of the most difficult steps. Blunt dissection and magnification greatly facilitate dural elevation. The dura is elevated from the posterior to anterior direction to prevent accidental injury to an exposed geniculate ganglion and greater superficial petrosal nerve. Bipolar coagulation is used to cauterize dural reflections within the petrosquamous suture before transection with scissors.

The elevation proceeds until the petrous ridge is identified medially and the arcuate eminence, meatal plane, and greater superficial petrosal nerve are exposed anteriorly. No attempt is made to identify the middle meningeal artery and accompanying troublesome bleeding veins. The tip of a self-retaining retractor (Fisch, Leibinger) is placed at the petrous ridge anterior to the arcuate eminence and medial to the meatal plane (Figure 24–12). A medium diamond bur (2 to 3 mm) and a suction-

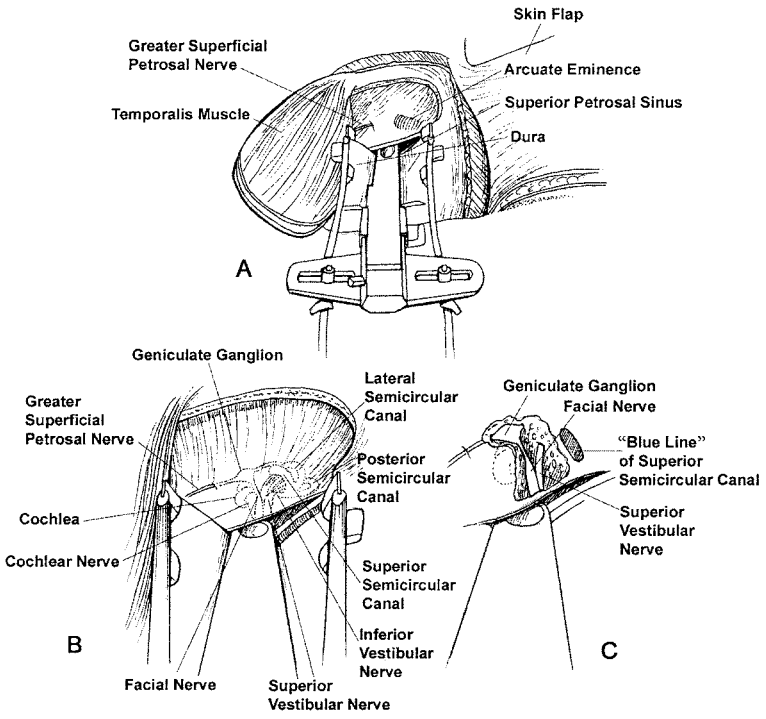


FIGURE 24–12. Surgical view of the right middle cranial fossa exposure after craniotomy, temporal lobe retraction, and bony exposure are complete. *A*, Placement of the retractor with exposure of the arcuate eminence and greater superficial petrosal nerve. *B*, The facial nerve, cochlea, and labyrinth within the temporal bone are shown ghosted beneath the bone as viewed through the middle cranial fossa approach. *C*, Drilling of the internal auditory canal for facial nerve decompression. Reproduced with permission; copyright © 2001 P. A. Wackym.

irrigation apparatus are used to identify the blue line of the SSC. A preoperative Stenvers projection radiograph helps to determine the level of the SSC in relation to the floor of the middle fossa and the degree of pneumatization above the SSC (Figure 24–13).

Drilling begins posterior to the arcuate eminence over the mastoid air cells until the dense yellow bone of the otic capsule is identified. Otic capsule bone is slowly removed until the blue outline of the SSC is seen. The IAC is located by removing bone with a 60-degree angle anterior to the blue line of the SSC and with the vertex based at the SSC ampulla. This dissection is continued until approximately 180 degrees of the IAC are exposed for facial nerve decompressions (see Figure 24–12) or 270 degrees of the IAC are exposed for schwannomas. Because of the close proximity of the SSC and the basal turn of the cochlea, only approximately 120 degrees of the circumference of the IAC can be safely removed in its lateral 5 mm or so. The facial nerve occupies the anterosuperior portion of the IAC. Laterally, the vertical crest (Bill's bar) marks the division between the superior vestibular nerve and the meatal foramen containing the facial nerve.

The entrance to the fallopian canal is the narrowest, most delicate portion of the facial nerve and consequently the most challenging portion of the dissection. At the meatal foramen, the facial nerve turns anterior and slightly superior. The basal turn of the cochlea can be within 1 mm inferiorly, and the ampulla of the SSC can be directly posterior to the nerve. The labyrinthine segment is followed to the geniculate ganglion. If the facial nerve needs to be exposed distal to the geniculate ganglion (eg, as with facial neuromas or with some traumatic injuries to the facial nerve), the tegmen tympani is removed with care to avoid injury to the head of the malleus and incus. The tympanic segment is easily seen to turn abruptly posterior; it is followed to where it courses inferior to the lateral semicircular canal. It is advisable to leave a thin shell of bone covering the nerve until its entire course is identified. Small blunt elevators are used to remove the final layer of bone. The nerve is tightly confined within the labyrinthine segment of the fallopian canal; larger curettes should be avoided to prevent compression injury. If the nerve is to be decompressed, a neurolysis is the final step. A disposable microscalpel (Beaver No. 59-10) is used to slit the periosteum and epineural sheath.

Alternative methods to locate the facial nerve may be necessary, especially in traumatic cases. The greater superficial petrosal nerve can be traced pos-

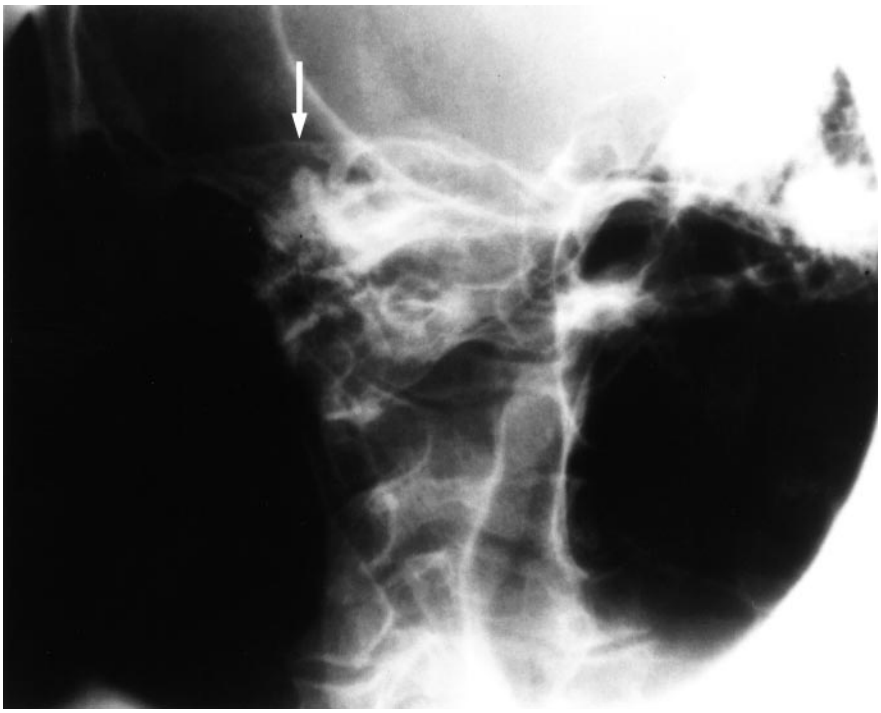


FIGURE 24–13. Stenvers projection radiograph demonstrating the anatomic variation of pneumatized air cells beneath the floor of the middle cranial fossa and the superior semicircular canal (*arrow*). Reproduced with permission; copyright © 1995 P. A. Wackym.

teriorly to the geniculate ganglion, or the tegmen tympani may be fractured and the tympanic segment visible through the fracture. The tympanic segment is then used to locate the geniculate ganglion and labyrinthine segments.

At the end of the procedure, a free temporalis muscle graft is placed within the IAC and a corner piece of the bone flap is fashioned to cover the defects in the tegmen tympani and IAC (Figure 24–14). Alternatively, titanium mesh (Synthes Maxillofacial, Paoli, Pennsylvania) between layers of Alloderm can be used. This prevents herniation of the temporal lobe into the middle ear or IAC. The temporalis fascia previously harvested is placed over the free bone graft to help seal the dural defect at the IAC. The craniotomy defect is then repaired using titanium mesh (Synthes Maxillofacial) and hydroxyapatite cement (BoneSource, Leibinger), and the temporalis muscle is closed with interrupted absorbable sutures. The skin is closed in layers with particular care in closing the galea. No drain is placed. A mastoid-type pressure dressing is applied.

Advantages and Uses The middle cranial fossa route is the only method that can be used to expose the entire IAC and labyrinthine segment with preservation of hearing. This, in combination with the retrolabyrinthine and transmastoid approaches, enables visualization of the entire course of the facial nerve and still preserves function of the inner ear. The middle cranial fossa technique is the most commonly used for decompression of the facial nerve in Bell's palsy[52,57] and longitudinal temporal bone fractures. However, as described earlier in this chapter, this approach may be useful in the management of patients with schwannomas of CN VII or CN VIII, as well as with patients with Melkersson-Rosenthal syndrome.

Postoperative Care/Complications and Their Management The anatomy of the floor of the middle cranial fossa is quite variable and presents some difficulty in identification of landmarks. The Stenvers projection radiograph provides important anatomic information regarding the degree of pneumatization above the SSC and should be performed in all cases to minimize the risk of surgical injury to the SSC. In addition, the surgeon must have a precise knowledge of three-dimensional anatomy of the temporal bone. Many hours in a temporal bone dissection laboratory are required to attain the delicate microsurgical skills that are necessary for this type of surgery.

Middle cranial fossa facial nerve decompression can result in conductive and/or sensorineural hearing loss. Conductive hearing loss can be sec-

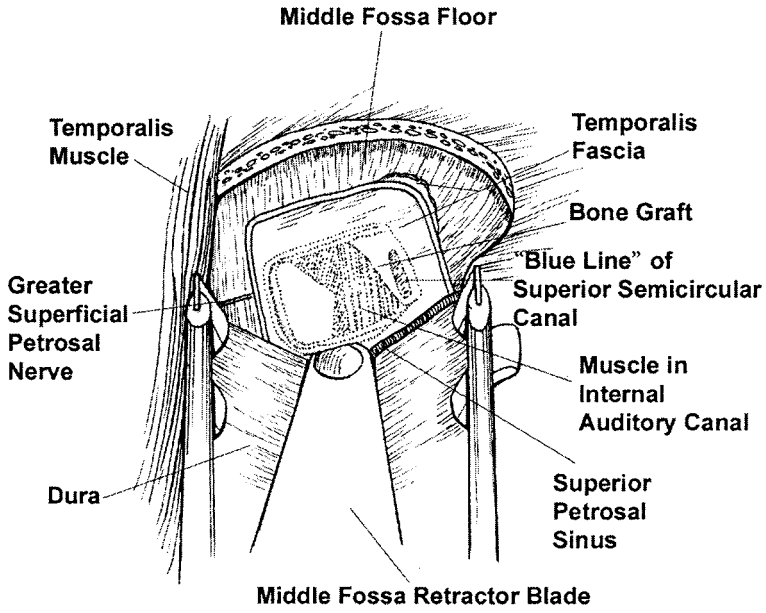


FIGURE 24–14. Reconstruction following right middle cranial fossa facial nerve decompression. Temporalis muscle is placed within the internal auditory canal (IAC) surgical defect to fill the dural and bony defect at the conclusion of the facial nerve surgical decompression. Temporalis fascia or Alloderm (LifeCell Corp., Woodlands, Texas) is placed over the floor of the middle cranial fossa and muscle graft. A free bone graft or titanium mesh (Synthes Maxillofacial, Paoli, Pennsylvania) is placed perpendicular to the axis of the IAC to prevent herniation of the temporal lobe onto the facial, cochlear, and vestibular nerves. Temporalis fascia or Alloderm is then used to seal the temporal lobe dura and cover the bone graft or titanium mesh. Reproduced with permission; copyright © 2001 P. A. Wackym.

ondary to temporal lobe herniation or ossicular disruption during dissection in the attic. A free bone graft, as already described, prevents temporal lobe herniation. Sensorineural hearing loss can result from direct injury to the inner ear by the drill exposing the cochlea or semicircular canals or from translational injury by the drill striking an ossicle. Should the SSC be entered during the surgical dissection, the fenestration should be immediately occluded with bone wax. Injury to the internal auditory vessels within the IAC can also result in loss of inner ear function. Loss of vestibular function can occur by the same mechanisms.

Postoperative intracranial complications including meningitis, temporal lobe edema, and epidural hematoma formation are possible. Perioperative antibiotics administered over 48 hours are recommended. Fluid restriction and dexamethasone (Decadron) are used for the first 3 days postoperatively to minimize temporal lobe edema following intraoperative retraction. In addition, our longer craniotomy flap decreases the amount of temporal lobe retraction required for complete exposure of the IAC and fallopian canal. With adequate intraoperative hemostasis using the bipolar cautery, oxidized cellulose (Oxycel), and dural tacking sutures, we have never had a clinically significant postoperative epidural hematoma develop.

Leakage of cerebrospinal fluid (CSF) must be avoided to prevent meningitis. All exposed mastoid air cells must be obstructed with bone wax. A temporalis muscle–free graft is placed into the superior aspect of the IAC to separate the posterior fossa from the extradural floor of the middle cranial fossa. Temporalis fascia or Alloderm is then used to provide a second layer of closure between the posterior fossa and the extradural middle fossa. Meticulous care must be taken to ensure that there are no dural dehiscences overlying the temporal lobe through which CSF may drain. If these are identified, a temporalis fascia, muscle, or Alloderm patch must be used to repair the dural tears to prevent CSF leaks. After a three-layer watertight closure of the temporalis muscle, galea, and scalp, a mastoid-type dress-

ing is applied daily for 5 days postoperatively. Should CSF leakage persist, a temporary lumbar fluid drain is placed, and the patient is kept at bed rest. If the CSF leakage does not resolve within 5 to 7 days after placement of the lumbar drain, re-exploration of the surgical field is indicated to identify and seal the area of CSF egress.

Uncontrolled bleeding or injury to the AICA poses the most serious complication during the operation; however, this is rare. The middle cranial fossa approach does not provide adequate access to the entire cerebellopontine angle. The AICA and accompanying veins can loop into the IAC. Control of bleeding of these vessels may require a suboccipital exposure. Injury to the AICA results in brainstem and cerebellar infarction of a variable degree, depending on its size and the area of its terminal arterial supply.

NERVE REPAIR

Whenever the continuity of the facial nerve has been disrupted by trauma, iatrogenic injury, or tumor invasion, every effort should be made to restore its continuity. In some instances, an end-to-end reapproximation can be accomplished, but if any tension occurs at the anastomotic site, an interposition nerve graft has a better chance of providing facial movement. All nerve repair techniques produce synkinesis, but sphincteric function of the mouth and eye is usually restored. Newer microsuture techniques and instrumentation should be employed to enhance return of function.

In general, the injured ends of the nerve should be freshened at a 45-degree angle. Experimental evidence has shown that cutting the nerve at this angle exposes more neural tubules and improves regrowth of the nerve.[81,84] In addition, a fresh razor blade induces less crush injury to the nerve than a scalpel blade or scissors does. We have found that the perineurium of CN VII does not hold 9-0 sutures, and attempting to suture it increases trauma to the neural tubules. Removing a portion of the epineurium before suturing prevents connective tissue growth at the anastomotic site. If the epineurium is cleaned from the end of the nerve for approximately only 0.5 mm, sutures can still be placed in the epineurium for reapproximating the nerve segments. Three or four 9-0 nylon sutures are placed with jeweler's forceps or longer instruments (19 cm microforceps) for anastomosis in the cerebellopontine angle. At the brainstem, two or three sutures are placed (Figure 24–15).

When an interposition graft is required, the greater auricular and sural nerves are the preferred graft donor sources. The greater auricular nerve is readily available near the operative field if it is not involved in resection of a neoplasm and has approximately the same diameter as that of the facial nerve. It is easily located midway, perpendicular to a line drawn between the mastoid tip and the angle of the mandible. If a graft of greater than 8 to 10 cm is required, the sural nerve should be used. The sural nerve has another advantage in that the peripheral portion of the nerve has many branches that can be used to reconstruct the branching pattern of the facial nerve. There is little discomfort from removing the sural nerve since it provides only a small area of sensation to the lateral lower leg and foot. The sural nerve is found immediately posterior to the lateral malleolus, along the saphenous vein. The nerve graft should be 10 to 20% larger in diameter than the facial nerve and long enough to ensure a tension-free anastomosis.

FACIAL REANIMATION PROCEDURES

Contemporary strategies for management of patients with facial paralysis have been a product of a gradual evolution of past clinical successes and

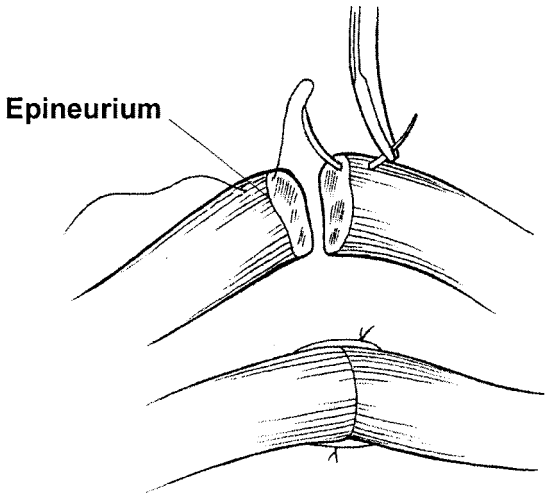


FIGURE 24–15. Epineurial end-to-end nerve repair using 9-0 nylon sutures. Reproduced with permission; copyright © 2001 P. A. Wackym.

failures over the past 100 years. The surgeon must consider the functional and cosmetic goals of reconstructive surgery, as well as the patient's desires, expectations, and motivations. Functional deficits include incomplete eye closure, speech difficulties, oral incompetence, and nasal airway obstruction. The cosmetic deficiencies of facial asymmetry and dysmorphism can be emotionally devastating for some patients. It is the achievement of facial balance and muscle coordination that continues to be the more challenging and elusive goal.

Management of Upper Third of the Face

Eye Care Protection of the eye is paramount (Figure 24–16). It is necessary to protect the cornea from foreign bodies and drying. Dark glasses should be worn during the day, artificial tears instilled at the slightest evidence of drying, and a bland eye ointment used during sleep. Patients who demonstrate a poor Bell's phenomenon or have trigeminal nerve deficits are particularly at risk for corneal damage. Taping of the eye closed is not usually recommended, but early-exposure keratitis may require patching or, rarely, a tarsorrhaphy. A formal ophthalmologic examination is recommended prior to any surgical intervention.

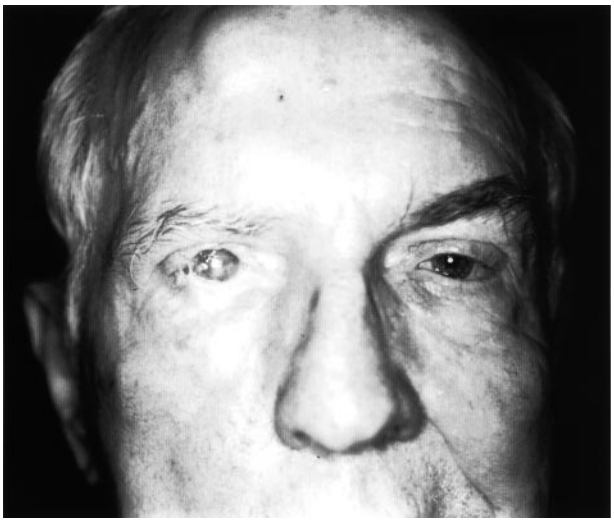


FIGURE 24–16. Inadequate eye care following a complete facial paralysis owing to acoustic neuroma resection at another institution resulted in corneal opacification. Reproduced with permission; copyright © 2000 P. A. Wackym.

Eyebrow Ptosis of the eyebrow can have functional and cosmetic consequences. In the elderly patient, the functional loss of the frontalis and orbicularis oculi muscles is compounded by the loss of tissue elasticity and decrease in the bulk of the subcutaneous tissue. This can lead to significant brow ptosis and hooding of the upper eyelid, which may cause lateral visual field compromise.

The two most commonly used procedures to correct brow ptosis are the midforehead lift and the direct brow lift. Both procedures require direct skin and subcutaneous tissue excision, followed by suspension of the orbicularis oculi muscle to the frontal bone periosteum. Slight overcorrection is needed as the brow position will settle during the next few weeks. The endoscopic approach for cosmetic browlifting is not commonly used owing to the severity of the brow ptosis associated with the complete loss of frontalis muscle function. However, there have been some recent anecdotal reports of successful outcomes using the endoscopic browlift technique.[1,85]

Upper Eye Lid Historically, tarsorrhaphy had been the standard of care in patients with facial paralysis. Today, this procedure should be reserved for only those patients with a severe risk for exposure keratitis or those who have failed upper eyelid reanimation procedures. The most commonly used procedure is the insertion of a prosthetic, specifically a gold weight implant or a palpebral wire spring.

In experienced hands, the palpebral wire spring can produce excellent results, affording the capability of mimicking, to some extent, the spontaneous blink. However, the insertion of the palpebral wire spring is technically more difficult, with a higher reported extrusion and infection rate. In addition, these springs often need postoperative adjustment for optimal function.[1]

Gold weight implantation is a relatively simple procedure that is highly successful, well tolerated by patients, and easily reversible if facial muscle function returns (Figure 24–17). The ideal candidates for gold weight placement are those with the following factors: some existing ability to lower the upper lid, good Bell's phenomenon, normal corneal sensation and tearing, prominent supratarsal lid crease, and nonprotruding eyes.

Prefabricated gold weight implants come in weights ranging from 0.8 to 1.6 g. Custom weighting

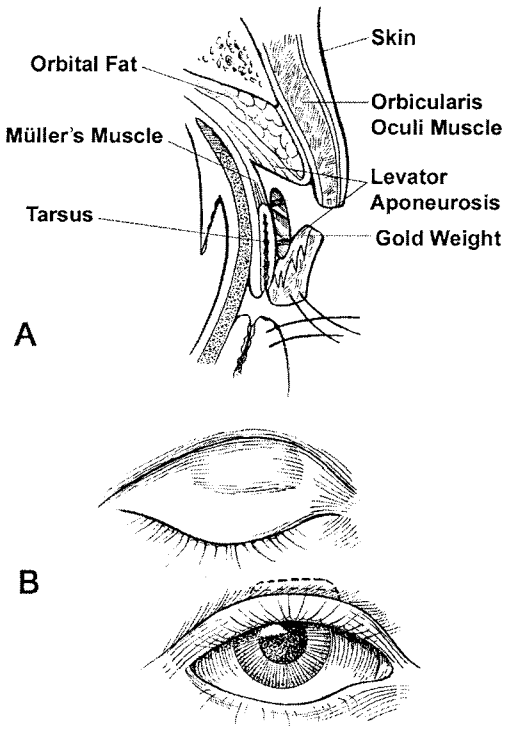


FIGURE 24–17. A, Gold weight implant in position, superficial to the tarsal plate, deep to the levator aponeurosis. *B*, Correct gold weight position in relation to the lid margin and pupil. Reproduced with permission; copyright © 2001 P. A. Wackym.

can be requested, if needed. All patients should be correctly sized by taping different size weights to the eyelid in an upright position. The largest weight allowing eyelid closure without causing more than slight lid ptosis should be chosen; the most common weights are 1.0 and 1.2 g. The patient should be informed that the weights are often not helpful when lying supine. In fact, eyelid closure may be worse in some cases owing to the weight of the implant pulling the eyelid open when supine. Eye care, as described above, should be continued following the reanimation procedure.

After selecting the proper size implant, the procedure is performed under local anesthesia. An incision is made along the supratarsal fold down to the orbicularis oculi muscle. Sharp scissor dissection is carried down through the muscle to the tarsal plate. Dissection is continued inferiorly over the surface of the tarsal plate to within 3 mm of the lash line. An exact pocket is made for the implant to be placed just slightly medial to the center of the pupil. The gold weight is then inserted and suture fixated to the tarsus using 6-0 clear prolene sutures. The incision is closed in layers, 6-0 chromic suture for the levator aponeurosis/orbicularis muscle layer and 6-0 fast gut suture for the skin.

The most common problems associated with gold weight implantation are lid underclosure, excessive lid ptosis, and implant extrusion. The reported extrusion rate in the literature is greatly varied, from as low as 1% to as high as 43%.[1,86,87] The risk of implant extrusion can be minimized by meticulous surgical technique and proper patient selection. Removal of the implant is performed once facial nerve function has returned; however, the gold weight can also be used as a permanent means to achieve eye closure.

Lower Eye Lid The goals for lower lid management are to improve lower lid margin approximation to the globe, correct ectropion, and maximize the efficiency of the tear drainage system. Lid-tightening procedures must not disturb the delicate interface between the lacrimal punctum and the globe. A lateral traction test, simulating a lid-tightening procedure, will demonstrate the effect of the procedure on lid position and the displacement of the punctum. As a general rule, up to one-eighth of the lid can be resected without disturbing the relationship of the inferior punctum to the globe. If the lateral traction test indicates excessive punctum displacement (greater than 2 mm) or does not provide proper lid support, then alternative or adjunctive surgical procedures are indicated. Excessive punctum lateralization indicates medial canthal tendon laxity, thus indicating the necessity for a medial canthoplasty. If further elevation of the lower lid is needed, then "spacer" grafts (palate mucosa, conchal cartilage) are used to provide vertical height to the eyelid.

Lower lid–tightening procedures include the Bick procedure, tarsal strip, and midlid wedge resection. The Bick procedure is our procedure of choice for lower lid tightening. The procedure involves resecting a lateral wedge of the lower lid, developing a tongue of tarsus, and resuturing the lateral edge of the lid to the lateral orbital rim (Figure 24–18). The procedure allows for fine adjustment of the tension on the lower lid by resecting a precise amount of tissue. Care must be taken not to overshorten the lower

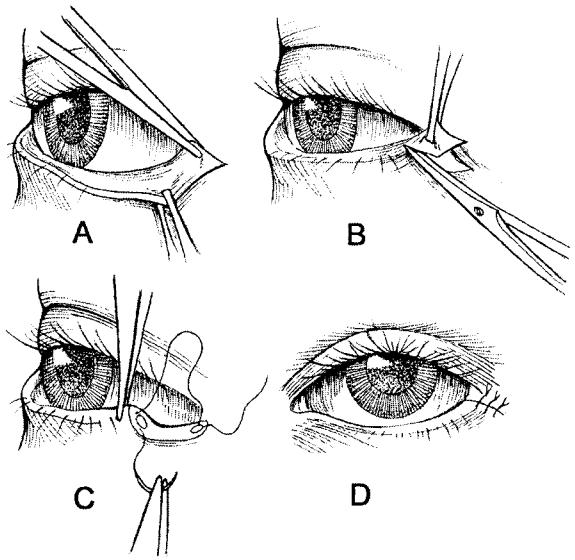


FIGURE 24–18. Bick procedure for lower lid ectropion repair. *A*, Lower lid is grasped and canthotomy and cantholysis are performed. *B*, Appropriate amount of lax lower lid (skin, muscle, conjunctiva, tarsus) is sectioned. *C*, Lateral margin of tarsus is sutured to canthal tendon and orbital periosteum. *D*, Canthotomy is closed with absorbable sutures. Reproduced with permission; copyright © 2001 P. A. Wackym.

lid, creating a hammock effect, with the lower lid actually bowing down under the globe.

Management of Lower Two-Thirds of the Face

The ultimate goal in treatment of the lower two-thirds of the face is to create symmetric, mimetic movement of the facial musculature. The best chance for this outcome is with primary repair of the facial nerve, with or without nerve interposition grafting. However, primary nerve repair is not always possible, and alternative procedures must be entertained. The choice of procedure is dictated by a number of factors including duration of paralysis, prognosis of the underlying illness, concomitant CN deficits, comorbid medical conditions, and the patient's wishes and motivations.

Role of Electromyography in Planning Facial Reanimation Procedures The degree of motor end-plate degeneration and prognosis for spontaneous recovery of the facial nerve can be assessed with EMG. The presence of normal or polyphasic action potentials at 1 year following facial nerve injury portends a favorable outcome, and no reanimation procedures are indicated. If fibrillation potentials are found, this indicates intact motor end plates but no evidence of reinnervation. This finding supports the use of a nerve substitution procedure to take advantage of the potential neurotized tone and movement of the intrinsic facial musculature. Electrical silence obtained from EMG indicates complete denervation and atrophy of the motor end plates. Neurotized reanimation procedures are contraindicated, and other reanimation procedures should be entertained.

Nerve Substitution Procedures Nerve substitution procedures are indicated when primary facial nerve repair is not possible. In addition, the following conditions are required for these procedures: intact proximal donor nerve, intact distal facial nerve, and viable motor end-plate function. Although these procedures provide facial tone and resting symmetry, they do not restore involuntary, independent mimetic facial expression. However, the majority of patients do achieve some voluntary facial movement with rehabilitation. The most commonly described nerve substitution procedures are the XI–VII crossover, VII–VII crossfacial, XII–VII crossover, and XII–VII jump graft.

The XI–VII procedure (accessory to distal facial nerve) has largely been abandoned owing to the morbidity of loss of trapezius muscle function. The VII–VII crossfacial grafting involves linking a functional branch of the facial nerve on the nonparalyzed side to a division of the facial nerve on the paralyzed side by using a long interposition nerve graft (sural, medial antebrachial cutaneous) (Figure 24–19). The disadvantages of this procedure include the sacrifice of a portion of the normal facial function on the contralateral side, a long interval for innervation (9 to 12 months), and a lack of substantial neural "firepower" owing to the relatively few number of axons grafted. Theoretically, the advantage of this technique is the possibility of symmetric, mimetic movement. Disappointing results from several investigators[1,88] have made this procedure less appealing.

Hypoglossal-facial (XII–VII) crossover is most appropriately performed for complete and permanent facial paralysis up to 2 years after injury (Figure

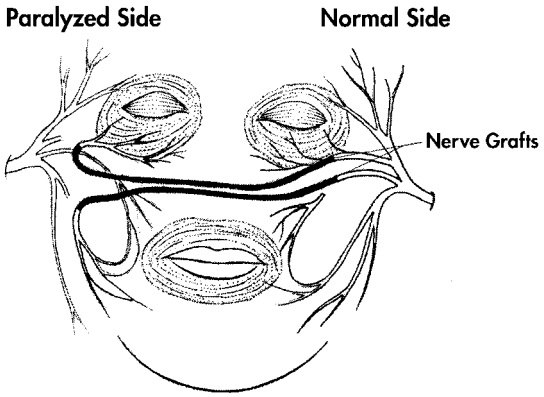


FIGURE 24–19. Facial-crossfacial nerve substitution technique. Reproduced with permission; copyright © 2001 P. A. Wackym.

24–20). This situation may arise following radical parotidectomy, temporal bone resection, skull base surgery, severe temporal bone trauma, or cerebellopontine angle tumor resection. An EMG finding consistent with viable motor end plates of the facial muscles is a prerequisite for surgery. Poor candidates for the XII–VII crossover include patients with multiple lower CN deficits. Sacrifice of the hypoglossal nerve may not be well tolerated and compensated for in the presence of other cranial neuropathies.

Improved facial tone and symmetry occur in 90% of patients following XII–VII anastomosis.[1] The results are more impressive in the midface and less in the frontalis and lower portions of the face. The return of muscle tone is seen within 4 to 6 months following neurorrhaphy, with better results seen in earlier repairs.[89] Voluntary facial movements follow with progressive improvements over the next 1 to 2 years. True spontaneous facial expressions are rare, although through motor sensory re-education, patients may develop spontaneous animation with speech.

Pensak et al, in their series of 61 cases of XII–VII crossover, rated their results as excellent in 3%, good in 39%, fair in 49%, and poor in 10% of patients.[90] Conley and Baker, in their series of 122 cases, reported good or excellent results in 65%, fair in 18%, and poor in 17%.[89] They also noted better results in those patients who underwent "immediate" repair (< 2 years) than in those who were "delayed" (> 2 years).

Synkinesis, hypertonia, and hemilingual atrophy are all noted deficiencies of the classic XII–VII crossover. Most studies report minimal disabling sequelae owing to loss of unilateral hypoglossal function unless concomitant ipsilateral CN deficits exist.

The XII–VII jump graft technique was devised by May et al to offset some of the above noted disadvantages of the classic XII–VII crossover.[91] The procedure entails placing an interpositional nerve graft between a partially transected hypoglossal nerve trunk, distal to the hypoglossal descendens,

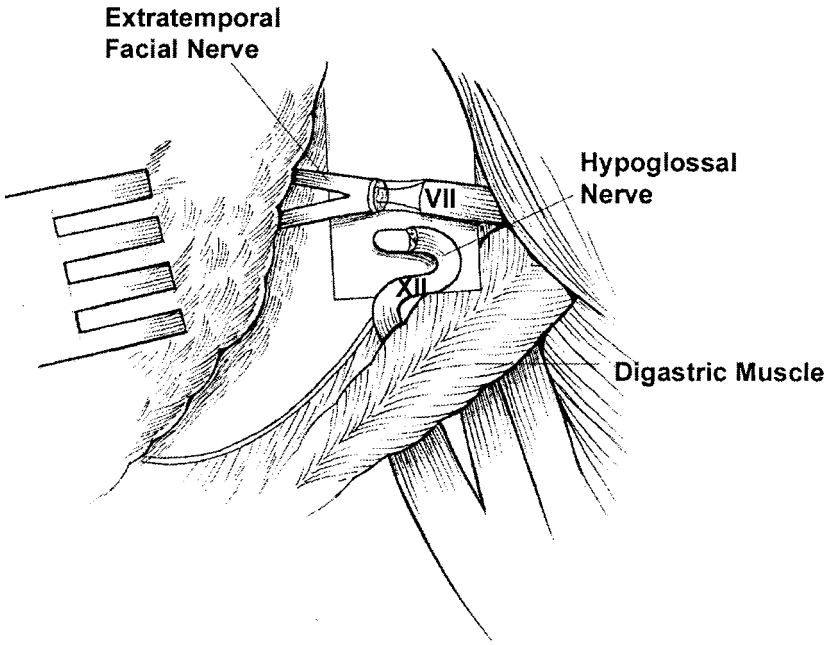


FIGURE 24–20. Hypoglossal-facial nerve crossover technique. Hypoglossal nerve is brought medial to digastric muscle to gain additional length and decrease tension at the suture line. Reproduced with permission; copyright © 2001 P. A. Wackym.

and the distal facial nerve trunk (Figure 24–21). Several authors have reported functional results comparable to the XII–VII crossover; however, the problems of hypertonia and mass facial movements have not been encountered.[91,92] Kartush and Lundy further modified this approach with anastomosis to only the lower division of the facial nerve, thereby reducing potential synkinesis.[93]

Muscle Transposition Procedures (Dynamic) Regional muscle transposition can provide dynamic reanimation of the mouth in patients with long-standing facial paralysis (over 2 years). It is indicated for patients with congenital facial paralysis (Möbius' syndrome) or when facial nerve grafting or nerve substitution techniques are contraindicated. It can also be performed in conjunction with facial nerve grafting or nerve substitution procedure in select cases to augment results. The temporalis muscle is most commonly used because of its length, contractility, and favorable vector of pull. Masseter muscle transposition can be useful following radical parotid surgery or when the temporalis muscle is not available.

Temporalis Muscle Transposition The temporalis is a fan-shaped muscle arising from the temporal fossa. The motor branch of the trigeminal nerve innervates it, and branches off the internal maxillary artery provide its blood supply. Preoperative evaluation includes the assessment of the function and strength of the temporalis muscle.

To perform a temporalis muscle transposition procedure, the temporalis muscle and fascia are both exposed by extending a modified facelift incision into the parietal region of the scalp (Figure 24–22). The pretragal incision is carried down to the subdermal region, above the superficial musculoaponeurotic system of the face. A tunnel is created from the zygomatic arch in a subdermal plane

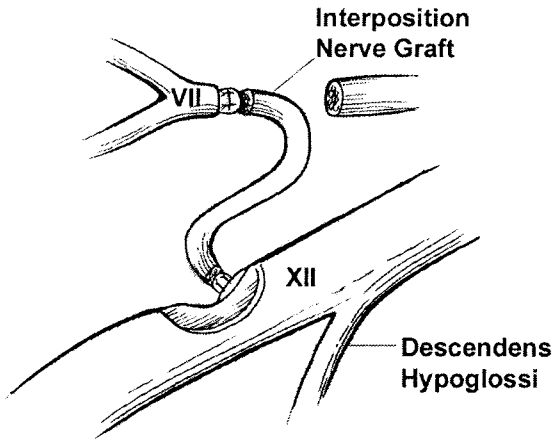


FIGURE 24–21. Hypoglossal-facial interpositional nerve graft technique. Approximately one-third of the hypoglossal nerve (XII) is sectioned, distal to the hypoglossal descendens. The nerve graft is sewn end to side with the hypoglossal nerve and end to end with the facial nerve (VII). Reproduced with permission; copyright © 2001 P. A. Wackym.

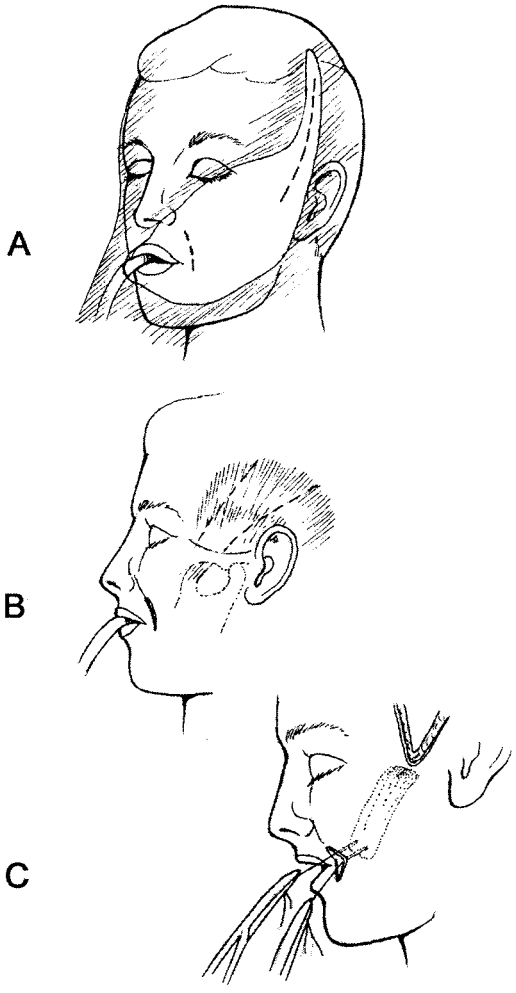


FIGURE 24–22. Temporalis muscle transfer for lower face reanimation. *A*, *Dotted lines* indicate the lip-cheek and temporal incisions. *B*, Midportion of the temporalis muscle is harvested. *C*, Temporalis muscle pulled through subcutaneous cheek tunnel into the lip-cheek incision. Reproduced with permission; copyright © 2001 P. A. Wackym.

between the temporal fossa and the cheek-lip crease incision (adult) or the oral commissure incision (children). The lateral portions of the zygomatic arch can be burred down to reduce the bulging appearance of the muscle at this fulcrum point. Only the midportion of the temporalis muscle is used. A portion of the pericranium is dissected with the muscle and its fascia to create additional length. The muscle is then transposed over the zygomatic arch, through the subcutaneous tunnel, and into the lip-cheek incision. The muscle-fascia-pericranium complex is sutured to the orbicularis oris muscle near the submucous layer of the corner of the mouth. Additional sutures are placed from the muscle to the subdermal layer of the upper aspect of the lip-cheek incision to accentuate the lip-cheek crease. Overcorrection is necessary for an optimal final result. The depression in the temporal area is corrected by rotating the remainder of the temporalis muscle into the deficient area. The redundant facial skin is resected as in a facelift, and wounds are closed in layers. Two closed suction drains and a compressive dressing are used for 48 hours postoperatively.

Ideally, the overcorrection of the lateral oral commissure and lip-cheek crease will resolve by 3 to 6 weeks. The results of the transposition should be evident by 4 to 6 weeks, with the patient able to produce a broad smile by tensing the temporalis muscle. Complications include infection, hematoma, and seroma. The most common reasons for failure of the procedure are inadequate overcorrection and suture dehiscence at the orbicularis oris–temporalis muscle interface.[94]

Suspension Procedures (Static) Static suspension procedures are indicated for those patients who are not candidates for nerve substitution or dynamic reanimation procedures. These procedures can provide permanent support, or, in cases in which reinnervation of the facial muscles is expected, static procedures can provide temporary or additional support until reinnervation of the facial muscles is complete. A variety of materials are available for static suspension procedures, ranging from autografts (palmaris longus tendon, fascia lata tendon) to alloplasts, such as Gore-Tex (GORE S.A.M., WL Gore & Associates Inc., Flagstaff, Arizona) and acellular human dermis (Alloderm).[95,96] The choice of material is dependent on patient factors and desires, as well as the surgeon's preference and experience.

The static suspension procedure is most often used to lateralize the corner of the mouth and lip-cheek crease. The procedure is similar in approach to the temporalis muscle transposition. The autograft or allograft is placed in a subcutaneous tunnel from the lip-cheek crease to the zygomatic arch. In cases in which reinnervation of the facial muscles is expected, it is important not to injure the deeper facial nerve branches when creating the tunnel. Sutures are used to secure the graft to the orbicularis oris muscle and lip-cheek incision, whereas wires or screws are used for anchoring at the zygomatic arch (Figure 24–23).

A separate local suspension can also be performed to lateralize the nose in cases of alar valve compromise. The nasal base is exposed with incisions along the alar-facial and nasolabial creases. An allograft or autograft is then sutured to the deep aspect of the nasal base. The nasal base is then pulled laterally to its desired position. After the periosteum of zygomaticomaxillary buttress is exposed, the other end of the graft is fixated to the bone with a screw.

Innervated Free Muscle Transfer

The ideal indication for innervated free muscle transfer is in the patient with Möbius' syndrome, for whom both facial nerve and musculature are not available. It is also indicated as an alternative to regional muscle transfers or static procedures in patients with long-standing facial paralysis (> 2 years). It is usually performed as a two-stage procedure in which an initial crossfacial nerve graft is combined with a subsequent free muscle transfer (most commonly, gracilis or serratus anterior).[97] Alternatively, the innervated free flap may be grafted to the hypoglossal nerve, performed as a single-stage procedure.[98]

In select patients, this procedure provides the possibility for dynamic, mimetic movement that cannot be achieved by static procedures. The disadvantages, however, are manifold, including donor site morbidity, risk of vascular thrombosis, lengthy operative time, long interval for reinnervation, and muscle bulkiness.

NEUROMUSCULAR FACIAL RETRAINING TECHNIQUES

For patients who have experienced some recovery of facial nerve function, and also for those patients who

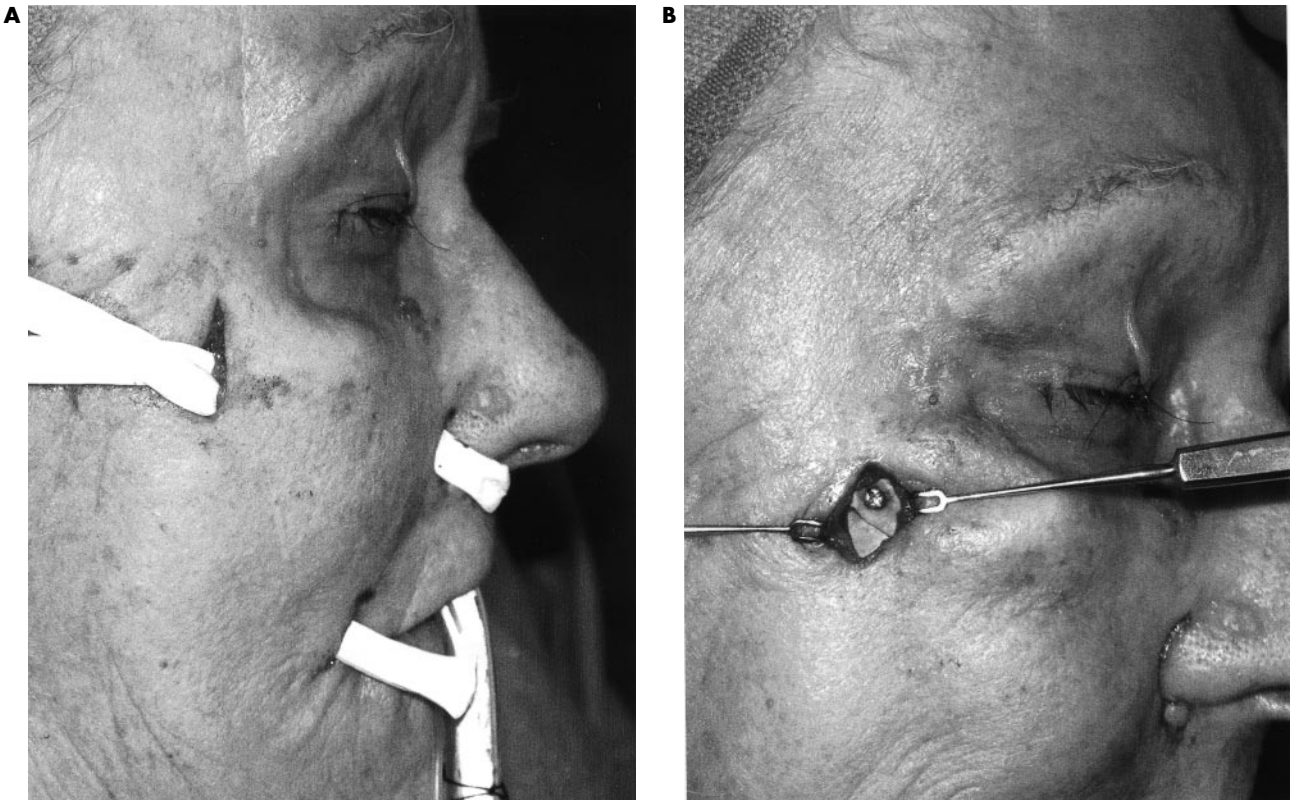


FIGURE 24–23. Static suspension procedure using allograft (Gore 2 mm nonreinforced sheet, W.L. Gore & Associates, Inc., Flagstaff, Arizona). *A*, Strips of allograft placed in the subcutaneous cheek tunnel. *B*, Allograft anchored to the malar eminence with a single titanium screw.

experience synkinesis, neuromuscular facial retraining therapy is an important treatment modality. These techniques can be applied before and after reanimation procedures to optimize outcome. In general, these techniques can be used to address loss of strength, loss of isolated motor control, muscle tension hypertonicity, and/or synkinesis. This method combines techniques such as patient education in basic facial anatomy, physiology, and kinesiology; relaxation training; sensory stimulation; EMG biofeedback; voluntary facial exercises with mirror feedback; and spontaneously elicited facial movements.[99–101] Botulinum toxin injections may be used to augment the results of neuromuscular facial retraining when dealing with synkinesis and hypertonicity.[102]

SUMMARY

The etiology and pathogenesis of several disorders that produce facial paralysis have been reviewed. Likewise, the diagnostic methods and algorithms that are used to make decisions regarding medical versus surgical management in acute facial paralysis have been summarized. Anatomy and surgical principles were highlighted and treatment alternatives for reanimation or neuromuscular facial retraining outlined.

REFERENCES

1. May M, Schaitkin BM, editors. The facial nerve. 2nd ed. New York: Thieme Medical Publishers; 2000.
2. Harris L. Public attitudes toward people with disabilities. Report of a U.S. survey to the National Organization on Disability. Washington (DC): 1991.
3. McKusick VA. Mendelian inheritance in man. A catalog of human genes and genetic disorders. 12th ed. Baltimore: The Johns Hopkins University Press; 1998.
4. Huebner O. Über angeborenen Kernmangel (infantiler Kernschwund, Möbius). Charite Ann 1900;25: 211–43.

5. Hanissian AS, Fuste F, Hayes WT, Duncan JM. Moebius syndrome in twins. Am J Dis Child 1970;120: 472–5.
6. Pitner SE, Edwards JE, McCormick WF. Observations on the pathology of the Möbius syndrome. J Neurol Neurosurg Psychiatry 1965;28:362–74.
7. Ebbesen F, Petersen W. Goldenhar's syndrome: discordance in monozygotic twins and unusual anomalies. Acta Paediatr Scand 1982;71:685–7.
8. Sieff CA, Chessells, Levinsky RJ, et al. Allogenic bone-marrow transplantation in infantile malignant osteopetrosis. Lancet 1983;1:437–41.
9. Toh EHY, Wackym PA. Injury to the temporal bone and facial nerve in children. Facial Plast Surg Clin North Am 1999;7:223–9.
10. Feuerman T, Wackym PA, Gade GF, Becker DP. Value of skull radiography, head computed tomographic scanning, and admission for observation in cases of minor head injury. Neurosurgery 1988;22:449–53.
11. Glasscock ME, Schwaber MK, Nissen AJ, Jackson CG. Management of congenital ear malformations. Ann Otol Rhinol Laryngol 1983;92:504–9.
12. Lambert PR. Major congenital ear malformations: surgical management and results. Ann Otol Rhinol Laryngol 1988;97:641–9.
13. Takahashi H, Nakamura H, Yui M, Mori H. Analysis of fifty cases of facial palsy due to otitis media. Arch Otorhinolaryngol 1985;241:163–8.
14. Sheehy JL. Cholesteatoma surgery in children. Acta Otorhinolaryngol Belg 1980;34:98–106.
15. Glasscock ME, Pensak ML, Gulya AJ, Baker DC. Lyme disease. A cause of bilateral facial paralysis. Arch Otolaryngol Head Neck Surg 1985;111:47–9.
16. Hunt JR. On herpetic inflammations of the geniculate ganglion. A new syndrome and its complications. J Nerv Ment Dis 1907;34:73–96.
17. Head H, Campbell AW. Pathology of herpes zoster and its bearing on sensory localization. Brain 1900;23:353–523.
18. Devriese PP. Facial paralysis in cephalic herpes zoster. Ann Otol Rhinol Laryngol 1968;77:1101–19.
19. Crabtree JA. Herpes zoster oticus. Laryngoscope 1968;78:1853–78.
20. Byl FM, Adour KK. Auditory symptoms associated with herpes zoster or idiopathic facial paralysis. Laryngoscope 1977;87:372–9.
21. Robillard RB, Hilsinger RL, Adour KK. Ramsay Hunt facial paralysis: clinical analysis of 185 patients. Otolaryngol Head Neck Surg 1986;95:292–7.
22. May M, Fria TJ, Blumenthal F, Curtin H. Facial paralysis in children: differential diagnosis. Otolaryngol Head Neck Surg 1981;89:841–8.
23. Uri N, Greenberg E, Meyer W, Kitzes-Cohen R. Herpes zoster oticus: treatment with acyclovir. Ann Otol Rhinol Laryngol 1992;101:161–2.
24. Hunt JR. The sensory field of the facial nerve. A further contribution to the symptomatology of the geniculate ganglion. Brain 1915;38:415–45.
25. Hunt JR. Geniculate neuralgia (neuralgia of the nervus facialis). A further contribution to the sensory system of the seventh nerve and its neuralgic conditions. Arch Neurol Psychiatry 1937;37: 253–85.
26. Korzec K, Sobol SM, Kubal W, et al. Gadolinium-enhanced magnetic resonance imaging of the facial nerve in herpes zoster oticus and Bell's palsy: clinical implications. Am J Otol 1991;12:163–8.
27. Blackley B, Friedmann I, Wright I. Herpes zoster auris associated with facial nerve palsy and auditory nerve symptoms: a case report with histopathological findings. Acta Otolaryngol (Stockh) 1967;63: 533–50.
28. Etholm B, Schuknecht HF. Pathological findings and surgical implications in herpes zoster oticus. Adv Otorhinolaryngol 1983;31:184–90.
29. Guldberg-Möller J, Olsen S, Kettel K. Histopathology of the facial nerve in herpes zoster oticus. Arch Otolaryngol 1959;69:266–75.
30. Wackym PA. Molecular temporal bone pathology: II. Ramsay Hunt syndrome (herpes zoster oticus). Laryngoscope 1997;107:1165–75.
31. Aleksic SN, Budzilovich GN, Lieberman AN. Herpes zoster oticus and facial paralysis (Ramsay Hunt syndrome): clinico-pathologic study and review of literature. J Neurol Sci 1973;20:149–59.
32. Payten RJ, Dawes DDK. Herpes zoster of the head and neck. J Laryngol Otol 1972;86:1031–55.
33. Wackym PA, Popper P, Kerner MM, et al. Varicella-zoster DNA in temporal bones of patients with Ramsay Hunt syndrome. Lancet 1993;342:1555.
34. Meier JL, Straus SE. Comparative biology of latent varicella-zoster virus and herpes simplex virus infections. J Infect Dis 1992;166:S13–23.
35. Furuta Y, Takasu T, Fukuda S, et al. Detection of varicella-zoster virus DNA in human geniculate ganglia by polymerase chain reaction. J Infect Dis 1992;166: 1157–9.
36. Aviel A, Marshak G. Ramsay Hunt syndrome: a cranial polyneuropathy. Am J Otolaryngol 1982;3:61–6.

37. Schuknecht HF. Pathology of the ear. 2nd ed. Philadelphia: Lea & Febiger; 1993.
38. Gilbert GJ. Herpes zoster ophthalmicus and delayed contralateral hemiparesis: relationship of the syndrome to central nervous system granulomatous angiitis. JAMA 1974;229:302–4.
39. Pratesi R, Freemon FR, Lowry JL. Herpes zoster ophthalmicus with contralateral hemiplegia. Arch Neurol 1977;34:640–1.
40. Coker NJ. Facial electroneurography: analysis of techniques and correlation with degenerating motoneurons. Laryngoscope 1992;102:747–59.
41. Inamura H, Aoyagi M, Tojima H, et al. Effects of aciclovir in Ramsay Hunt syndrome. Acta Otolaryngol Suppl (Stockh) 1988;446:111–3.
42. Dickins JRE, Smith JT, Graham SS. Herpes zoster oticus: treatment with intravenous acyclovir. Laryngoscope 1988;98:776–9.
43. Stafford FW, Welch AR. The use of acyclovir in Ramsay Hunt syndrome. J Laryngol Otol 1986;100: 337–40.
44. Beutner KR, Friedman DJ, Forszpaniak C, et al. Valaciclovir compared with acyclovir for improved therapy for herpes zoster in immunocompetent adults. Antimicrob Agents Chemother 1995;39: 1546–53.
45. Harner SG, Heiny BA, Newell RC. Herpes zoster oticus. Arch Otolaryngol 1970;92:632–6.
46. Burgess RC, Michaels L, Bale JF, et al. Polymerase chain reaction amplification of herpes simplex viral DNA from the geniculate ganglion of a patient with Bell's palsy. Ann Otol Rhinol Laryngol 1994;103: 775–9.
47. Murakami S, Mizobuchi M, Nakashiro Y, et al. Bell's palsy and herpes simplex virus: identification of viral DNA in endoneural fluid and muscle. Ann Intern Med 1996;124:27–30.
48. Furuta Y, Takasu T, Sato KC, et al. Latent herpes simplex virus type 1 in human geniculate ganglia. Acta Neuropathol (Berl) 1992;84:39–44.
49. Bauer CA, Coker NJ. Update on facial nerve disorders. Otolaryngol Clin North Am 1996;29:445–54.
50. Adour KK, Byl FM, Hilsinger RL Jr, et al. The true nature of Bell's palsy: analysis of 1000 consecutive cases. Laryngoscope 1978;88:787–811.
51. Pitts DB, Adour KK, Hilsinger RL. Recurrent Bell's palsy: analysis of 140 patients. Laryngoscope 1988;98:535–40.
52. Fisch U. Surgery for Bell's palsy. Arch Otolaryngol 1981;107:1–11.
53. Adour KK, Ruboyianes JM, Von Doersten PG, et al. Bell's palsy treatment with acyclovir and prednisone compared with prednisone alone: a double-blinded, randomized controlled trial. Ann Otol Rhinol Laryngol 1996;105:371–8.
54. Dennis JM, Coker NJ. Electroneuronography. Adv Otorhinolaryngol 1997;53:112–31.
55. Gantz BJ, Rubinstein JT, Gidley P, Woodworth GG. Surgical management of Bell's palsy. Laryngoscope 1999;109:1177–88.
56. Gantz BJ, Wackym PA. Facial nerve abnormalities. In: Bumstead RM, Smith JD, editors. Pediatric facial plastic and reconstructive surgery. New York: Raven Press; 1993. p. 337–47.
57. Wackym PA, Andrews JC. Middle cranial fossa approach. In: Samii M, Cheatham M, Becker D, editors. Atlas of cranial base surgery. Philadelphia: WB Saunders; 1995. p. 26–31.
58. Fisch U. Prognostic value of electrical tests in acute facial paralysis. Am J Otol 1984;5:494–8.
59. Gantz BJ, Gmur A, Fisch U. Intraoperative evoked electromyography in Bell's palsy. Am J Otolaryngol 1982;3:273–8.
60. Bergman I, May M, Wessle HB, Stool SE. Management of facial palsy caused by birth trauma. Laryngoscope 1986;96:381–4.
61. Lipkin AF, Coker NJ, Jenkins HA, Alford BR. Intracranial and intratemporal facial neuroma. Otolaryngol Head Neck Surg 1987;96:71–9.
62. Wackym PA, King WA, Meyer GA, et al. Endoscope-assisted surgery of the trigeminal, facial, cochlear or vestibular nerve. In: Wackym PA, Rice DH, Schaefer SD, editors. Minimally invasive surgery of the head, neck, and cranial base. Philadelphia: Lippincott Williams & Wilkins; 2002. p. 101–16.
63. Schwaber MK. Vascular compression syndromes. In: Jackler RK, Brackmann DE, editors. Neurotology. St. Louis: Mosby-Year Book; 1994. p. 881–903.
64. Rohrer D, Burchiel K. Trigeminal neuralgia and other trigeminal dysfunction syndromes. In: Barrow D, editor. Surgery of the cranial nerves of the posterior fossa. Park Ridge (IL): American Association of Neurological Surgeons; 1993. p. 201–17.
65. Apfelbaum R. Surgery for tic douloureux. Clin Neurosurg 1984;31:346–50.
66. Burchiel K, Clarke H, Haglund M. Long-term efficacy of microvascular decompression in trigeminal neuralgia. J Neurosurg 1988;69:35–8.
67. Sweet W. Trigeminal neuralgia: problems as to cause and consequent conclusions regarding treat-

ment. In: Wilkins RH, Rengachary S, editors. Neurosurgery update II. New York: McGraw-Hill; 1991. p. 366–72.
68. Barker FG, Jannetta PJ, Bissonette DJ, et al. The long-term outcome of microvascular decompression for trigeminal neuralgia. N Engl J Med 1996;334: 1077–83.
69. Mendoza N, Illingworth RD. Trigeminal neuralgia treated by microvascular decompression: a long-term follow-up study. Br J Neurosurg 1995;9:13–9.
70. Kureshi SA, Wilkins RH. Posterior fossa reexploration for persistent or recurrent trigeminal neuralgia or hemifacial spasm: surgical findings and therapeutic implications. Neurosurgery 1998;43: 1111–7.
71. Wilkins RH. Cranial nerve dysfunction syndromes: evidence for microvascular compression. In: Barrow D, editor. Surgery of the cranial nerves of the posterior fossa. Park Ridge (IL): American Association of Neurological Surgeons; 1993. p. 155–63.
72. Liao JJ, Cheng WC, Chang CN, et al. Reoperation for recurrent trigeminal neuralgia after microvascular decompression. Surg Neurol 1997;47:562–70.
73. Magnan J, Caces F, Locatelli P, Chays A. Hemifacial spasm: endoscopic vascular decompression. Otolaryngol Head Neck Surg 1997;117:308–14.
74. Graham MD, Kemink JL. Total facial nerve decompression in recurrent facial paralysis and the Melkersson-Rosenthal syndrome: a preliminary report. Am J Otol 1986;7:34–7.
75. Eicher SA, Coker NJ, Alford BR, et al. A comparative study of the fallopian canal at the meatal foramen and labyrinthine segment in young children and adults. Arch Otolaryngol Head Neck Surg 1990; 116:1030–5.
76. Ge XX, Spector GJ. Labyrinthine segment and geniculate ganglion of the facial nerve in fetal and adult temporal bones. Ann Otol Rhino Laryngol 1981;90: 1–12.
77. Jackson CG, Johnson GD, Hyams VJ, Poe DS. Pathologic findings in the labyrinthine segment of the facial nerve in a case of facial paralysis. Ann Otol Rhinol Laryngol 1990;99:327–9.
78. Sunderland S. Nerves and nerve injuries. 2nd ed. New York: Churchill Livingstone; 1978. p. 258–9.
79. Seddon HJ. Three types of nerve injury. Brain 1943;66:237–88.
80. House JW, Brackmann DE. Facial nerve grading system. Otolaryngol Head Neck Surg 1985;93: 146–7.
81. Gidley PW, Gantz BJ, Rubinstein JT. Facial nerve grafts: from cerebellopontine angle and beyond. Am J Otol 1999;20:781–8.
82. Gantz BJ. Intraoperative facial nerve monitoring. Am J Otol 1985;6 Suppl:58–61.
83. House W. Surgical exposure of the internal auditory canal and its contents through the middle cranial fossa. Laryngoscope 1961;71:1363–85.
84. Yamamoto E, Fisch U. Experiments on facial nerve suturing. ORL J Otorhinolaryngol Relat Spec 1974;36:193–204.
85. Rhee J, Gallo JF, Costantino PD. Endoscopic facial rejuvenation. In: Wackym PA, Rice DH, Schaefer SD, editors. Minimally invasive surgery of the head, neck, and cranial base. Philadelphia: Lippincott Williams & Wilkins; 2002. p. 355–66.
86. May M. Gold weight and wire spring implants as alternatives to tarsorraphy. Arch Otolaryngol Head Neck 1987;113:656–60.
87. Kelly SA, Sharpe DT. Gold eyelid weights in patients with facial palsy: a patient review. Plast Reconstr Surg 1992;89:436–40.
88. Samii M. Rehabilitation of the face by facial nerve substitution. In: Fisch U, editor. Facial nerve surgery. Birmingham (AL): Aesculapius Publishing Co.; 1977. p. 224–45.
89. Conley J, Baker D. Hypoglossal-facial nerve anastomosis for reinnervation of the paralyzed face. Plast Reconstr Surg 1979;63:63–72.
90. Pensak ML, Jackson GG, Glasscock ME, Gulya AJ. Facial reanimation with the XII-VII anastomosis: analysis of the functional and psychological results. Otolaryngol Head Neck Surg 1986;94: 305–10.
91. May M, Sobel SM, Mester SJ. Hypoglossal-facial nerve interpositional jump graft for facial reanimation without tongue atrophy. Otolaryngol Head Neck Surg 1991;204:818–26.
92. Hammerschlag PE. Facial reanimation with jump interpositional graft hypoglossal facial anastomosis and hypoglossal facial anastomosis: evolution in management of facial paralysis. Laryngoscope 1999;109 Suppl:1–23.
93. Kartush JM, Lundy LB. Facial nerve outcome in acoustic neuroma surgery. Otolaryngol Clin North Am 1992;25:623–47.
94. May M, Drucker C. Temporalis muscle for facial reanimation: a 13-year experience with 224 procedures. Arch Otolaryngol Head Neck Surg 1993; 119:378–82.

95. Fisher E, Froedel JL. Facial suspension with acellular human dermal allograft. Arch Facial Plast Surg 1999;1:195–9.
96. Konior RJ. Facial paralysis reconstruction with Gore-Tex Soft Tissue Patch. Arch Otolaryngol Head Neck Surg 1992;118:1188–94.
97. Ueda K, Harii K, Yaruada A. Long-term follow-up of nerve conduction velocity in cross-face nerve grafting for the treatment of facial paralysis. Plast Reconstr Surg 1994;93:1146–9.
98. Ueda K, Harii K, Yaruada A. Free neurovascular muscle transposition for the treatment of facial paralysis using the hypoglossal nerve as a recipient motor source. Plast Reconstr Surg 1994;94:808–17.
99. Balliet R, Shinn J, Bach-y-Rita P. Facial paralysis rehabilitation. Re-training selective muscle control. Int Rehabil Med 1982;4:67–74.
100. Segal B, Zompa I, Danys I, et al. Symmetry and synkinesis during rehabilitation of unilateral facial paralysis. J Otolaryngol 1995;24:143–8.
101. Brach JS, VanSwearingen JM, Lenert J, et al. Facial neuromuscular retraining for oral synkinesis. Plast Reconstr Surg 1997;99:1922–33.
102. May M, Croxson GR, Klein SR. Bell's palsy: management of sequelae using EMG rehabilitation, botulinum toxin, and surgery. Am J Otol 1989;10: 220–9.

CHAPTER 25

Surgery of the Skull Base

Richard J. Wiet, MD, Richard Hoistad, MD

The skull base is a complex anatomic region that can be the site of a diverse group of pathologic processes. Surgical access to this area is necessary for a variety of neoplasms as well as nontumorous conditions such as vertigo, facial paralysis, and hemifacial spasm.

Treatment of skull base disorders can be quite complex, but recent advances allow more accurate assessment and more reliable outcomes. Improved imaging techniques (especially magnetic resonance imaging [MRI]) allow more accurate planning of surgery and counseling of the patient. Diagnostic techniques such as auditory brainstem responses (ABRs) and electronystagmography (ENG) allow early identification of abnormalities. Intraoperative neural monitoring allows quicker, more reliable localization of important structures during surgery. As more experience is reported in this area, ideas and treatments are being refined and improved.

An important aspect of skull base surgery, and a large factor for optimal outcome, is cooperation among physicians. Consultation with a neuroradiologist maximizes the benefit of imaging techniques and allows interventional procedures such as embolization. Neurosurgical assistance is desirable when intracranial exposure is necessary. In selected cases, participation of other physicians such as vascular surgeons may also be beneficial.

Recent advances in technology, including image guidance systems and endoscopes, are being applied to skull base surgical procedures.

Image guidance systems consist of either an electromagnetic tracking system or an optical infrared tracking system.[1,2] In the field of otolaryngology, the more common application has been in sinus surgery, but recent interest has arisen for application to skull base procedures. Image guidance systems are particularly appealing for middle fossa approaches to skull base lesions.[3] Both systems have been found to be reliable and accurate within 2 mm for anatomic localization.

Adjunctive use of endoscopy for posterior fossa surgery and acoustic neuroma surgery has also been found to be useful.[4] Endoscopy provides better visualization than the microscope in certain areas of the skull base, particularly in identifying neurovascular structures, inspecting acoustic neuroma tumor removal along the internal auditory canal (IAC), and identifying open petrous air cells, thus lowering the postoperative cerebrospinal fluid (CSF) leak rate.[5,6]

ANATOMY

Superiorly, the floor of the cranial cavity, or skull base, serves as a support for the brain and intracranial structures. Inferiorly, it attaches to the facial bones anteriorly and serves as the attachment for the head and neck muscles laterally and posteriorly. The floor of the cranial cavity is composed of the frontal, ethmoid, sphenoid, temporal, and occipital bones (Figure 25–1) and is broadly classified into the anterior, middle, and posterior fossae. Numerous foramina allow passage of vascular and neural structures between the intracranial and extracranial space. The area below the midlateral portion of the skull base (essentially the temporal bone) is designated the infratemporal fossa.

Anterior Cranial Fossa

The anterior fossa is composed of the frontal bone anteriorly and laterally, the cribriform plate of the ethmoid bone centrally, and the body and lesser wing of the sphenoid bone posteriorly and centrally. Immediately anterior to the cribriform plate is the foramen cecum, through which an emissary vein

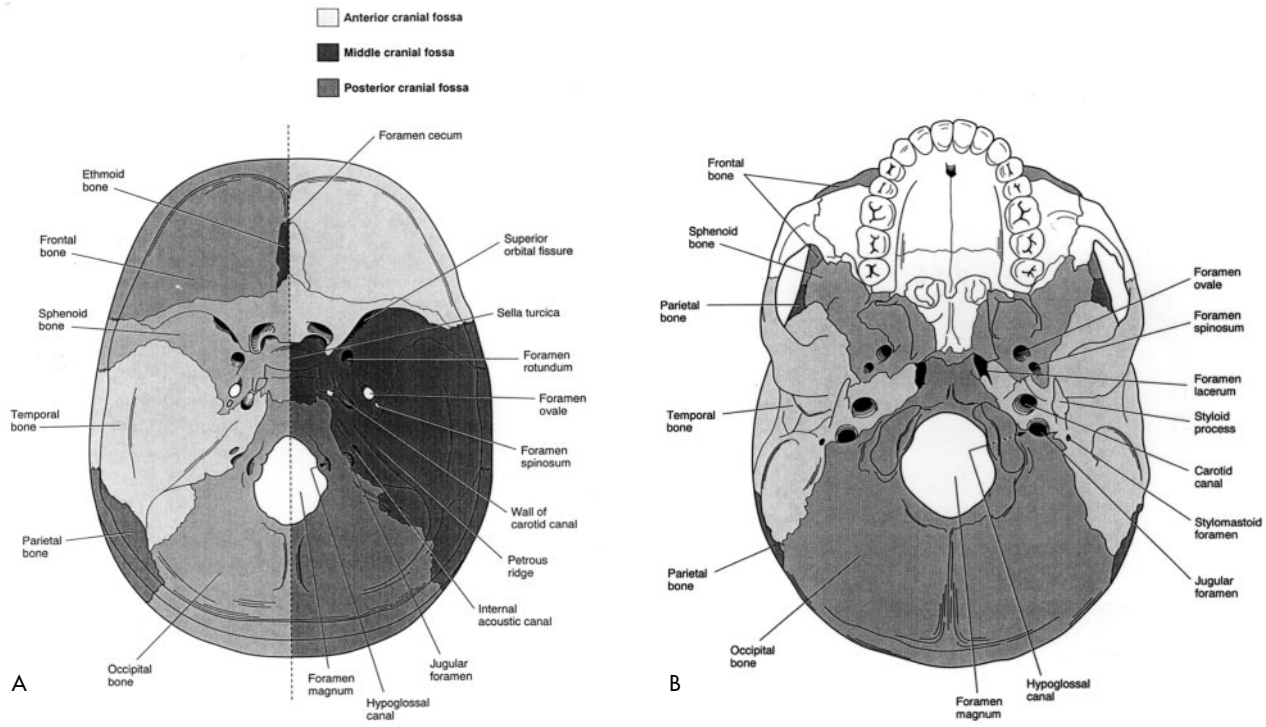


FIGURE 25–1. *A*, View of the skull base from above shows the foramina discussed in the text (*right*), the three divisions of the skull base (*right*), and the bones comprising this area (*left*). *B*, View of the skull base from below with the foramina labeled on the right side of the figure and individual bones comprising the skull base on the left.

passes to the nasal cavity. The perforated cribriform plate allows passage of the olfactory nerve fibers traveling from the olfactory epithelium of the superior part of the nasal cavity to the olfactory bulb. The sella turcica contains the pituitary gland in the posterior central area of the anterior fossa at the junction with the middle and posterior fossae. The optic chiasm rests anterior and superior to the pituitary in the chiasmic groove. The cavernous sinus lies lateral to the pituitary gland and sella turcica.

Middle Cranial Fossa

The middle fossa is formed by the greater wing of the sphenoid bone anterocentrally and the temporal bone centrally, laterally, and posteriorly. It extends from the superior orbital fissure (greater wing of the sphenoid) anteriorly to the superior petrosal sinus along the petrous ridge posteriorly and from the cavernous sinus medially to the squamous portion of the temporal bone laterally. Anteromedially, the foramen spinosum allows passage of the middle meningeal artery into the cranial space, where it courses laterally along the outer surface of the dura. The trigeminal nerve originates in the posterior cranial fossa and enters the middle cranial fossa through a fold of dura over the medial part of the petrous ridge, Meckel's cave. The fifth cranial nerve (gasserian) ganglion lies immediately anterior to this and rests over the horizontal portion of the carotid artery. From the gasserian ganglion, the three branches arise. The mandibular division of the trigeminal nerve exits the cranial cavity anteromedial to the foramen spinosum through the foramen ovale. Further anteromedially, the foramen rotundum allows for the exit of the maxillary branch of the trigeminal nerve in an anterior and horizontal direction to the pterygopalatine space. The ophthalmic branch of the trigeminal nerve, along with cranial nerves III, IV, and VI and the ophthalmic vein, pass through the cavernous sinus and then enter the orbit through the superior orbital fissure. After entering the central portion of the temporal bone from the neck, the carotid artery curves medially 90 degrees and runs a near-horizontal course in a medial direction in the foramen lacerum before entering the intracranial space and passing through the cavernous sinus.

The middle cranial fossa is the one portion of the skull base with significant structures contained within the underlying bone, namely, the seventh cranial nerve and the labyrinth. Their anatomy is presented in Chapter 1. Access to and around these structures is a frequent issue in skull base surgery.

Posterior Cranial Fossa

The majority of the floor of the posterior fossa is composed of the occipital bone, whereas the posterior face of the petrous portion of the temporal bone defines the anterior lateral extent. Centrally, the foramen magnum allows passage of the spinal cord and the vertebral arteries. Immediately lateral to the foramen magnum, each hypoglossal canal allows passage of a hypoglossal nerve from the intracranial cavity. Lateral to this, the jugular foramen lies at the junction of the occipital bone (medial posterior margin) and the temporal bone (anterolateral margin). The jugular vein (laterally) and the ninth, tenth, and eleventh cranial nerves (medially) traverse this foramen.

Superior to the jugular foramen, the IAC serves as an entrance of the seventh and eighth cranial nerves into the temporal bone. In the cerebellopontine angle (CPA), the seventh nerve lies anteromedial to the eighth nerve. As the eighth nerve enters the IAC, it trifurcates into the auditory, superior vestibular, and inferior vestibular nerves. In the lateral end of the IAC (fundus), these three nerves, along with the seventh nerve, fill each of the quadrants of the IAC (Figure 25–2).

Infratemporal Fossa

The area below the temporal bone is designated the infratemporal fossa. The internal carotid artery and the internal jugular vein pass through here in a vertical direction.

Medial to the internal jugular vein, cranial nerves IX, X, and XI exit the jugular foramen and pass through the infratemporal fossa. The deep lobe of the parotid abuts the anterolateral margin of the infratemporal fossa.

EVALUATION

History and Physical Examination

Key aspects of history taking and physical examination as they relate to skull base disorders are highlighted. A past medical history should always be taken to include use of medications, allergies, alcohol and tobacco use, and any past history of malignancy or neurologic disorder, such as multiple sclerosis. It is very important to rule out the possibility of metastasic neoplasms or systemic illness.

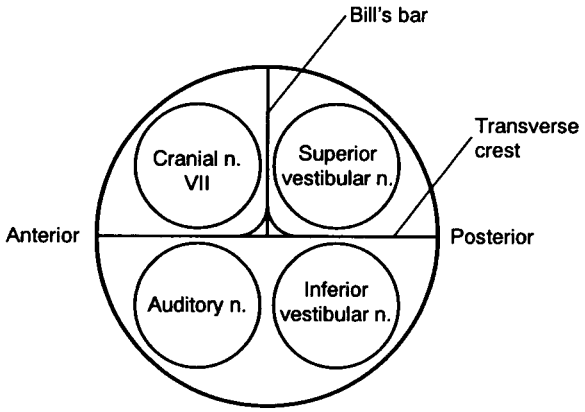


FIGURE 25–2. Schematic drawing of the position of the four nerves at the lateral aspect of the internal auditory canal. The transverse crest and Bill's bar are important identifying landmarks.

Although some skull base lesions may be an incidental finding (eg, on MRI), most patients will be symptomatic. Their complaints can range from nonspecific symptoms, such as headache, to very localized findings, such as unilateral hearing loss. In directing questions to the patient, one should remember that most symptoms are related to encroachment on nervous and vascular structures along the skull base.

Evaluation of a patient with a skull base disorder requires a thorough head and neck examination, otologic/neurotologic examination, and general neurologic examination with emphasis on the cranial nerves. Specific signs symptoms and findings will be discussed with each disorder.

Diagnostic Testing

Physiologic tests are also useful in diagnosis and treatment of skull base disorders.

Patients with audiovestibular complaints or suspected cerebellopontine disorders undergo basic audiologic evaluation including pure-tone averages, speech reception thresholds, word recognition scores, and tympanometry. These tests supply valuable information concerning the function of the

external auditory canal, tympanic membrane, middle ear and ossicles, cochlea, and auditory nerve. If a retrocochlear lesion is suspected, stapedial reflexes are obtained and rollover (decreasing discrimination with increasing intensity) is evaluated (Table 25–1). If there is suspicion of a retrocochlear lesion after initial evaluation, further testing (either ABR or MRI scan) is done (Figure 25–3).

Auditory brainstem response testing is a very useful, noninvasive screening test used most commonly when an acoustic neuroma is suspected. Although absolute latencies of the five obtained waves are helpful, the most commonly used measurement in evaluation of an acoustic neuroma is the interaural comparison of wave V latencies. A difference of greater than 0.2 ms is considered significant. The sensitivity of ABR, using a click stimulus, is reported between 93 and 100% [7]

False positives range from 8 to 25%.[8] Care must be used in evaluating a patient with pure-tone thresholds greater than 65 dB at 4 kHz as an abnormal response may be attributable simply to the inability to activate the auditory system adequately.

The ABR is also useful in the identification of other disorders. In seventh nerve vascular compressive syndromes, the wave I–III latencies tend to be prolonged. This is felt to be caused by the irritation/compression from the pulsation of the blood vessel against the nerve. Abnormalities can also be seen in other disorders, such as multiple sclerosis. Up to 93% of patients with multiple sclerosis have ABR abnormalities.[9] Various wave amplitude and latency abnormalities are possible, with wave V amplitude and wave II–V latencies being the most common.

The ENG is used to evaluate the vestibular labyrinth, the vestibular portion of the eighth nerve, and the central portions of the vestibular system.

TABLE 25–1. Audiologic Evidence of Retrocochlear Lesions

Asymmetric sensorineural hearing loss
Stapedial (acoustic) reflex abnormalities
Tone decay
Rollover in discrimination scores
Decreased discrimination out of proportion to hearing loss

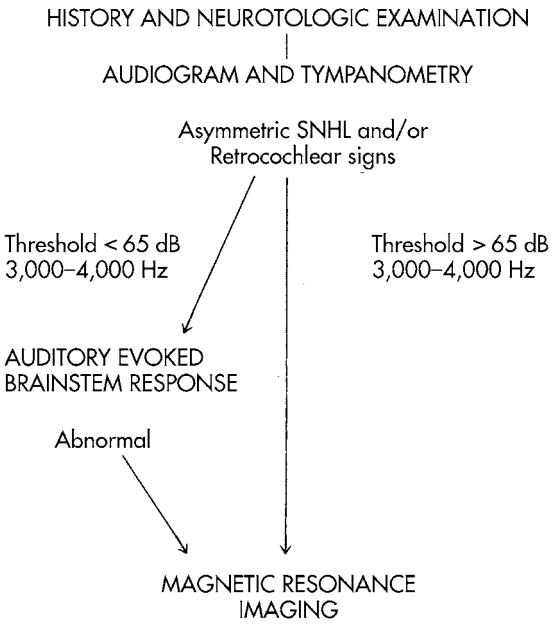


FIGURE 25–3. Flow diagram demonstrating our procedure for evaluation of patients suspected of having retrocochlear disease. SNHL = sensorineural hearing loss.

The caloric portion of the ENG allows each labyrinth to be evaluated individually. The information is very helpful when attempting to document objectively the side of vestibular abnormality when considering an ablative vestibular neurectomy.

Seventh cranial nerve status is frequently important in skull base disorder decision making. Most commonly, the level of clinical function is the important feature. At times, in the completely paralyzed face, it is important to evaluate the integrity of the nerve. This provides useful prognostic information and is also helpful in Bell's palsy or herpes zoster oticus (Ramsay Hunt syndrome) when surgical decompression of the seventh nerve is being considered. Several tests are available including electroneuronography (evoked electromyography), maximal stimulation test, nerve excitability test, and electromyonography. (The descriptions, indications, and details of each of these tests are discussed in Chapter 24.)

Diagnostic Imaging

Imaging is frequently the key diagnostic method in skull base abnormalities. In mass lesions, it is especially critical in treatment planning.

Magnetic resonance imaging came into regular use in the late 1980s. Soft tissue detail is significantly improved from previous techniques. Gadolinium diethylenetriamine pentaacetic acid (DTPA), a paramagnetic contrast agent, allows definition of many tissues and lesions. It is given intravenously, and, unlike iodine compounds, adverse reactions to it are extremely rare. In the CPA, MRI has allowed identification of smaller acoustic neuromas than in the past. Even very small acoustic neuromas that do not show significant tissue distortion readily enhance with gadolinium (Figure 25–4, A and B). One distinct advantage of MRI is that, in one "pass" through the scanner, complete information is entered into the computer and an image can be created in any plane, most commonly the axial, coronal, or sagittal planes. A drawback of MRI is the complete lack of visualization of bone, making it difficult to evaluate any bony abnormality or soft tissue– bony interfaces. Magnetic resonance imaging is contraindicated in patients with metal implants such as pacemakers and surgical clips. Claustrophobia and obesity can be relative contraindications, although newer "open" scanners frequently allow enough room for obese patients and help alleviate claustrophobic symptoms.

Computed tomography (CT) became the prevalent imaging tool in the 1970s, and although it allowed the best soft tissue detail at the time, it has subsequently been surpassed by MRI. High-resolution CT scanning of the temporal bone and skull base remains the best study of bony anatomy and of bone destruction by space-occupying lesions (Figure 25–5). Intravenous iodine-based contrast can be used to enhance soft tissue visualization.

Angiography can be useful as both a diagnostic and a treatment modality. Primary vascular abnormalities such as aneurysms, arteriovenous malformations, and arteriovenous fistulae are readily seen on angiography. Vascular tumors such as glomus jugulare (Figure 25–6) and glomus vagale can be clearly identified, with their blood supplies delineated. Angiography is also useful to establish the patency, location, and relationship of the blood vessels adjacent to the lesions.

In lesions that involve or are adjacent to the internal carotid artery, preoperative occlusion studies are of predictive value when considering resection of the artery. After catheterization of the internal carotid artery, the vessel is balloon occluded, and the patients are monitored. Patients tolerating occlusion are at lower risk for neurogenic sequelae postoperatively if the artery is resected or damaged during surgery.[10]

In vascular lesions such as glomus jugulare tumors, angiography with embolization has been used as an isolated treatment modality but more

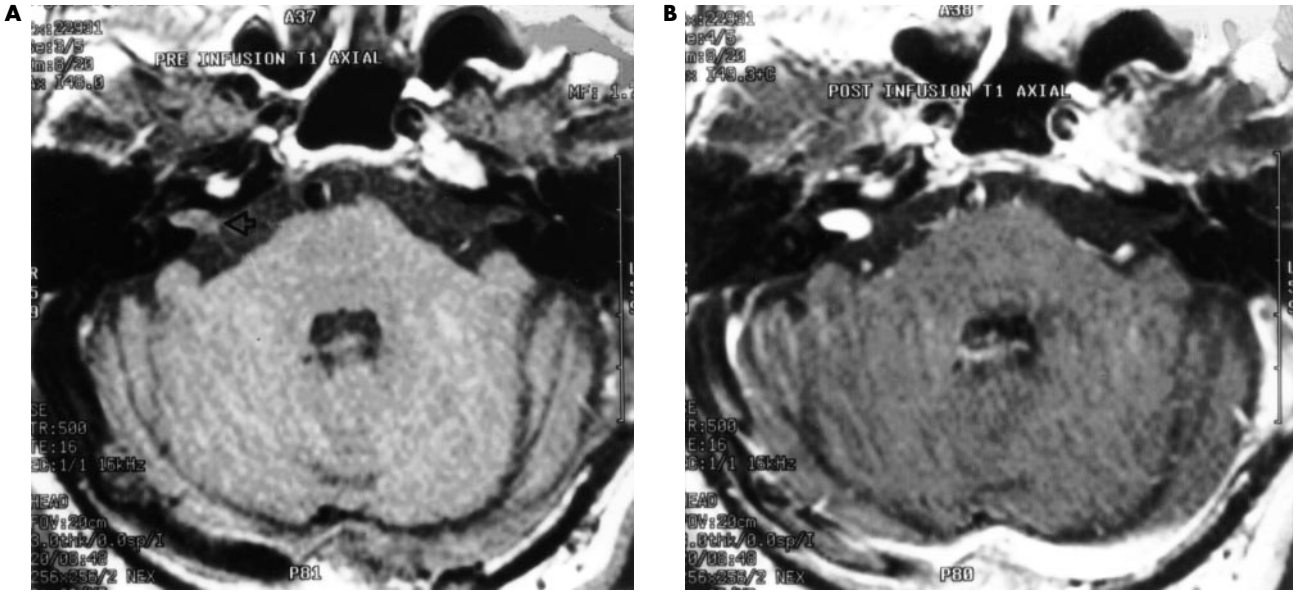


FIGURE 25–4. Magnetic resonance image of a patient with a small right intracanalicular acoustic neuroma. *A*, The T_1 image shows subtle contrast changes without significant internal auditory canal enlargement (*arrow*). *B*, Gadolinium enhancement of the T1 image clearly delineates the tumor.

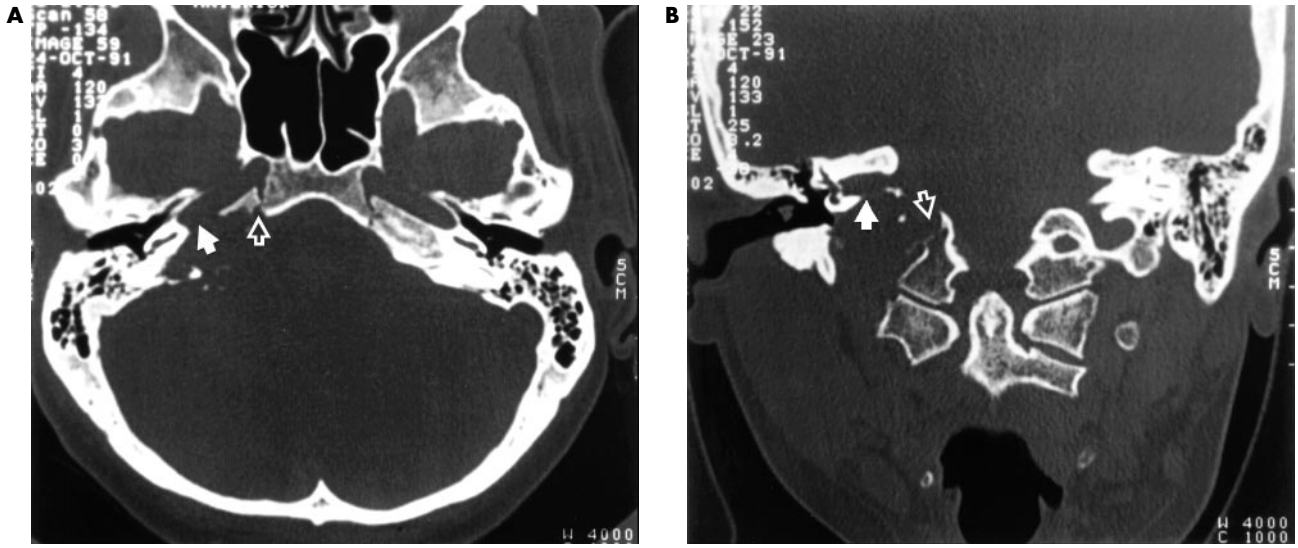


FIGURE 25–5. Computed tomographic scan of the temporal bone in which a chondrosarcoma arises from the right foramen lacerum. There is significant bone erosion of the petrous apex. *A*, Axial view demonstrates erosion of bone along the horizontal portion of the carotid canal (*black arrow*) and the clivus portion of the sphenoid bone (*open arrow*). *B*, Coronal view reveals bone erosion of the inferior aspect of the canal of the internal carotid artery (*black arrow*) and hypoglossal canal (*open arrow*).

commonly is used as a preoperative adjunct. Preoperative embolization has been shown to decrease significantly intraoperative blood loss in these patients.[11,12]

SKULL BASE DISORDERS

There are a large variety of skull base disorders (see Table 25–2 for differential diagnosis of skull base lesions). Although many of these are mass lesions, other entities, such as infectious diseases, inflammatory abnormalities, and vascular compression syndromes, must be considered. Although malignant lesions do occur, fortunately, most of the skull base entities are benign. When evaluating a patient with a malignant skull base lesion, it is important to rule out the possibility of a metastasis from a regional or distant site.

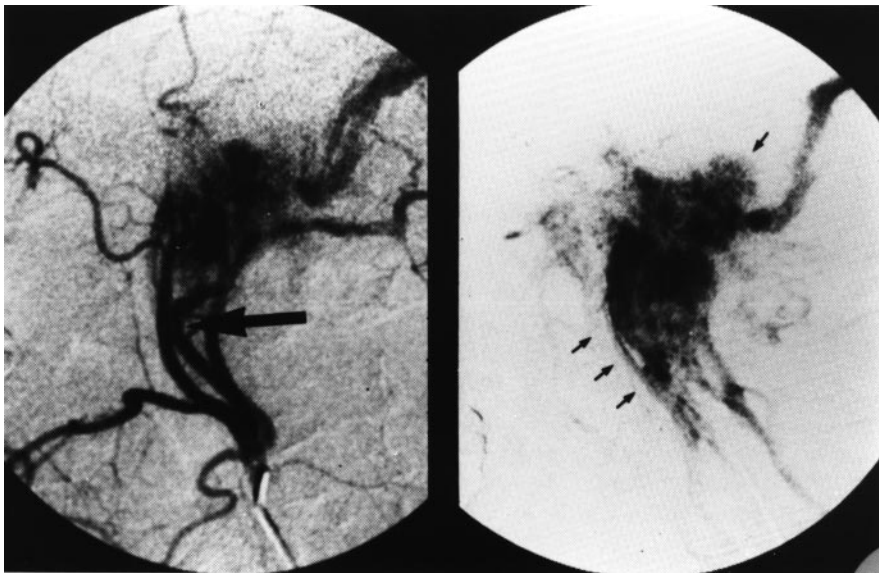


FIGURE 25–6. Angiogram showing a large glomus jugulare tumor. The ascending pharyngeal artery (*large arrow*) is the dominant feeding vessel. Tumor fills the jugular vein (*triple arrow*) and extends into the posterior fossa (*small single arrow*). Reproduced with permission from Young N et al.[11]

TABLE 25–2. Differential Diagnosis of Skull Base Lesions

Glomus tumors
Meningiomas
Neural tumors
Neurilemmoma (acoustic neuroma, facial nerve schwannoma)
Neurofibroma
Chordoma
Carcinomas
Primary
Squamous cell carcinoma
Basal cell carcinoma
Metastatic
Squamous cell carcinoma (upper aerodigestive tract, lung)
Adenocarcinoma (kidney, breast, lung, salivary gland)
Nasopharyngeal
Primary cholesteatomas
Rare lesions
Rhabdomyosarcoma
Plasmacytoma
Melanoma
Giant cell tumor
Osteoblastoma
Lipoma
Chondrosarcoma
Leukemia
Eosinophilic granuloma
Dermoid/teratoma
Histiocytosis
Hemangiopericytoma
Ewing's sarcoma
Petrositis

Some mass lesions may be unique to specific areas of the skull base, whereas others can arise in more than one site. The following section will briefly review lesions by location. The more important entities will subsequently be discussed individually.

Lesions of the Anterior Part of the Skull Base

Mass lesions of the anterior part of the skull base include benign fibro-osseous lesions and osseous lesions as well as malignant entities such as sinonasal tumors and esthesioneuroblastoma.

Fibro-osseous lesions, fibrous dysplasia and ossifying fibroma, are common benign lesions of the anterior part of the skull base. Fibrous dysplasia may occur in monostotic (one site) or polyostotic (more than one site) forms. The monostotic form is the most common. Ossifying fibromas have a higher growth rate than fibrous dysplasia. They are differentiated on CT from fibrous dysplasia by the appearance of a bony capsule. Another common lesion is an osseous lesion, an osteoma. Osteomas frequently arise in the frontal and ethmoid sinuses and involve the anterior part of the skull base secondarily. Pathologic differentiation among these lesions is sometimes difficult. Treatment of these lesions is necessary only when functional problems arise or cosmetic deformity occurs.[13]

Two malignant lesions that can involve the anterior part of the skull base are paranasal sinus carcinoma and esthesioneuroblastoma. Ethmoid and, less commonly, frontal sinus tumors (most of which are squamous cell carcinomas) involve the anterior part of the skull base through direct extension. Wide surgical removal, usually in combination with radiation or chemotherapy, is necessary when treatment for cure is planned. Esthesioneuroblastomas are of neural crest origin and represent about 2% of all malignant nasal neoplasms. The cribriform plate is involved early as this lesion arises from the olfactory epithelium.[14] Total gross removal may require craniofacial resection as well as orbital exenteration in addition to sinus eradication.

Middle Fossa/Petrous Apex Lesions

Mass lesions of the middle cranial fossa and petrous apex include benign entities such as cholesterol granuloma, meningiomas, cholesteatomas, and schwannomas. The latter three lesions are discussed in more detail later in this chapter. Malignant lesions such as chordoma, chondrosarcoma, and nasopharyngeal carcinoma can also affect this area.

Cholesterol granulomas are fluid-filled cysts that may arise in any portion of the air cell system of the temporal bone, more commonly in the petrous apex. They occur in well-pneumatized temporal bones in air cells that are poorly ventilated. The negative pressure causes a hemorrhage in the air cell. Once the blood begins to break down, there is a for-

eign body reaction to one of the breakdown components, cholesterol.[15] Cholesterol granulomas have a propensity to affect cranial nerves. The patient usually presents with hearing loss, although vertigo, facial paresis, and diplopia may be present. Diagnosis is usually confirmed with CT and MRI (Figure 25–7). This lesion appears very bright on both T_1- and T_2-weighted images. Treatment consists of surgical drainage with maintenance of ventilation.[16]

Of the potentially malignant lesions, chordoma is relatively insidious. This neoplasm is thought to arise from the primitive notochord, and the majority involve the petroclival area. Symptoms are caused by progressive involvement of adjacent cranial nerves. The typical CT shows an irregular mass with bony destruction. Magnetic resonance imaging reveals a low-intensity mass on T_1 and T_2, with variable enhancement. Extensive surgery in combination with high-dose radiation therapy is the treatment of choice.[17]

Chondrosarcomas are rare, comprising less than 1% of skull base tumors.[18] They are thought to arise from trapped embryologic rests in and around the foramen lacerum. Hearing loss is the initial complaint of most patients, although complaints may include tinnitus, vertigo, unsteadiness, diplopia, and facial paresis. Computed tomography shows irregular bony destruction with the contrast enhancement (see Figure 25–5). The lesion also brightly enhances on MRI. Total removal of the tumor is the optimum treatment. Radiation therapy is controversial.

Nasopharyngeal cancer is discussed in Chapter 60. Larger lesions can invade the skull base and present with symptoms of a skull base lesion.

Posterior Fossa/Internal Auditory Canal Lesions

Most mass lesions of the posterior fossa are benign. Acoustic neuromas account for roughly 90% of the lesions, whereas meningiomas account for about 5%. These will be discussed later. Other mass lesions include epidermoids, lipomas, and arachnoid cysts. The majority of the malignant lesions that do occur in this area are metastatic. Epidermoids arise from trapped rests of keratinized squamous epithelium and comprise 1% of CPA tumors.[19] Careful removal of all squamous cell elements is necessary to prevent aseptic meningitis or recurrence.

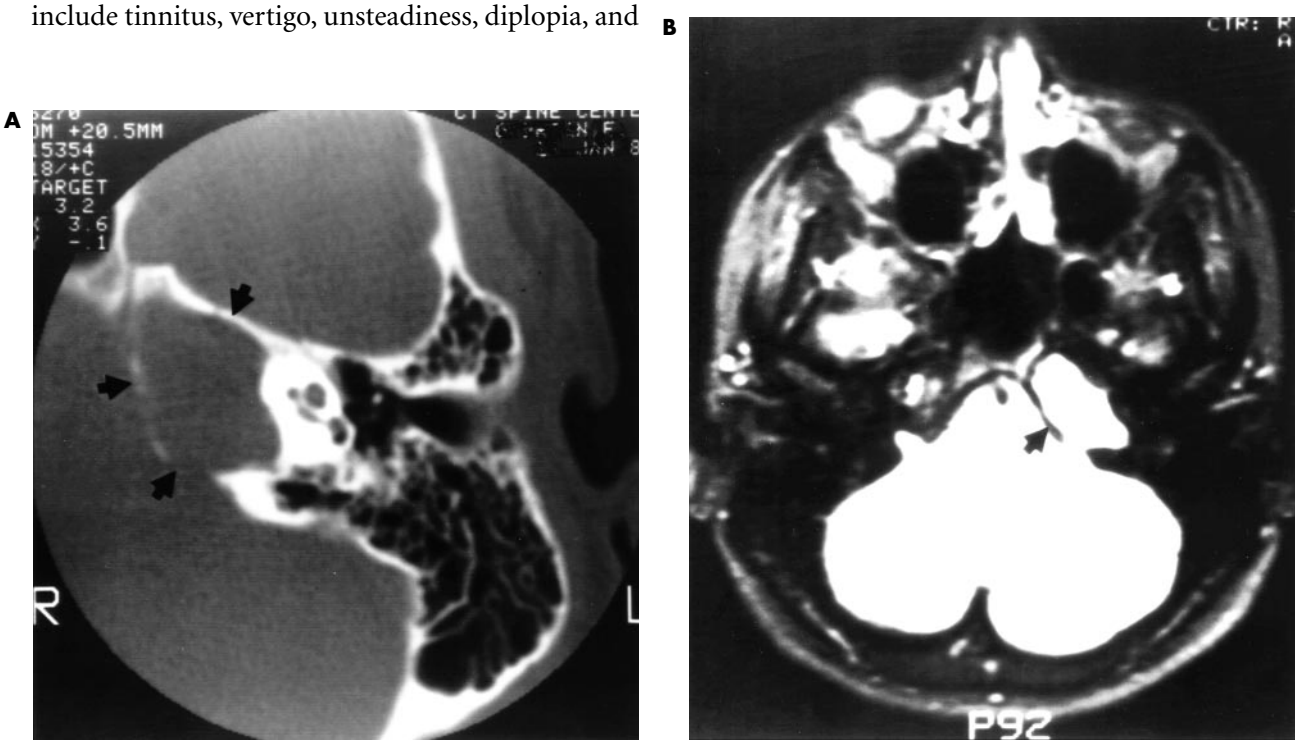


FIGURE 25–7. Patient with a large cholesterol granuloma of the left petrous apex. *A*, Bone window axial computed tomographic scan shows soft tissue density replacing the entire left petrous apex (*arrows*). *B*, T_1-weighted magnetic resonance image shows this bright lesion (*arrow*).

Lipomas can occur in the IAC and CPA. Internal auditory canal lipomas tend to present earlier because of neural compression, whereas CPA lipomas may not cause symptoms until they are much larger. Diagnosis is made with MRI.[20] Arachnoid cysts, which are collections of CSF trapped by arachnoid granulations, comprise about 1% of CPA lesions. They are very frequently asymptomatic and are incidental findings on MRI (Figure 25–8). If symptomatic, surgical treatment involves opening of the cyst wall.[21]

Metastatic lesions from any other location to the CPA is possible. Fortunately, this occurrence is rare. Tumors most likely to metastasize are breast tumors, adenocarcinomas, prostrate cancers, and clear cell tumors of the genital organs.[22]

SKULL BASE DISORDERS BY DIAGNOSIS

Neuroma/Schwannoma

Acoustic Neuroma Acoustic neuromas are the most common lesion of the CPA, representing 90% of all tumors in this area. Acoustic neuromas constitute about 6% of all intracranial tumors. In the United States, there is an incidence of approximately 1 in 100,000, resulting in roughly 2,500 new cases diagnosed each year. The name is a misnomer because they actually arise from the Schwann cells surrounding the vestibular portion of the eighth cranial nerve. Since Schwann cells are peripheral nerve entities, then, by definition, these must arise in the IAC.

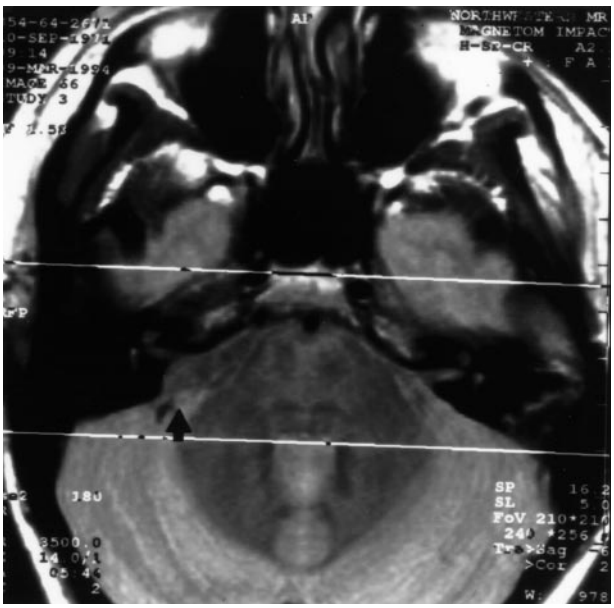


FIGURE 25–8. T_1-weighted magnetic resonance image shows an apparent soft tissue mass (*arrow*) in the cerebellopontine angle in a patient who presented with headaches and vertigo. At surgical resection, an arachnoid cyst was identified.

Acoustic neuromas occur in two forms: sporadic and those associated with neurofibromatosis 2. The sporadic form constitutes well over 90% of all acoustic neuromas. Neurofibromatosis 2 is a rare autosomal dominant genetic disorder whose gene is located on chromosome 22, resulting in bilateral acoustic neuromas and other central nervous system neoplasms.

Unilateral hearing loss is the most common presenting symptom. Other symptoms that may occur are tinnitus, dysequilibrium, vertigo, headache, and aural fullness. Later symptoms can include facial numbness, facial weakness, and diplopia. When the tumors are very large, neuropathies of cranial nerves III, IV, VI, IX, X, and XI, as well as cerebellar and brainstem compression, can occur.

Acoustic neuromas are relatively slow-growing tumors. Grossly, they appear yellow to gray and may have a cystic structure. Histopathologically, there are morphologic patterns: Antoni A and Antoni B. Antoni Type A is represented by small, densely packed spindle-shaped cells with dark staining nuclei. In Antoni A, the cells may have a whirled appearance referred to as Verocay bodies. They have no clinical significance. Antoni Type B refers to loosely packed vacuolated cells. Nuclei appear large and atypical and are densely staining. This is the predominant form in large tumors. There may also be a mixed pattern. The pattern does not correlate with the clinical presentation. Staining for S-100 immunoperoxidase is confirmatory for a Schwann cell.

The introduction of MRI in the 1980s has revolutionized the diagnosis of acoustic neuromas. Prior to MRI, CT was employed. Computed tomography was a valuable diagnostic tool if the tumor extended into the CPA or there was widening of the porus acusticus by the tumor. However, smaller intracanalicular tumors had to be diagnosed by air-contrast CT, requiring a lumbar puncture and injecting air in the subarachnoid space. Magnetic resonance imaging has made this procedure obsolete. Acoustic tumors readily enhance with the injection of gadolinium (see Figure 25–4). Tumors as small as 2 to 3 mm

are easily seen. This allows for the diagnosis of acoustic neuromas prior to the onset of severe clinical symptoms. Also, asymptomatic lesions that are contained in the IAC can be readily followed.

Surgical removal is the treatment of choice when the tumor is symptomatic or begins to extend beyond the IAC.[23] Total removal with the preservation of the facial nerve should be attempted. The choice of procedure, translabyrinthine, suboccipital, or middle fossa, depends on the size of the tumor and the status of the hearing. These approaches will be discussed later in this chapter. In very large tumors and in high-risk patients, there is a role for subtotal removal of the tumor to relieve symptoms.

Recently, stereotactic radiation or gamma knife has been advocated for treatment of acoustic neuromas. This involves a one-time dose of radiation focused on the center of the mass. The usual dose is 10 to 25 Gy. This causes shrinkage and necrosis of the mass. It has been employed in the elderly and other high-risk patients. The drawbacks are lack of elimination of the tumor and complications, including hydrocephalus and cranial nerve neuropathies. For further elaboration, see Chapter 20.

Facial Nerve Neuroma Facial nerve neuromas are one of the few primary lesions of the nerve. Initial symptoms may include facial paresis or facial muscular fasciculations. They are found primarily in the intratemporal portion of the nerve, with the geniculate ganglion area being the most common. Magnetic resonance imaging is helpful in the workup of patients with this lesion. These lesions are usually removed when facial weakness is present. Patients are usually reluctant to have them removed when they still have facial function, especially since the result will be a temporary total paralysis. These lesions are usually slow growing and can be watched for a period of time. Treatment is surgical removal with nerve grafting. For further elaboration, see Chapter 24.

Other Neuromas Schwannomas of the other cranial nerves are relatively unusual, but fifth and ninth, eleventh, and twelfthth nerve tumors must occasionally be treated. The petrous apex (especially with the fifth nerve) and the infratemporal fossa can be involved. Symptoms are generally related to the involved nerve but can be caused by compression of adjacent structures. Trigeminal and vagal schwannomas are more common, whereas ninth, eleventh, and twelfth cranial nerve schwannomas are rare. They are diagnosed with contrast MRI, whereas CT can reveal expansion of the bone. Surgical removal is the treatment of choice for symptomatic lesions.

Meningiomas

Meningiomas are the most common benign intracranial tumors and can be found in all areas of the skull base. In the CPA, they comprise 5% of all tumors. They are thought to arise from the arachnoid villi. Unlike acoustic neuromas, they do not originate in the IAC but arise broadly from the petrous face. At presentation, meningiomas are usually larger than acoustic neuromas, probably because they do not compress the contents of the IAC. Symptoms are related to the impingement on surrounding structures; however, compared with acoustic neuromas, they more commonly spare hearing. The gross appearance of a meningioma is a yellow to gray globular mass.

Microscopically, it is characterized by sheets of polygonal cells. The boundaries of the cell are obscure. Nuclei contain chromatin and are pale staining. Histologically, calcifications known as psammoma bodies can be seen. On MRI, meningiomas appear as well-circumscribed heterogeneous masses that enhance with gadolinium, as well as have the dural tail sign (Figure 25–9). On CT scan, the calcifications may be visible. Angiography may also be helpful as these lesions tend to be vascular. Surgical removal is necessary when symptoms appear.[24,25]

Cholesteatoma

Cholesteatomas of the petrous apex are thought to arise from epithelial rests trapped during development. The squamous cell debris begins to organize as a cyst, which starts a destructive process that may erode bone. Another danger is secondary infection, which can cause intratemporal as well as intracranial abscess formation. Presenting symptoms can include hearing loss, otorrhea, facial paresis, or diplopia. This diagnosis is also confirmed with a combination of CT and MRI. Computed tomography shows a smooth expansile lesion with bony destruction. On MRI, the T_1 intensity is low, whereas the T_2 intensity is usually increased. This lesion does

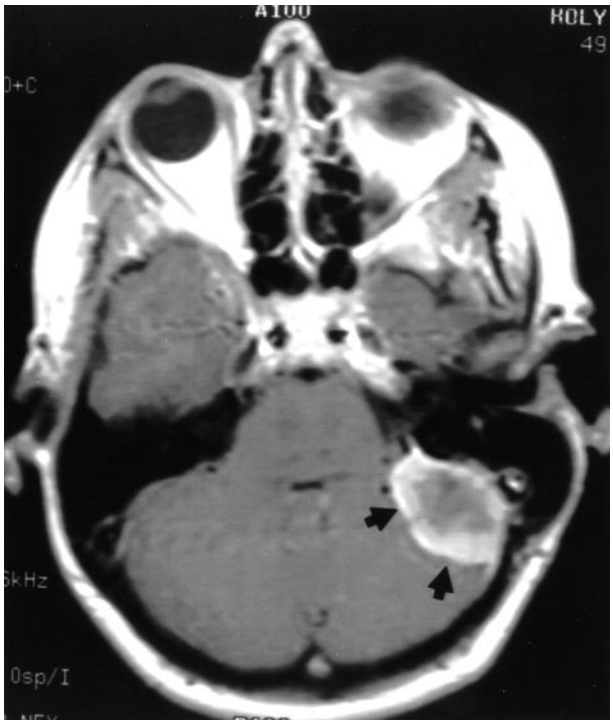


FIGURE 25–9. Gadolinium-enhanced T_1-weighted magnetic resonance image of a large meningioma (*arrows*) originating from the posterior face of the left petrous pyramid.

not enhance with gadolinium. Surgical extirpation is the treatment of choice.[26]

Glomus Tumors

Glomus tumors are slow-growing, vascular lesions, with the majority (97%) being benign. They are paragangliomas that arise from groups of chemoreceptor cells of neural crest origin, which are predominantly found along the course of Jacobson's nerve and within the adventitia of the jugular bulb. Tumors arising from Jacobson's nerve and limited to the promontory are classified as glomus tympanicum tumors, whereas those arising from the glomus bodies of the jugular bulb are classified as glomus jugulare tumors. The classification of these tumors can be found in Table 25–3. Glomus tumors have been reported to be multicentric in up to 10% of patients.[27] Glomus vagale tumors arise from the chemoreceptors along the vagus nerve. Glomus tumors may secrete vasoactive substances or be associated with other tumors, such as pheochromocytomas, which secrete these substances.

The diagnosis of glomus tumors is often difficult because the presenting signs and symptoms can be nonspecific, mimicking other otologic and neurologic conditions. By history, patients with glomus tumors typically present with pulsatile tinnitus, followed by the development of hearing loss. The hearing loss is usually conductive because the middle ear is filled with tumor, although a sensorineural hearing loss may develop if the cochlea is invaded.

Patients with glomus tympanicum tumors present with smaller tumors because the middle ear becomes involved earlier than with glomus vagale or jugulare tumors. On the other hand, patients with glomus jugulare tumors may present with dysfunction of jugular foramen contents (Vernet's syndrome, involving cranial nerves IX, X, and XI).

Careful otoscopic evaluation, preferably with an operating microscope, is essential. Classically, a red tumor mass beneath an intact tympanic membrane is seen. If limited to the promontory with all tumor margins visible, it probably represents a small glomus tympanicum tumor. If the tumor extends beyond the level of the annulus (especially inferiorly), its classification cannot be made visually.

A small mass seen in the floor of the middle ear may represent the "tip of the iceberg" of a large glomus jugulare tumor that may extensively involve the skull base. These tumors may blanch with positive pressure on pneumotoscopy (Brown's sign) or have cessation of tumor pulsation with ipsilateral carotid artery compression (Aquino's sign). The tumor also may extend through the tympanic membrane and present as an aural polyp. An audible bruit or evidence of vascular pulsations during tympanometry may also be present in these patients.

A glomus tumor filling the middle ear may also cause a diffusely discolored appearance of the tympanic membrane, which may be mistaken for serous fluid. Other entities that can be confused with glomus tumors include a dehiscent jugular bulb or, rarely, a dehiscent carotid artery. Cholesterol granuloma can easily be confused with large glomus tumors because they can make the tympanic membrane appear reddish-blue. Judicious use of the appropriate diagnostic imaging techniques readily differentiates these possibilities.

Any attempt at biopsy should be deferred until a complete workup is done. A systematic protocol allows one to make the diagnosis and to define the extent of the tumor (Figure 25–10).

TABLE 25–3. Classifications of Glomus Tumors

Fisch Classification of Glomus Tumors of the Temporal Region

Type A	Tumors limited to the middle ear cleft
Type B	Tumors limited to the tympanomastoid area with no infralabyrinthine compartment involvement
Type C	Tumors involving the infralabyrithine compartment of the temporal bone and extending to the petrous apex
Type D1	Tumors with an intracranial extension less than 2 cm in diameter
Type D2	Tumors with an intracranial extension greater than 2 cm in diameter

Glasscock-Jackson Classification of Glomus Tumors

Glomus Tympanicum

Type I	Small mass limited to the promontory
Type II	Tumor completely filling the middle ear
Type III	Tumor filling the middle ear and extending into the mastoid
Type IV	Tumor filling the middle ear, extending into the mastoid or through the tympanic membrane to fill the external auditory canal; may also extend anterior to the internal carotid artery

Glomus Jugulare

Type I	Small tumor involving the jugular bulb, middle ear, and mastoid
Type II	Tumor extending under the internal auditory canal; may have intracranial extension
Type III	Tumor extending into the petrous apex; may have intracranial extension
Type IV	Tumor extending beyond the petrous apex into the clivus or infratemporal fossa; may have intracranial extension

De La Cruz Glomus Tumor Classification with Associated Surgical Approach

Classification	Surgical Approach
Tympanic	Transcanal
Tympanomastoid	Mastoid—extended facial ridge
Jugular bulb	Mastoid-neck (possible limited facial nerve rerouting)
Carotid artery	Infratemporal fossa ± subtemporal
Transdural	Infratemporal fossa/intracranial
Craniocervical	Transcondylar
Vagal	Cervical

Fisch Classification of Glomus Tumors

A	Limited to the middle ear
B	Involves the middle ear and mastoid, without infralabyrinthine involvement
C	Extends into the infralabyrinthine area and to the petrous apex; originates in the jugular bulb and may erode overlying bone
C1	May erode carotid foramen but does not involve the artery
C2	Involves vertical part of the carotid canal
C3	Invades vertical and horizontal parts of the carotid canal but does not reach the foramen lacerum
C4	Involves the whole intrapetrous part of the carotid from the foramen lacerum to the cavernous sinus
D	Intracranial extension
De1	Extradural < 1 cm
De2	Extradural > 1 cm
Di	Intracranial, intradural
Di3	Intracranial, unresectable

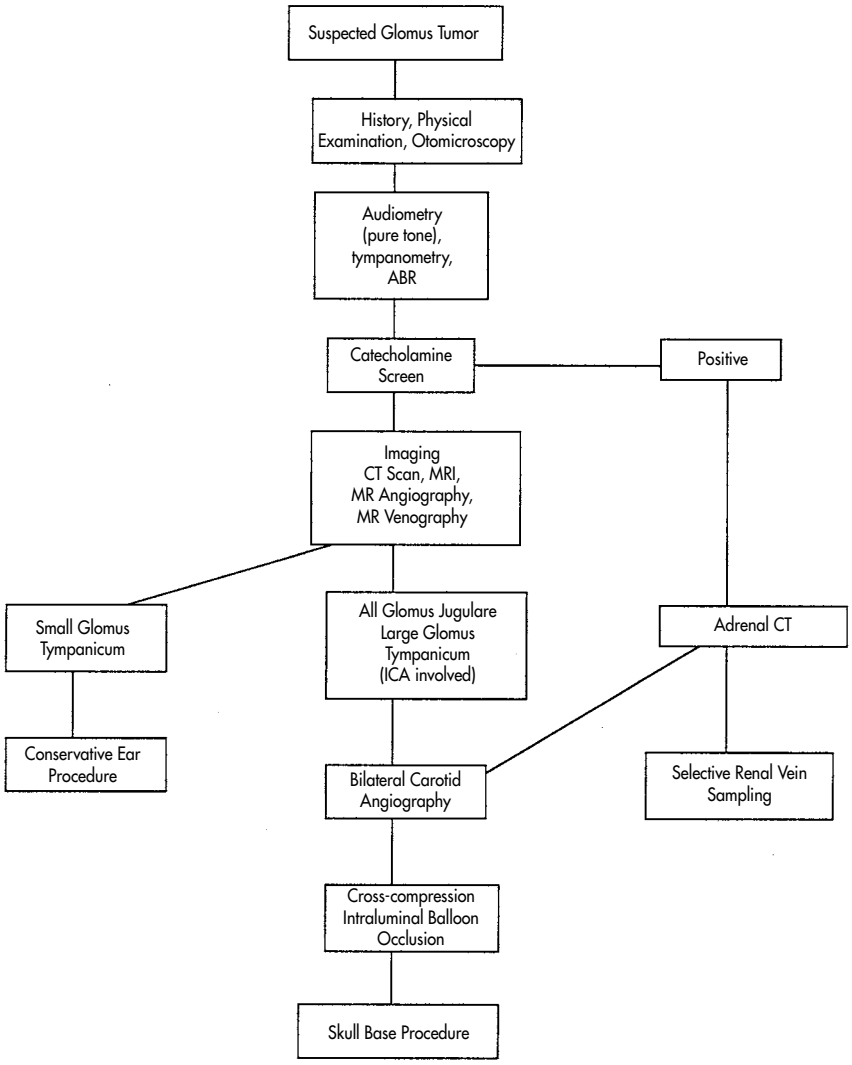


FIGURE 25–10. Glomus tumor diagnostic and treatment algorithm. ABR = auditory brainstem response; CT = computed tomography; MRI = magnetic resonance imaging; ICA = internal auditory canal.

These patients should undergo high-resolution CT with contrast injection. If the tumor is limited to the middle ear and there is no bony destruction, the tumor is classified as a glomus tympanicum tumor, and no further radiologic workup is needed.[28] If the CT shows any bony destruction or connection with the jugular bulb, an MRI is obtained, which helps to define any intracranial extension or jugular vein or carotid artery involvement and also screens for the presence of a second glomus tumor in the head and neck.[29,30] If workup reveals a glomus jugulare tumor and the patient is a surgical candidate, preoperative four-vessel angiography is indicated.

Glomus Tumor Algorithm The characteristic salt-and-pepper appearance on MRI corresponds to flow voids. Special subtraction MRI sequences may help distinguish a glomus tympanicum tumor from a glomus jugulare tumor.[31] Flow-sensitive weighting of the MRI can produce a magnetic resonance arteriography and magnetic resonance venography that are very useful in imaging glomus tympanicum lesions. The flow-sensitive modalities can indicate occlusion of the jugular bulb and vein, which can be particularly helpful for identifying involvement of the jugular bulb itself. It must be remembered, however, that an enhanced MRI involves T_1 imaging that produces a bright signal characteristic for fat, which is a major constituent of bone marrow. Thus, this characteristic may overestimate the degree of tumor involvement when there is paraganglioma adjacent to bone marrow–containing spaces such as the petrous apex.[32] A particular value of flow-weighted MRI is that it can be a useful screening test for synchronous paragangliomas.[33] This study can identify

glomus tumors in the temporal bone and along the carotids without the need for intra-arterial infusion.[34] After angiography, superselective embolization of the tumor-feeding vessels is performed at the same time. During cerebral angiography, the patency of the venous drainage system of the opposite side and the patency of the cerebral crossflow are evaluated. A combination of CT, MRI, and angiography clearly defines the extent of the disease and therefore limits unexpected findings at the time of operation.

Surgical removal is the treatment of choice. Preoperative embolization can decrease operative morbidity. Radiation therapy can be used alone; however, it is usually reserved for patients with concurrent medical problems or the elderly, who may be at higher risk for surgical complications.

Encephaloceles

Encephaloceles are defined as intracranial contents beyond the confines of the cranial vault. Various components may be extruding. In meningoceles, only the meninges and CSF have herniated. Encephalomeningoceles represent herniations of the brain along with meninges, whereas hydroencephalomeningoceles represent the herniation of ventricles along with the brain.

Encephaloceles can be divided into acquired and spontaneous lesions. An acquired lesion is one that occurs secondary to an event such as trauma, chronic infection, or surgery. Lesions are referred to as spontaneous if no preceding event is present. Spontaneous lesions can be further subdivided into congenital and idiopathic lesions. Congenital lesions occur in 1 in 3,000 to 10,000 live births. They are more common in the anterior and posterior fossa. Congenital temporal lobe herniations are rare but do occur. There are many theories to try to explain temporal lobe herniations. Most patients with a herniation are found to have multiple defects in the temporal bone. Arachnoid granulations greater than 3 mm are thought to cause bony erosion, leading to herniation.[35]

Presentation is similar in all types of lesions. They can present with CSF rhinorrhea or otorrhea, recurrent meningitis, serous otitis media, or a conductive hearing loss. Diagnosis is made with a combination of CT for bony detail and MRI for soft tissue detail. Surgical repair is necessary if they become symptomatic. For lesions greater than 1.5 cm, the recommended treatment involves a transmastoid exposure of the lesion from below combined with a middle fossa craniotomy to expose the superior aspect of the defect.

Infectious Lesions

Petrous apicitis is a coalescence of the cells in the apex usually associated with a chronic infectious process, mastoiditis, or cholesteatoma. Infectious agents include *Staphylococcus, Haemophilus*, or *Pseudomonas*. Symptoms include aural discharge, retro-orbital pain, and a sixth nerve palsy (Gradenigo's syndrome). Possible sequelae can include intracranial abscess formation, meningitis, and other cranial nerve palsies. Computed tomography will reveal bony destruction with irregular borders. On MRI, the T_1 image shows low signal intensity; however, the lesion is very bright on T_2. Gadolinium infusion shows a ring of enhancement around the lesion. Treatment includes intravenous antibiotics and a surgical procedure to establish adequate and sustained drainage and ventilation of the petrous apex.

Vascular Compression Syndromes

Vascular compression syndromes are abnormal excitation of the nerve owing to sustained contact with adjacent blood vessels. Prolonged contact of a blood vessel to a neural sheath causes chronic excitation and reorganization of the cranial nerve nucleus, resulting in hyperfunction.[36] Depending on the involved cranial nerve, this condition may result in hemifacial spasm, trigeminal neuralgia, or cochleovestibular compression syndrome.

Hemifacial spasm consists of involuntary facial muscle spasms. These usually begin with the orbicularis oculi and proceed to involve the lower parts of the face. Eventually, the patient can develop facial weakness thought to be caused by demyelination secondary to the compression. The unilaterality of this disorder helps to differentiate it from essential blepharospasm.

Trigeminal neuralgia is characterized by piercing facial pain. Tactile stimulation of the face, pain, cold, or dental work can cause hyperactivity in the nucleus, resulting in the pain. Cochleovestibular compression syndrome has been reported to result in continuous vertigo, hearing loss, nausea, tinnitus, and imbalance.

In these patients, all other potential causes must be ruled out. For those failing more conservative treatment (ie, medical therapy to relieve symptoms), surgical exploration of the nerve and decompressing the offending vascular structure are an option. In the CPA, this is usually the anterior inferior cerebellar artery or posterior inferior cerebellar artery, but veins may also be the culprits.

MANAGEMENT

It is beyond the scope of this chapter to review the treatment issues for each of the skull base disorders, but a general overview of available modalities and techniques is designed to give the reader some basis for understanding their roles.

Some of the lesions discussed are malignant by definition. The majority are benign but clinically "malignant" either because of the debility of symptoms or their proximity to neural and vascular structures. The histologic classification, severity of symptoms, and proximity to important structures play a role in treatment decision making. Treatment options generally include surgery, radiation, chemotherapy, or simple observation. Oncologic chemotherapy has a limited role in the management of skull base lesions aside from an occasional primary or metastatic cancer. This discussion centers around mass lesions that require intervention or pathophysiologic disorders that have failed medical therapy.

Radiation

External beam radiation has been used as an adjunct to surgery in certain lesions such as olfactory neuroblastoma. Primary radiation traditionally has been used in the more common benign lesions (such as acoustic neuroma) in the elderly or medically disabled, but treatment results vary. Primary risks include damage to the brain and nerves and the superficial structures such as skin as well as osteoradionecrosis of the bone.

More recently, stereotactically directed gamma radiation ("gamma knife") has been discussed as a treatment option.[37] Numerous cobalt sources (approximately 200) are positioned in a hemispheric pattern oriented toward a common point. In this technique, high doses of radiation are given to a more localized area in shorter periods of time than in traditional radiation therapy. Commonly, patients receive their total radiation dose in a single treatment. Thus, the risks should be less than traditional radiation, and the time involved is greatly reduced (1 day versus approximately 6 weeks). Its use is limited to somewhat smaller tumors (acoustic neuromas < 2 cm), and it appears to be an option for surgically unsuitable patients with appropriate-size tumors. The obvious benefit is avoidance of a large, usually intracranial surgical procedure. In the specific example of acoustic neuroma, the initial risks are markedly less than surgical intervention, but over time, the risks to adjacent structures such as the facial or cochlear nerves approach or exceed the surgical risks. Halting or limiting the growth of the tumor is the usual outcome, with regression less commonly noted. Leskell first used the gamma knife in 1969.[38] Local control rates on tumor growth were first reported to be 86% for unilateral tumors[39]; Flinkinger et al reported a local control rate of 89%.[40] Maire et al reported a 6-year actuarial rate of 82% local control (no growth) with external beam radiation.[41] Owing to the expense of gamma-knife radiation recently, development of linear accelerator–based stereotactic radiation has ensued. It is less expensive and used for conventional radiation. Mendenhall et al treated patients with a linear accelerator–based stereotactic radiation and reported a local control rate of 98% and actuarial control rate of 95%.[42]

Although surgery is the mainstay for the treatment of patients with acoustic neuromas, radiation therapy has improved and now is considered an alternative option for a certain subset of patients, especially patients who are not surgical candidates.[41] An occasional problem arises when tumors do not respond to radiation and require surgical treatment. The surgical risks are increased since the radiation-damaged tissue is generally more difficult to dissect.

Surgery

Considerations The skull base is an area dense with important neural, vascular, and structural entities. Preoperative discussions with patients must include a thorough review of the pathologic entity as well as the potential and planned morbidity of the treatment. Each procedure and skull base area have their own individual structures at risk.

In the majority of accessible tumors, surgical removal is the treatment of choice. Surgery also

allows access to particular structures such as the seventh cranial nerve for microvascular decompression and the eighth cranial nerve for vestibular neurectomy. Success depends on many factors. Two important aspects are accurate assessment of the clinical entity preoperatively and, of course, the skill of the surgeon. The choice of surgical approach depends on the pathologic process and the expertise of the surgeons(s).

Expected morbidity, such as loss of any residual hearing in a translabyrinthine approach for acoustic neuroma removal, should be discussed and weighed against the benefits of surgery in general and specifically of a particular procedure/approach. In the case of the translabyrinthine approach, decreased risk to facial nerve and less cerebellar retraction are the primary advantages.

Complications of Skull Base Surgery See Table 25–4 for complications of skull base surgery.

The more important complications result from alteration of function of intracranial structures. Infarction of intracranial structures, such as the cerebellum or brainstem, from excessive retraction or blood vessel damage can result in significant morbidity or death.

Intracranial hemorrhage may occur in the immediate or late postoperative period. Treatment usually includes immediate opening of the wound and return to the operating room for complete drainage of the blood with identification and control of the bleeding site. If identified and treated early, there may be no residual sequelae, but important functional defects are possible.

Complications may arise simply as a result of entrance into the CSF-containing space. Meningitis is a known risk and tends to occur between the third and seventh postoperative days. Because of the frequent postoperative malaise, head/neck discomfort, and fevers from meningeal irritation, meningitis can be difficult to identify in this situation. A low threshold of suspicion and early lumbar puncture will allow early identification and rapid resolution with the institution of appropriate intravenous antibiotics.[43]

Cerebrospinal fluid leaks can occur at any time after the operation, but the risk is greatest in the first 10 days. In high-risk procedures, CSF lumbar drains may be placed at the time of operation to keep intracranial pressures low and decrease tension on wound closures. Drainage may occur through the nose (CSF rhinorrhea), the ear (CSF otorrhea), or the incision. In the anterior approaches (craniofacial), the fluid will generally drain through the nose anteriorly, whereas in the posterior approaches, the CSF reaches the nose through the eustachian tube and may result in either anterior or posterior rhinorrhea. If a tympanic membrane perforation is present at the time of operation in the middle or posterior fossa or occurs in the postoperative period, CSF otorrhea may occur. Initial treatment of CSF leak consists of head elevation and either lumbar puncture or placement of an indwelling lumbar spinal drain. If these conservative measures are unsuccessful, then surgical exploration and reclosure of the wound are necessary.

TABLE 25–4. Complications and Sequelae of Skull Base Surgery

Cerebrospinal fluid leakage
Meningitis
Chemical
Bacterial
Intracranial hemorrhage
Wound infection
Wound hematoma
Cerebrovascular accident/coma
Cranial nerve deficits
External auditory canal stenosis
Urinary tract infection
Aspiration
Pulmonary embolism
Death
Pneumonia
Adult respiratory distress syndrome
Pneumothorax
Cerebellar dysfunction
Bulbar dysfunction
Hydrocephalus
Seizures

Injury to any of the vascular structures (ie, carotid artery) or neural structures (ie, cranial nerves) is possible and must be dealt with intraoperatively if the structure has been interrupted and postoperatively if there are significant sequelae. In the lateral and posterior approaches, altered function of the seventh cranial nerve is a frequent postoperative problem. Transection of the nerve intraoperatively requires immediate primary repair or cable grafting, if possible. In this case, delayed and incomplete recovery would be expected. If the continuity is maintained, but injury occurs secondary to stretching or contusion, recovery generally occurs, but time and completeness are variable. Risk to the cornea, which results from incomplete eye closure, is the most important issue. For expected transient weaknesses, local eye care such as use of lubricants is all that is generally needed. For longer periods of altered function or permanent paralysis, procedures such as upper eyelid gold weight implantation, lower eyelid ectropion repair, or tarsorrhaphy are frequently indicated for adequate corneal protection. In cases of permanent paralysis, nerve grafting, twelfth to seventh end-to-side (jump graft) or end-to-end grafting, or muscle transposition (ie, temporalis) may be useful for rehabilitation of the entire face.

Intraoperative Monitoring Preservation of neural structures is a crucial part of skull base surgery. Although there is no substitute for surgical knowledge, neural monitoring is a useful adjunct. The motor nerve integrity can be followed by monitoring EMG activity in the appropriate muscles. Increased activity will be seen as the nerve is manipulated. The nerve of interest can also be electrically stimulated in the operative field to verify its identity. The seventh nerve is probably the most commonly and easily monitored motor nerve.

In the hearing preservation approaches in acoustic neuroma surgery, auditory function can be monitored in several ways. Most commonly, continuous ABR can be elicited. This has the advantage of being readily available and easy to set up but the disadvantage of requiring many averaged trials to detect an alteration. This can translate into a significant delay between the occurrence of a compromising action and the detection of the change in the ABR. The recording of an electrocochleogram during a procedure is frequently used. As in ABR, stimulation is with repeated auditory clicks through an external auditory canal insert. Since recordings are generally obtained with a needle electrode on the promontory, the responses are more robust, require fewer trials, and ideally require less time between the operative maneuver and the change in the observed waveform. Direct eighth nerve monitoring records nerve impulses traveling in the auditory nerve. The electrode is placed on the nerve proximal to the operative site so that any alteration in the physiologic status of the peripheral part of the auditory system will be immediately detected.[44]

Surgical Approaches Because of the complexity of the skull base and the diversity of the possible abnormalities, many surgical approaches have been developed. They are continually being modified, and new ones are being described.[45,46]

Craniofacial Approach. For lesions that involve the roof of the ethmoid sinuses or the superior nasal vault (cribriform area), visualization from above (cranially) and below (facially) can be obtained through a craniofacial approach. The most common lesions include olfactory neuroblastoma and ethmoid sinus cancer.

The approach from above is initiated with a bicoronal incision originating anterior and superior to the auricles bilaterally followed by flap elevation over the forehead (Figure 25–11). An anterior craniotomy exposes the frontal sinus, which is completely removed, and the frontal lobes that are gently elevated, necessitating transection of the olfactory nerves. This allows visualization of the floor of the anterior cranial fossa, including the roofs of the orbits and cribriform plate area. Visualization can be obtained back to the area over the anterior portion of a fully pneumatized sphenoid sinus. The cribriform plate and various amounts of bone can be resected. Ideally, the tumors are extradural, although dura (and even brain parenchyma) can be resected.

The approach from below is more variable but generally involves at least an ethmoidectomy. Most commonly, a lateral rhinotomy incision is used to allow a generous view of the nasal vault and associated paranasal sinuses. If necessary, the ethmoidectomy can be combined with maxillectomy or orbital exenteration. Reconstruction and closure vary, depending on the resection performed. At the com-

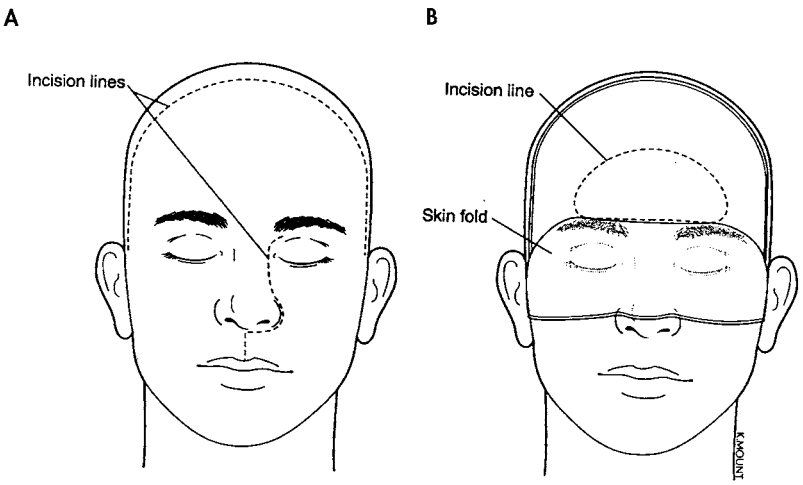


FIGURE 25–11. A, The skin incisions for a craniofacial resection are shown. *B*, After elevation of the forehead skin and retraction inferiorly, an approximate outline of the frontal craniotomy (*dotted line*) can be seen.

pletion of the procedure, dural repair can be accomplished with temporalis fascia, fascia lata, or a local pericranial flap. In the most limited resections, the anterior cranial bone flap is replaced, and the upper and lower incisions are closed primarily. For more extensive resections, local or distant flaps may be required to reconstruct the defect, and/or surgical prostheses may be necessary.

Infratemporal Fossa Approach. For access to the infratemporal fossa and the jugular foramen, the infratemporal fossa approach is available. Glomus jugulare and vagale tumors, vagal neuromas, and deep lobe parotid tumors may arise in this area.

A large curvilinear incision is made to allow access to the lateral skull and the upper cervical areas (Figure 25–12). The skin flap is elevated anteriorly. The sternocleidomastoid muscle must be detached from its mastoid insertion to allow the generous exposure that is usually necessary. The seventh nerve is identified exiting from the stylomastoid foramen. The ninth, tenth, and eleventh cranial nerves can be identified at the skull base exiting from the jugular foramen into the neck. The hypoglossal nerve is seen crossing lateral to the carotid at approximately the level of the carotid bifurcation.

The internal jugular vein and internal carotid artery are identified and isolated so they can be easily controlled.

Superiorly, the structures are located through a transmastoid approach. After the mastoidectomy, the sigmoid sinus, posterior and middle cranial fossae dura, and vertical portion of the facial nerve are identified (Figure 25–13). In the majority of these procedures, the external auditory canal is removed, and the lateral canal skin is closed on itself. The facial nerve is skeletonized in the mastoid and middle ear and then completely mobilized. It is retracted anteriorly to allow direct visualization of the more

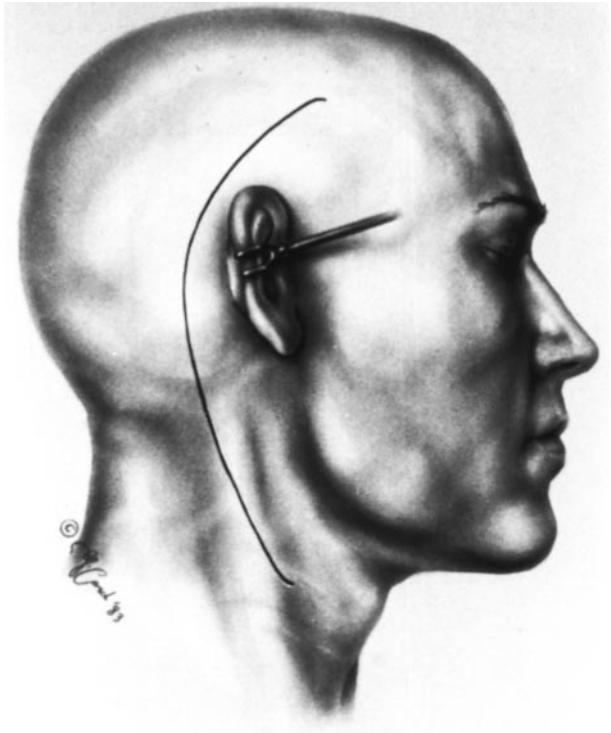


FIGURE 25–12. A large C-shaped incision extending from the parietal area across the mastoid and inferiorly into the cervical area is necessary for complete exposure in the infratemporal fossa approach. Reproduced with permission from Wiet RJ, Causse J, editors. Complications in otolaryngology: head and neck surgery. Vol 1. Philadelphia: BC Decker; 1986.

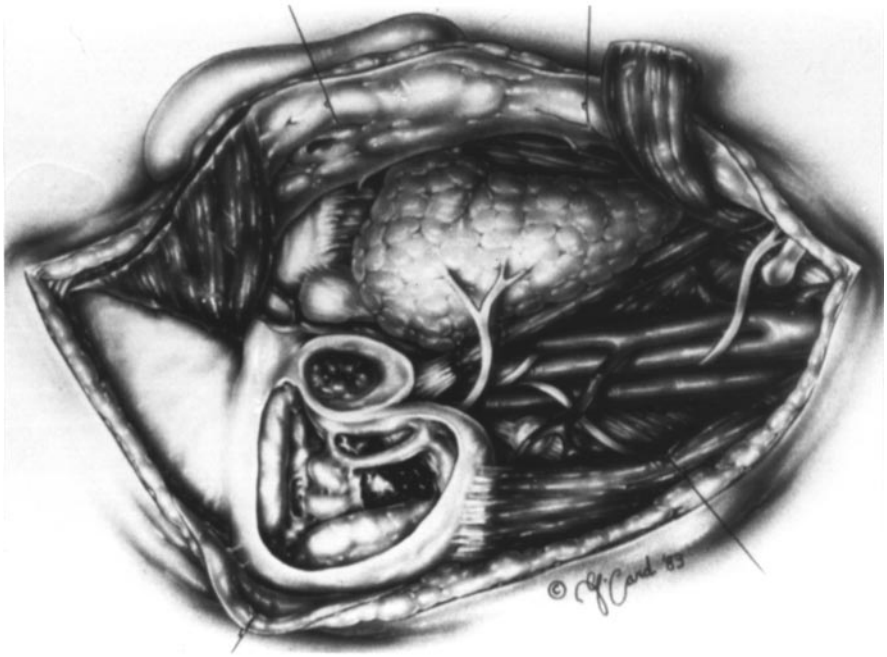


FIGURE 25–13. After cervical exposure and mastoidectomy, most of the important structures are visible. In this drawing, a glomus jugulare tumor can be seen medial to the facial nerve in the mastoid. Reproduced with permission from Wiet RJ, Causse J, editors. Complications in otolaryngology: head and neck surgery. Vol 1. Philadelphia: BC Decker; 1986.

medial structures (Figure 25–14). Having done this, the remaining infralabyrinthine bone is removed to allow access to the tumor, the entire sigmoid sinus/jugular bulb/internal jugular vein, and the internal carotid artery anteromedial to the jugular vein. Removal of the tumor is then generally possible. In glomus jugulare tumors, the sigmoid sinus, jugular bulb, and upper portion of the internal jugular vein are resected with the tumor. The operation may be continued intracranially in the same sitting or in a second stage for larger tumors. In glomus jugulare tumor operations, major blood loss is usually encountered, although it is usually markedly decreased with preoperative embolization.

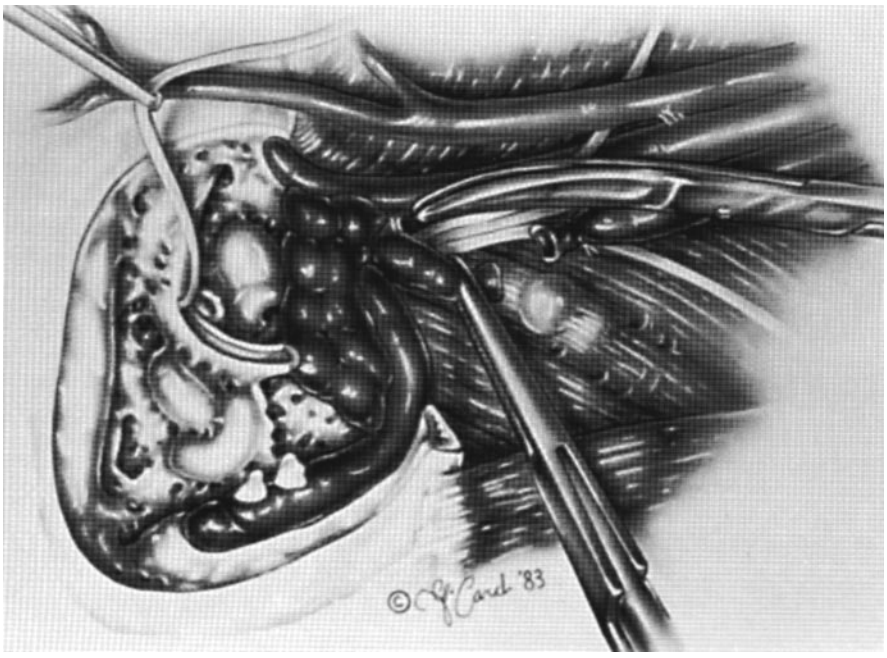


FIGURE 25–14. After mobilization of the facial nerve and removal of additional bone medially, excellent visualization of the glomus jugulare tumor is possible. Reproduced with permission from Wiet RJ, Causse J, editors. Complications in otolaryngology: head and neck surgery. Vol 1. Philadelphia: BC Decker; 1986.

In most situations, the wound can be closed primarily. In more extensive resections, abdominal fat grafting, local tissue flaps such as temporalis muscle, or free vascularized grafts may be necessary.

Internal Auditory Canal/Posterior Fossa Approaches. These surgical approaches overall are the most commonly used and are designed to allow exposure of the IAC, the CPA, and the petrous apex. The middle cranial fossa, translabyrinthine, and retrosigmoid approaches can all be used for resection of acoustic neuromas (Figure 25–15). The middle cranial fossa and retrosigmoid approaches are classified as hearing preserved, and in many instances, the cochlear division of the eighth nerve can be preserved, allowing at least the possibility of hearing postoperatively. In the translabyrinthine approach, the vestibular portion of the labyrinth is intentionally removed. Although this allows better visualization of the seventh nerve and the IAC, any residual hearing is predictably and completely lost. Each has its advantages and indications.

Middle Cranial Fossa Approach. The middle cranial fossa procedure approaches the IAC from above, and here the labyrinth can be preserved in an attempt to conserve hearing. This procedure is commonly used in the resection of acoustic neuromas and for vestibular neurectomy. In the IAC, the vestibular portions of the eighth nerve are distinct from the cochlear division, and, at least theoretically, the vestibular nerve can be sectioned more precisely. In acoustic neuroma surgery, the middle cranial fossa approach is used when two main criteria are fulfilled: the tumors are completely within the IAC (intracanalicular) and the patient has preoperative hearing that is good enough to justify a procedure that is somewhat more complex than the other approaches.

This procedure can also be used to access the area medial to the IAC (petrous apex) for lesions such as cholesterol granulomas. For larger temporal lobe encephaloceles, visualization through the middle cranial fossa approach is necessary to bring the brain back into the intracranial cavity and to cover

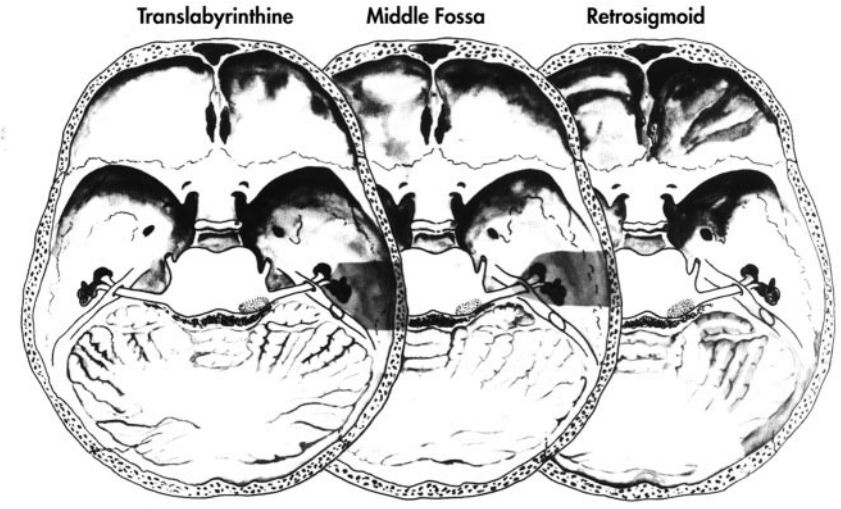


FIGURE 25–15. Diagram of the three main approaches to tumors of the cerebellopontine angel (CPA). The translabyrinthine approach gives the surgeon the best exposure but sacrifices hearing. The middle fossa offers the least exposure and thus is useful only for intracranial tumors in patients in whom hearing preservation will be attempted. The retrosigmoid approach gives good exposure of the CPA but offers the least number of landmarks for identification of the facial nerve. We reserve it for small tumors of the CPA in patients in whom hearing preservation is worthwhile. Reproduced with permission from Monsell EM, McElveen JT, Hitselberger WE, House WF. Surgical approaches to the human cochlear nuclear complex. Am J Otol 1987;8:450–5.

the bony defect adequately (generally with a bone flap). Additionally, this approach is used for exploration of the geniculate and labyrinthine portion of the seventh nerve in cases of temporal bone trauma and for bony decompression for Bell's palsy.

With the surgeon at the head of the table, the patient is placed in a supine position with the head turned so the involved side is up. Although different incisions have been used, the simplest is a vertical incision extending from a point in the preauricular area 1 cm anterior to the auricle directly superiorly for 6 to 8 cm. The incision is carried down to the bone.

A 3 × 3 cm bone flap is created in the underlying bone and is removed to expose the dura overlying the temporal lobe. The temporal lobe is elevated off the floor of the middle cranial fossa until the foramen spinosum and middle meningeal artery are seen anteriorly, the greater superior petrosal nerve is seen anteromedially, and the superior petrosal sinus is seen posteriorly (Figure 25–16). The pertinent surface of the temporal bone can now be visualized. The arcuate eminence (the bony elevation over the superior semicircular canal) and the greater superficial petrosal nerve are the most common overt landmarks. From these, the facial nerve and the IAC can be identified once drilling has begun. The superior semicircular canal (laterally) and the cochlea (medially), which lie on each side of the IAC, must not be violated or hearing will be lost.

In vestibular neurectomy, the superior and inferior vestibular nerves are identified in the IAC and transected. For seventh nerve decompression, exposure of the geniculate ganglion and labyrinthine portion of the seventh nerve with opening of the nerve sheath is completed. In acoustic neuroma resection, the tumor is carefully removed from the IAC, with care being taken to preserve the seventh nerve and the cochlear branch of the eighth cranial nerve.

Fascia and/or a muscle plug are used to close the dural defect, and the retracted temporal lobe is placed back in its anatomic position. The bone flap is replaced, and the incision is carefully closed.

Retrosigmoid Approach. In the retrosigmoid approach, the CPA is approached directly, and the IAC can be approached posteriorly. Because this approach allows a generous view of the CPA, it is useful in a variety of conditions and is used most frequently for CPA tumor resection, vestibular neurectomy, and microvascular cranial nerve decompression. Because this is a hearing preservation approach, it is used in situations in which serviceable hearing is present and the tumor is extracanalicular but is less than 2 cm. However, for

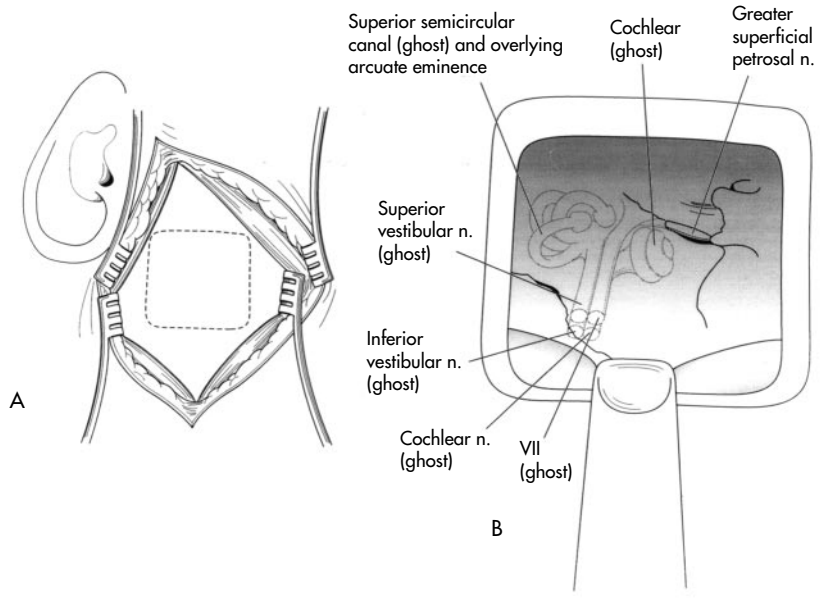


FIGURE 25–16. A, View of the left middle cranial fossa approach from the surgeon's position showing the area of planned craniotomy (*dotted line*) after a vertical incision has been opened. *B*, A view through the craniotomy allows exposure of the middle cranial fossa floor after retraction of the temporal lobe. The significant underlying structures are outlined.

intracanalicular tumors, the middle cranial fossa approach is more appropriate. In tumors greater than 2 cm, hearing preservation is unlikely, so the translabyrinthine approach is used.

First, a curvilinear incision is made 4 cm posterior to the auricle to create an anteriorly based flap (Figure 25–17). The bone of the mastoid and occipital area is exposed. A bone square is created with the sigmoid sinus defining the anterior limit and the transverse sinus defining the superior limit. Removal of the bone plug exposes the dura overlying the cerebellum. The dura is incised 5 to 10 mm posterior to the sigmoid sinus and inferior to the transverse sinus. The cerebellum is retracted posteriorly to expose the CPA. In acoustic neuroma resection, the posterior lip of the IAC must be drilled away to allow visualization of the intracanalicular portion of the tumor.

Since the fifth to eleventh nerves can be readily viewed, microvascular decompression of the fifth (for trigeminal neuralgia), seventh (for hemifacial spasm), or eighth (for tinnitus or disabling positional vertigo) nerves can be readily accomplished by gently placing a strip of Teflon felt between the offending blood vessel and the nerve to prevent the irritating contact. Vestibular neurectomy is performed by identifying the longitudinal septum between the vestibular and the cochlear portions of the eighth nerve and transecting the upper (vestibular) half. With vestibular neurectomy, control of vertigo in Meniere's disease is generally reported to be greater than 90% regardless of the approach used.[47,48] Cerebellopontine angle tumors such as acoustic neuromas and meningiomas can be resected through this approach.

On completion of the procedure, the dura is closed in a watertight fashion. If this is not possible, an abdominal fat graft is laid over the closure to seal the dural defect. Although some surgeons drill away the bone of the entire cranial defect, we prefer to preserve the bone square and replace it at the completion of the procedure. The skin is closed in the usual manner.

Translabyrinthine Approach. The translabyrinthine approach allows a generous view of the IAC, CPA, and petrous apex, if necessary. This approach is generally used in patients with poor hearing as any remaining hearing will be predictably and totally lost.

Less brain retraction is required than in the middle cranial fossa or the retrosigmoid approach, and unlike the retrosigmoid approach, the facial nerve can be identified prior to tumor identification/removal. This approach is used in essentially all acoustic neuromas that are larger than 2 cm or in any patient with poor hearing regardless of the tumor size. Tumors greater than 4 to 5 cm may require this approach combined with the retrosigmoid approach for complete resection (Figure 25–18).

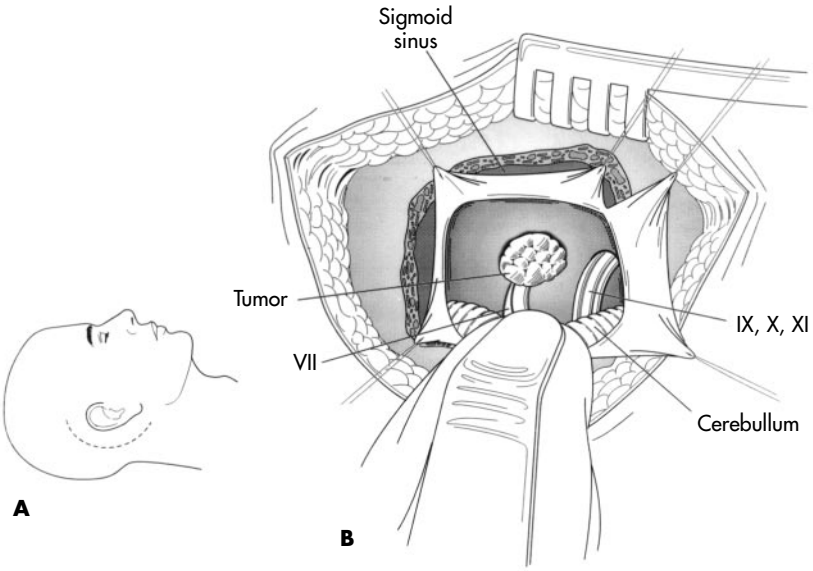


FIGURE 25–17. A, With the patient in the supine position, approximate incision line for the retrosigmoid approach is identified (*dotted line*). *B*, After craniotomy, dural incision, and retraction of the cerebellum, the cerebellopontine angle can be viewed. In this drawing, an acoustic neuroma is seen extending out of the internal auditory canal.

With the patient in the supine position and the head rotated away from the side of interest, a postauricular incision is made 2 to 3 cm behind the postauricular sulcus (Figure 25–19). The mastoid bone is completely exposed, and a drill is used to remove completely the bone of the mastoid. The sigmoid sinus, posterior fossa dura, and middle cranial fossa are completely decompressed. The vertical portion of the facial nerve is identified and left with a thin bony covering. The fossa incudus is clearly identified, and the incus is removed so that the middle ear space may be viewed and carefully packed at the end of the procedure. A labyrinthectomy is completed, and as drilling continues medially, the IAC is exposed. The dura of the posterior fossa and IAC is opened, and the IAC and CPA are viewed. In the fundus (lateral) portion of the IAC, a vertical crest of bone (Bill's bar) is present and is a useful landmark since the facial nerve is anterior and the superior vestibular nerve is posterior. After completion of the procedure, the middle ear space is carefully packed with small pieces of temporalis muscle, the aditus is covered with compressed bone pate and bone wax, and the remainder of the surgical defect is filled with fat obtained from the abdomen.[49] The incision is meticulously closed, and a mastoid dressing is placed.

In selected situations, the translabyrinthine approach can be "extended" to allow better access medially and anteriorly. In the *transcochlear* approach, drilling is continued after a complete mastoidectomy and labyrinthectomy.[50] The facial nerve is skeletonized and mobilized posteriorly from the stylomastoid foramen to the geniculate ganglion. The cochlea is then completely exenterated and the internal carotid artery skeletonized, allowing a generous view of the internal carotid artery, petrous apex, clivus, sixth and twelfth nerves, and the basilar and vertebral arteries, if necessary. A more extended view can be obtained through the *transotic* approach, with the tympanic membrane, malleus, incus, and posterior external auditory canal wall removed in addition to skeletonization and mobilization of the seventh nerve and exenteration of the cochlea.[51] In the *modified transotic* approach, the seventh nerve is left in the fallopian canal, decreasing the risk to it.[52]

Approaches to the Temporal Bone. A large amount of skull base pathology occurs within the temporal bone. Frequently, there are specific approaches available that are generally simpler than the intracranial procedures.

Tympanotomy is performed through the external auditory canal and allows exposure of the mesotympanum. Frequently, glomus tympanicum and congenital cholesteatoma are limited enough to be removed through this approach. Mastoidectomy

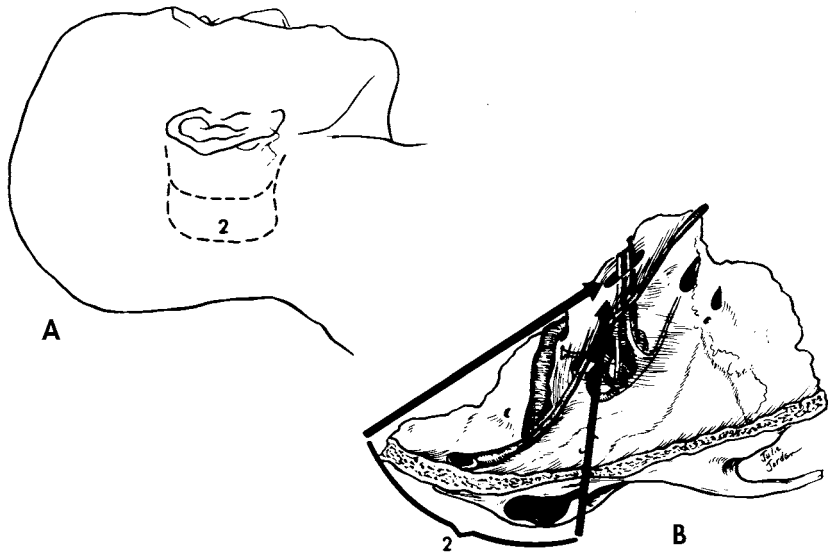


FIGURE 25–18. *A*, Postauricular incision for the translabyrinthine approach and its modification (2) for the combined approach. *B*, A combination of the translabyrinthine approach and retrosigmoid approach to the cerebellopontine angle may be used to excise larger tumors.

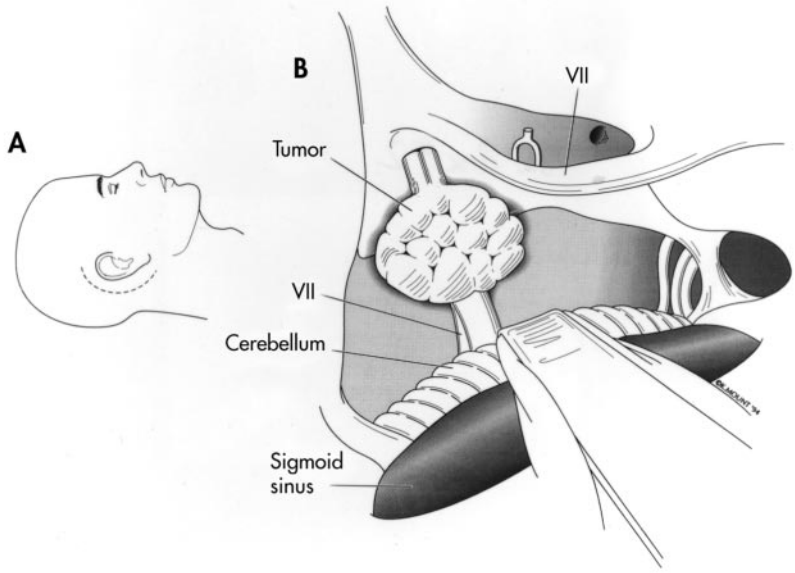


FIGURE 25–19. *A*, The approximate postauricular incision for the translabyrinthine approach is shown (*dotted line*). *B*, After the drilling is complete, generous visualization of the cerebellopontine angle is possible. An acoustic neuroma is seen arising from the eighth cranial nerve.

allows visualization of the mastoid air cell system and enables access to the antrum and epitympanum as well as the labyrinth and deeper structures of the temporal bone. Exploration and/or decompression of the second genu and vertical segment of the seventh nerve are possible.

Smaller tegmental encephaloceles can be repaired through this approach, although ones larger than 2 cm will also require a middle fossa approach. Access to the tympanum is possible from the mastoid between the facial nerve and the chorda tympani via the facial recess, opening the bone around the posterior and inferior parts of the external auditory canal. A good view of the posterior and inferior parts of the tympanum can then be obtained, especially for larger glomus tympanicum tumors.

In specific situations, limited access of the petrous apex is all that is necessary. In petrous apicitis and petrous cholesterol granuloma, drainage of the contents resolves the problem. If hearing is poor, a labyrinthectomy will allow a generous view medially. If useful hearing is present, the labyrinth must be avoided. The retrofacial approach is available after a complete mastoidectomy. Drilling is continued posterior to the facial nerve and inferior to the posterior semicircular canal. In the case of a cholesterol granuloma, the cyst is opened, and a stent is placed to maintain the drainage (Figure 25–20). Multiple other paths are available and are chosen based on the location of the lesion and the aeration of the temporal bone.

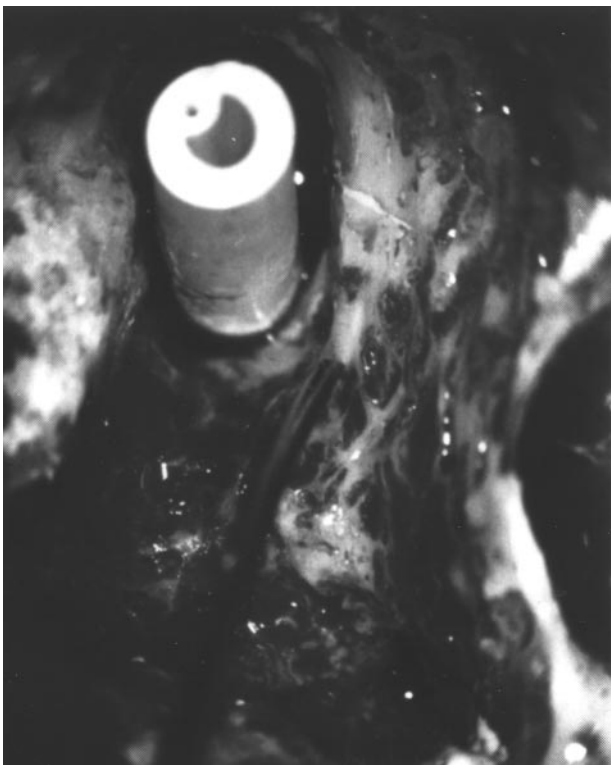


FIGURE 25–20. Infralabyrinthine approach and stent placement into a cholesterol granuloma of the left petrous apex. This patient presented with vertigo, which resolved postoperatively.

REFERENCES

1. Metson R, Gliklich RE, Stankiewicz JA, et al. Comparison of sinus computed tomography staging systems. Otolaryngol Head Neck Surg 1997;117: 372–9.
2. Metson R, Gliklich RE, Cosneza MJ. A comparison of image guidance systems for sinus surgery. Laryngoscope 1998;108:1164–70.
3. Sargent EW, Bucholz RD. Middle cranial fossa surgery with image-guided instrumentation. Otolaryngol Head Neck Surg 1997;117:131–4.
4. Wackym PA, King WA, Barker FG, Poe DS. Endoscope-assisted vestibular neurectomy. Laryngoscope 1998;108:1787–93.
5. King WA, Wackym PA. Endoscope-assisted surgery for acoustic neuromas (vestibular schwannomas): early experience using the rigid Hopkins telescope. Neurosurgery 1999;44:1095–102.
6. Wackym PA, King WA, Poe DS, et al. Adjunctive use of endoscopy during acoustic neuroma surgery. Laryngoscope 1999;109:1193–202.
7. Hart RG, Davenport J. Diagnosis of acoustic neuroma. Neurosurgery 1981;9:450–63.
8. House WF, Luetje M, editors. Acoustic tumors I: diagnosis. Baltimore: University Park Press; 1979.
9. Stockard JJ, Rossiter VS. Clinical and pathologic correlates of brain stem auditory response abnormalities. Neurology 1977;27:608–13.
10. Sekhar LN, Janecka IP. Intracranial extension of cranial base tumors and combined resection: the neurosurgical perspective. In: Jackson CG, editor. Surgery of skull base tumors. New York: Churchill-Livingstone; 1991.
11. Young NM, Wiet RJ, Russel EJ, Monsell EM. Superselective embolization of glomus jugulare tumors. Ann Otol Rhinol Laryngol 1988;97:613–20.
12. LaRouere MJ, Zappia JJ, Wilner HI, et al. Selective embolization of glomus jugulare tumors. Skull Base Surg 1994;4:21–5.
13. Harrison DF. Osseous and fibro-osseous conditions affecting the craniofacial bones. Ann Otol Rhinol Laryngol 1984;93:199–203.
14. Harrison DF. Surgical pathology of olfactory neuroblastomas. Head Neck Surg 1984;7:60.
15. Nager GT, Vanderveen TS. Cholesterol granuloma involving the temporal bone. Ann Otol Rhinol Laryngol 1976;85:204–9.
16. Thedinger BA, Nadol JB, Montgomery WW, et al. Radiographic diagnosis, surgical treatment and long term follow-up of cholesterol granuloma of petrous apex. Laryngoscope 1989;99:896–907.
17. Oot RF, Melville GE, New PF, et al. The role of MRI and CT in evaluating clival chordomas and chondrosarcomas. AJR Am J Roentgenol 1988;151: 567–75.
18. Kveton JF, Brackmann DE, Glasscock ME, et al. Chondrosarcoma of the skull base. Otolaryngol Head Neck Surg 1986;94:23–32.
19. Nager GT. Epidermoids involving the temporal bone: clinical, radiological and pathologic aspects. Laryngoscope 1975;85(12 Pt 2 Suppl 2):1–21.
20. Christensen WN, Long DM, Epstein JI. Cerebellopontine angle lipoma. Hum Pathol 1986;17:739–43.
21. Pappas DG, Brackmann DE. Arachnoid cysts of the posterior fossa. Otolaryngol Head Neck Surg 1981;89:328–32.
22. Hitselberger WE, Gardner G. Other tumors of the cerebellopontine angle. Arch Otolaryngol 1968;88: 712–4.
23. National Institutes of Health. Consensus statement on acoustic neuroma. Washington (DC): December 1991.
24. MacCarty MR, Symen L. Meningiomas of the clivus and apical petrous bone, report of 35 cases. J Neurosurg 1986;65:160.
25. Nager GT, Masica DN. Meningiomas of the cerebello-pontine angle and their relation to the temporal bone. Laryngoscope 1970;80:863–95.
26. Pyle GM, Wiet RJ. Petrous apex cholesteatoma: exteriorization vs. subtotal petrosectomy with obliteration. Skull Base Surg 1991;1:97.
27. Glasscock ME, Jackson CG, Dickins JR, Wiet RJ. Panel discussion: glomus jugulare tumors of the temporal bone. The surgical management of glomus tumors. Laryngoscope 1971;89:1640.
28. Phelps PD, Lloyd GA. Vascular masses in the middle ear. Clin Radiol 1986;37:359–64.
29. Vogl T, Bruning R, Schedel H, et al. Paragangliomas of the jugular bulb and carotid body: MR imaging with short sequences and GD-DTPA enhancement. AJR Am J Roentgenol 1989;153:583–7.
30. Phelps PD, Stansbie JM. Glomus jugulare or tympanicum? The role of CT and MR imaging with gadolinium DTPA. J Laryngol Otol 1988;102:766–76.
31. Warren WH, Calderelli DD, Lec I. Neuroendrocrine markers in paragangliomas of the head and neck. Ann Otol Rhinol Laryngol 1985;94:555–9.
32. Lloyd GA, Cheesman AD, Phelps PD, King CM. The demonstration of glomus tumors by subtraction MRI. Neuroradiology 1992:34:470–4.

33. Arriaga MA, Lo WW, Brackmann DE. Magnetic resonance angiography of synchronous bilateral carotid body paragangliomas and bilateral vagal paragangliomas. Ann Otol Rhinol Laryngol 1992: 101:955–7.
34. Stewart K, Kountakis S, Chang CY, Jahrsdoerfer RA. Magnetic resonance angiography in the evaluation of glomus tympanicum tumors. Am J Otolaryngol 1997:18;116–20.
35. Gacek RR. Evaluation and management of the temporal bone arachnoid granulations. Arch Otolaryngol Head Neck Surg 1992;118:327–32.
36. Møller AR, Jannetta PJ. On the origin of synkinesis in hemifacial spasm: results of intracranial recordings. J Neurosurg 1984;61:569–76.
37. Lunsford LD, Linskey ME. Stereotatic radiosurgery in the treatment of patients with acoustic tumors. Otolaryngol Clin North Am 1992;25:471–91.
38. Leskell L. A note on the treatment of acoustic tumors. Acta Chir Scand 1971;137:763–5.
39. Kamerer DB, Lunsford LD, Moller M. Gamma knife: an alternative treatment for acoustic neuromas. Ann Otol Rhinol Laryngol 1988;97:61–5.
40. Flinkinger JC, Lunsford LD, Linskey ME, et al. Gamma knife radiosurgery for acoustic tumors: multivariate analysis of four year results. Radiother Oncol 1993;27:91–8.
41. Maire JP, Floquet A, Darrouzet V, et al. Fractionated radiation therapy in the treatment of stage III and IV cerebellopontine angle neuromas: preliminary results in 20 cases. Int J Radiat Oncol Biol Phys 1992;23: 147–52.
42. Mendenhall WM, Friedman WA, Buatti JM, Bova FJ. Preliminary results of linear accelerator radiosurgery for acoustic schwannomas. J Neurosurg 1996;85: 1013–9.
43. Wiet RJ, Teixido M, Liang JG. Complications in acoustic neuroma surgery. Otolaryngol Clin North Am 1992;25:389–412.
44. Wiet RJ, Liang JG, Teixido MT. Electrocochleography and direct nerve monitoring during vestibular neurectomy. In: Wiet RJ, editor. Intraoperative auditory monitoring. Alexandria (VA): American Academy of Otolaryngol Head and Neck Surgery Foundation; 1992.
45. Sekhar S. Tumors of the cranial base. Diagnosis and treatment. Mount Kisco (NY): Futura Publishing; 1987.
46. Jackler RK, Brackmann DE. Neurotology. St. Louis: Mosby; 1994.
47. Glasscock ME, Kveton JF, Christiansen SG. Middle fossa vestibular neurectomy: an update. Otolaryngol Head Neck Surg 1984;92:216–20.
48. Monsell E, Wiet RJ, Young NM, Kazan RP. Surgical treatment of vertigo with retrolabyrinthine vestibular neurectomy. Laryngoscope 1988;98:835–9.
49. Wiet RJ, Mamikoglu B, Hoistad D, Battista R. A technique to prevent cerebrospinal fluid leakage after translabyrinthine approach. Laryngoscope 2000;110: 1234–6.
50. House WF, Hitselberger WE. The transcochlear approach to the skull base. Arch Otolaryngol 1976; 102:334–42.
51. Jenkins HA, Fisch U. The transotic approach to resection of difficult acoustic tumors of the cerebellopontine angle. Am J Otol 1980;2:70–6.
52. Gantz BJ, Fisch U. Modified transotic approach to the cerebellopontine angle. Arch Otolaryngol 1983; 109:252–6.

CHAPTER 26

Anatomy and Physiology of the Nose and Paranasal Sinuses

John Jacob Ballenger, MD

ANATOMY OF THE NOSE AND PARANASAL SINUSES

External Aspects of the Nose

From the tip or apex, the nose slopes upward and slightly posteriorly to reach the forehead at a junction marked by a slight depression termed the nasion (or radix nasi). Above each eye there are slightly projecting bony arches and between the arches a slightly flattened area known as the glabella. The rhinion is found at the lower end of the suture between the two nasal bones.

The membranous columella lies between the two nares and extends posteriorly from the nasal apex to the center of the upper lip. The philtrum, a gently rounded, convex, vertically oriented depression in the skin of the upper lip, extends upward from the center of the upper lip to the base of the membranous columella. On either side of the columella are the skin-lined nares, bounded by the lower lateral (alar) cartilages above, the floor of the nose below, and the columella medially. Just within the naris is a slight dilation, the nasal vestibule, lined by skin containing coarse hair or vibrissae and sebaceous and sweat glands.

The two nasal bones together with the two upper and lower lateral cartilages form the external framework of the nose (Figures 26–1 and 26–2). The caudal margins of the upper lateral cartilages usually overlie the upper margins of the lower lateral cartilages, and by elevating the tip of the nose,

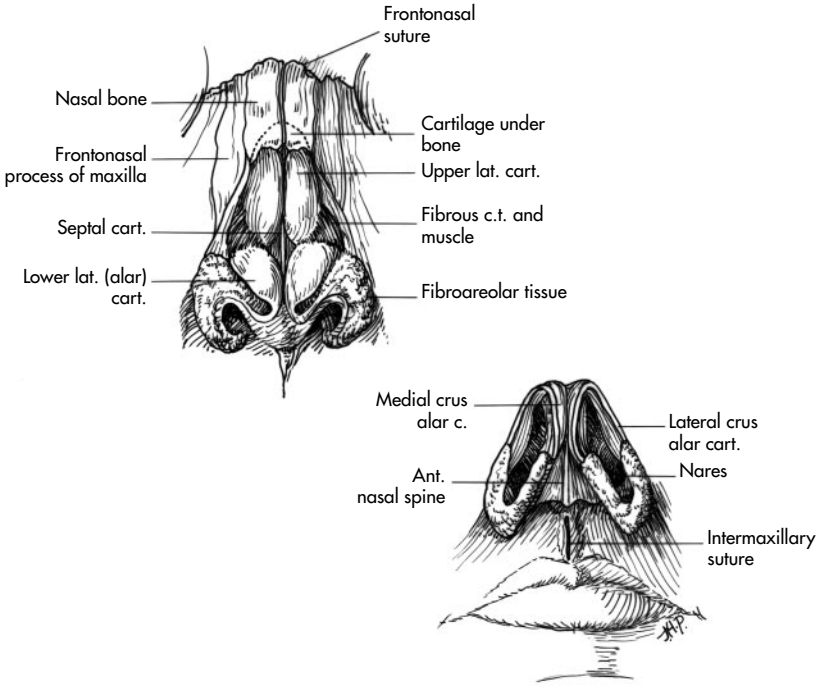


FIGURE 26–1. Dorsal and inferior surfaces of the nose.

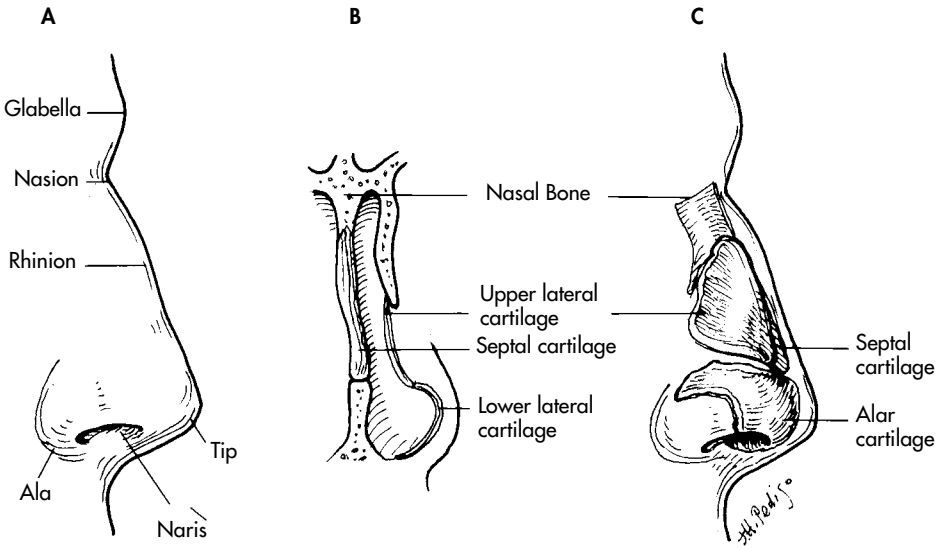


FIGURE 26–2. *A*, Lateral profile of the nose. *B*, Coronal section of the nose. *C*, Lateral view of nasal cartilages.

this margin can be demonstrated within the nose. The junction is called the limen nasi or nasal valve area; it is the narrowest part of the nasal airway and accounts for almost half of the total airflow resistance. At times, the medial margins of the lower lateral cartilages may lie close to but separate from the septum and thus provide less support for the nasal dorsum. Lying lateral to the two alar cartilages in the soft tissues of the nose are frequently found one or more unattached, small, functionless sesamoid cartilages.

The horseshoe-shaped lower lateral cartilage provides the framework of the naris. Its weak medial crus, together with its fellow from the other side, is incorporated into the substance of the columella.

In the bony skull, the pear-shaped opening to the nose is the pyriform aperature, and in the midline of the inferior border, a midline prominence, the anterior nasal spine, is located. The superior and lateral margins of the pyriform aperature are formed by the nasal bones and the frontal processes of the maxillae, and the base or floor by the alveolar processes of the maxillae.

There are two sets of alar muscles: the dilators (dilator naris, m. procerus, caput angulare) and constrictors (m. nasalis and depressor septi). They all receive innervation from cranial nerve VII.

Nasal Septum

The nasal septum is a midline structure derived from several bony and cartilaginous sources: superiorly and posteriorly by the perpendicular plate of the ethmoid bone and anteriorly by the septal (quadrilateral) cartilage, premaxilla, and membranous columella. Inferiorly, it is formed by the crests of the vomer, maxillary, and palatine bones and posteriorly by the sphenoidal crest (Figure 26–3).

Nasal Cavities: Nose

The palatal processes of the maxilla and horizontal processes of the palate bones form the floor of the nose. The roof of the nose is formed by alar cartilages, the nasal bones, the nasal processes of the frontal bones, and the bodies of the ethmoid and sphenoid bones. The cribriform plate (lamina cribrosa) forms the major portion of the roof of the nasal lumen. The inner surfaces of the maxillae, the lacrimal bones, the superior and middle turbinates, the inferior turbinate, and the medial pterygoid make up the lateral wall.

The three scroll-like, pitted turbinate bones, or the conchae, on each side of the nose divide the nasal lumen into meatus. The space between the inferior turbinate and floor of the nose is the inferior meatus, the space between the inferior and middle turbinates is the middle meatus, and the space above the middle turbinate is the superior meatus. Occasionally, there is a supreme turbinate. The middle and superior turbinates are extensions of the ethmoid bones, whereas the inferior turbinate is a separate bone attached by its superior border to the lateral nasal wall. At the anterior ends of the middle and inferior turbinates, a low cuboidal or squamous cell epithelium is found.

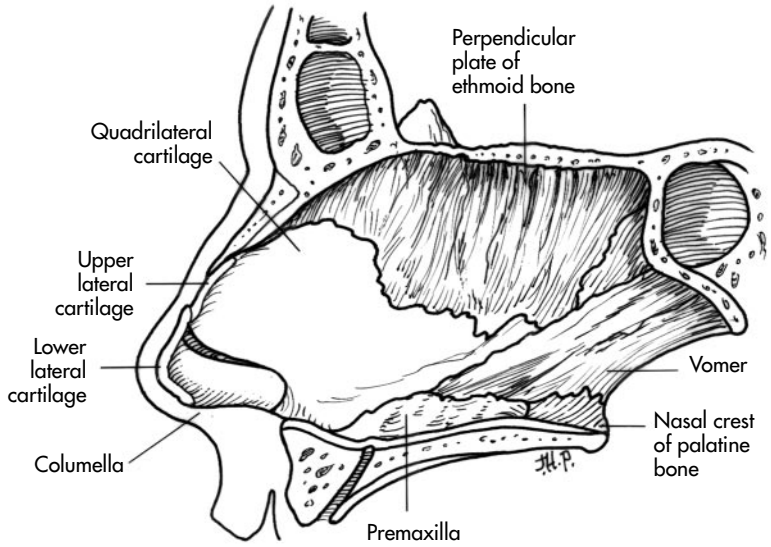


FIGURE 26–3. The nasal septum.

Posteriorly, the surfaces of both middle and inferior turbinates are covered with pseudostratified ciliated columnar (respiratory) epithelium. The epithelial stroma of the middle turbinate contains many glands. The large, tortuous, valveless, anastomosing veins, called sinusoids, are found mainly in the middle and inferior turbinates. By the degree of fluid contained in the sinusoids, they can influence the size of the nasal airway and, in effect, are capacitance structures. They respond to neural, mechanical, thermal, physiologic, and chemical stimuli.

Superior Meatus The superior meatus (also called the ethmoid fissure) is a slit-like space above the middle turbinate and is situated between the nasal septum and the ethmoid bone (Figure 26–4). The posterior ethmoid cells open into the central portion of this meatus. Above and posterior to the superior turbinate is the sphenoid recess, into which the sphenoid sinus opens.

Middle Meatus This meatus lies between the middle and inferior turbinates. It contains the orifices of the frontal and maxillary sinuses and also the anterior ethmoid cells. Hidden by the anterior half of the overhanging middle turbinate is a deep crescentic groove, the infundibulum. The crescent-shaped opening is called the hiatus semilunaris. The inferior medial wall of the infundibulum forms a shelf-like ledge known as the uncinate process and above the ledge a hemispheric prominence termed the ethmoid bulla.

The frontal, maxillary, and anterior ethmoid sinuses open into the infundibulum. Some ethmoid cells may open above the ethmoid bulla, and the frontonasal duct may have a separate opening. Additional details of endoscopic anatomy of the nose are given in Chapter 34.

Inferior Meatus The inferior meatus lies below the inferior turbinate. On its lateral surface, 3 to 5 cm beyond the naris, is found the orifice of the duct from the lacrimal gland. The floor of the meatus is congruent with the roof of the mouth.

Nares The anterior nares are formed medially, superiorly, and laterally by the lower lateral cartilages and the floor by the upper lip. Each naris measures about 1.5×1.0 cm, whereas the choanae are 2.5×1.5 cm. Just within the naris is a skin-lined enlargement, the nasal vestibule. The thin skin here contains coarse hair (the vibrissae) and sebaceous and sweat glands. The choanae are formed by the horizontal plate of the palatine bone inferiorly, the vomer medially, the vaginal process of the spenoid bone superiorly, and the medial pterygoid plate laterally.

Paranasal Sinuses

There are four paired paranasal sinuses: the maxillary, ethmoid, sphenoid, and frontal. In health, each is filled with air and communicates with the nasal lumen through an ostium. The bony ostium is slightly larger than the aperature in the mucous membrane. Each sinus is lined by a relatively thin, ciliated mucous membrane whose cilia beat the overlying blanket of mucus toward the sinus ostium

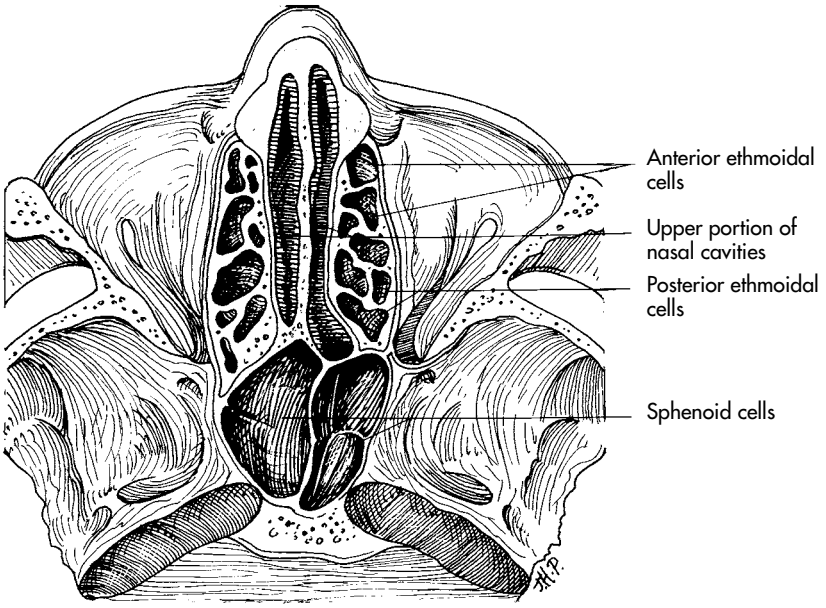


FIGURE 26–4. Horizontal section through the nose.

to join the mucous blanket in the nose. The cilia in the sinal mucous membrane become more numerous as the ostium is approached.

For clinical purposes, the sinuses are divided into two groups, anterior and posterior, depending on their location in reference to the line of attachment of the middle turbinate to the lateral wall of the nose. The anterior group consisting of the frontal, maxillary, and anterior ethmoid cells open into or near the infundibulum. The posterior group, made up of the posterior ethmoid cells and the sphenoid sinuses, opens above the middle turbinate.

Maxillary Sinus The growth of this sinus, the largest of the four, is biphasic (Table 26–1). The first period occurs during the first 3 years of life followed by a final period from ages 7 to 17 to 18 years. Much of the growth is an invasion into the alveolar process following eruption of permanent dentition.

The maxillary sinus (also called the antrum of Highmore) occupies the body of the maxilla. The sinus is generally pyramidal in shape, with its base formed by the lateral wall of the nasal cavity and its apex directed toward the zygomatic process. Its roof separates the sinus from the orbit. The floor is formed by the alveolar process of the maxilla and the hard palate. In children, the sinus floor lies at or above the level of the floor of the nasal interior, whereas in the adult, the sinus floor may lie 5 to 10 mm below. The anterior wall corresponds with the canine fossa and separates the sinus from the

TABLE 26–1. Maxillary Sinuses

Embryology	First of the paranasal sinuses to develop, beginning as a bud along the inferolateral surface of the ethmoidal portion of the nasal capsule about gestation day 65
Size	Birth: 7–8 × 4–6 mm
	Adult: 31–32 × 18–20 mm
	Volume (adult): 15 mL
Blood supply	Arterial branches of the internal maxillary including the infraorbital, lateral nasal branches of the sphenopalatine, descending palatine, and posterior and anterior superior alveolar arteries; most sinus walls drain into the maxillary vein, which shares communication with the pterygoid plexus
Innervation	Mucosal sensation from the lateroposterior nasal and superior alveolar branches of the infraorbital nerve, all derived from cranial nerve V_2

TABLE 26–2. Ethmoid Sinuses

Embryology	First appears during third and fourth fetal months as evaginations of the lateral nasal wall
Size	Adult: 20–24 × 20–24 × 10–12 mm (anterior group)
	20–21 × 20–22 × 10–12 mm (posterior group)
	10 to 12 cells each side
Blood supply	Nasal branches of the sphenopalatine artery and anterior and posterior ethmoid arteries, branches of the ophthalmic artery from the internal carotid system; maxillary vein, making connections with the cavernous sinus
Innervation	Posterior nasal branches of the maxillary nerve (cranial nerve V_2), the anterior and posterior ethmoidal branches of the ophthalmic nerve (cranial nerve V_1)

cheek. The posterior wall lies against the contents of the infratemporal space and pterygomaxillary fossa.

The antrum communicates with the infundibulum through an ostium located in the upper anterior part of the medial wall of the sinus. In 10 to 30%, an additional (accessory) ostium is present. The bony orifice is larger than the membranous one. The ostium serves as an entry conduit for most nerves and blood vessels.

The apices of the second upper bicuspid and first and second molar teeth are located in close relation to the floor of the sinus and may be separated only by mucous membrane, thus permitting easy spread of a dental infection into the sinus. The superior wall or roof is usually traversed in its central portion by the infraorbital nerve that may, at times, be protected only by a thin plate of bone.

Ethmoid Sinus At birth, usually three or four cells of the ethmoid sinus are present and, along with the maxillary sinus, are the only sinus cavities that are large enough to be of clinical importance (Table 26–2). The cells lie either in front of and below (anterior ethmoid cells) or posterior and above (posterior ethmoid cells) the attachment of the middle turbinate to the lateral nasal wall. They lie on either side of the superior halves of the nasal cavities and are separated from the orbits by the laminae papyraceae (Figures 26–5 and 26–6).

The anterior ethmoid cells open into the infundibulum of the middle meatus and the posterior cells into the superior meatus. The middle turbinate may be the site of an ethmoid cell, a concha bullosa, which can obstruct free sinus drainage but is usually asymptomatic. There are many variations of pneumatization of the ethmoid cells.

TABLE 26–3. Frontal Sinus

Embryology	Upward extension occurs at about 4 months gestation of the anterior portion of the nasal capsule in the region of the frontal recess
Size	Adult: 3 × 2.5 × 2 cm
	Adult volume: 6 to 7 mL
Blood supply	Supratrochlear and suborbital branches of the ophthalmic artery; venous drainage into the cavernous sinus
Innervation	Mucosal sensation derives from supratrochlear and supraorbital branches of the frontal nerve of cranial nerve V_1

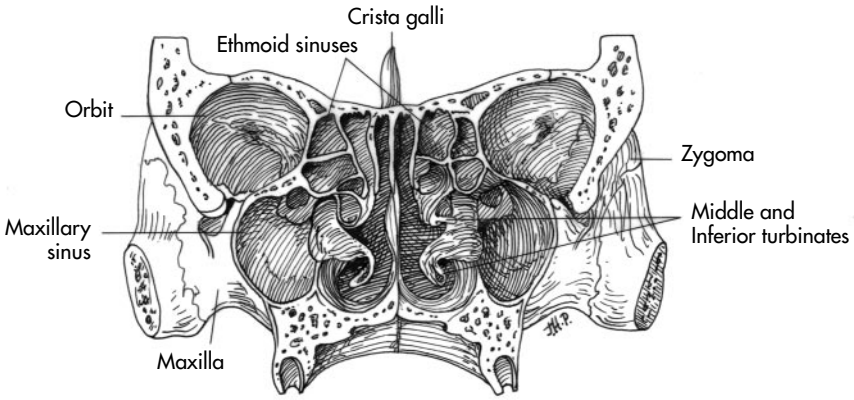


FIGURE 26–5. Coronal section through the nose.

Agar nasi cells are the most anterior of the anterior ethmoid cells and are located anterior to the anterosuperior attachment of the middle turbinate. They lie in close proximity to the frontal recess and can occasionally obstruct the nasofrontal duct. The anterior ethmoid cells frequently extend into the uncinate process and may invade the lumen of the frontal sinus as a frontal bulla or invade the body of sphenoid bone.

Haller cells are ethmoid cells that may obstruct maxillary ventilation in the region of the sinus ostium. They occur in about 10% of the population and may be without symptoms. Onodi cells are posterior ethmoid cells that can extend laterally or superiorly toward the nearby sphenoid sinus and lie near the optic nerve as well.

Frontal Sinus The frontal sinus can rarely be imaged before the second year of life (Table 26–3). At this time, it slowly invades the frontal bone and has much diversity of shape. The frontonasal duct enters the nose near the upper portion of the infundibulum. The anterior wall of the sinus is composed of diploic bone, and the posterior wall is composed of a compact, bony plate.

Sphenoid Sinus Pneumatization of the sphenoid bone occurs during middle childhood, proceeding rapidly after 7 years of age to its final form at 12 to 15 years (Table 26–4). Each sinus communicates with the sphenoethmoidal recess of the superior nasal meatus by means of a small aperature of 0.5 to 4 mm. The bony ostium is usually larger than the membranous and, for gravity drainage, is disadvantageously located 10 to 20 mm above the sinus floor.

Several vital structures lie closely adjacent to the sinus: the optic nerve and the hypophysis above, the pons posteriorly and external and lateral to the

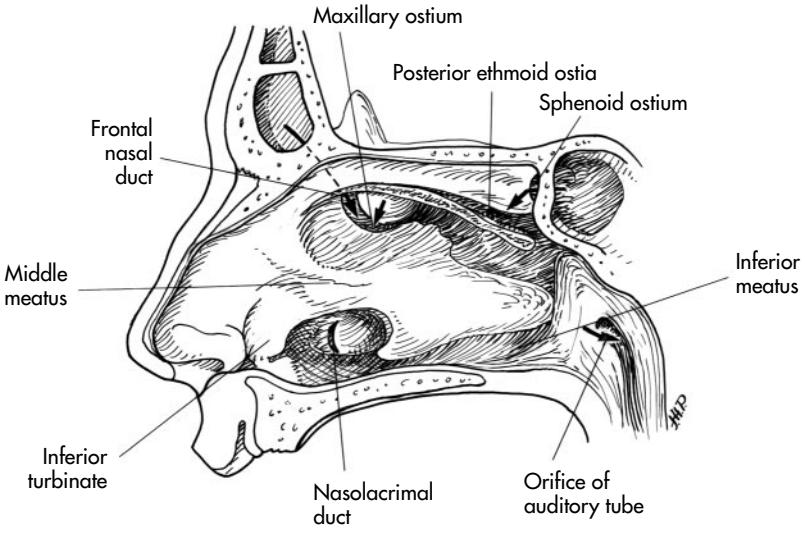


FIGURE 26–6. Schematic drawing of the right lateral nasal wall with the entire middle concha and the anterior third of the inferior part of the concha removed.

TABLE 26–4. Sphenoid Sinus

Embryology	Originates during third fetal month as an evagination of the mucosa in the sphenoethmoidal recess
Size	Adult: 20 × 22 × 16 mm; volume: 7.5 mL
Blood supply	Branches of the sphenopalatine and posterior ethmoid arteries
Innervation	Posterior ethmoid nerve from cranial nerve V_1, nasal and sphenopalatine branches of cranial nerve V_2

sinus, the cavernous sinus, the superior orbital fissure, the carotid artery, and several cranial nerves. The nerve of the pterygoid canal (vidian nerve) may encroach on the sinus floor.

Ostiomeatal Unit The ostiomeatal unit refers to the relationship between the middle meatus and anterior group of sinuses, particularly the anterior ethmoid cells (Figure 26–7). If there is an anatomic deformity (eg, concha bullosa) or a disease process that brings two mucosal surfaces into direct contact, local ciliary stasis occurs, and likely infection of one or more paranasal sinuses may be inaugurated. This pathogenesis is further described in Chapter 34.

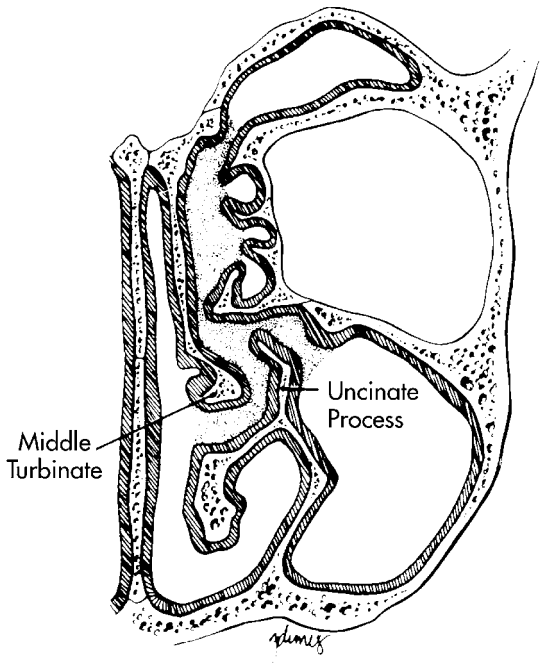


FIGURE 26–7. Concept of the ostiomeatal unit (*shaded*). (Courtesy of Dr. David W. Kennedy.)

Function of the Paranasal Sinuses

Prevailing theories of the function of the sinuses are numerous, but none are generally accepted. They include lightening the skull, a vocal resonating box, increased olfaction, humidification of the inspired air, and assistance in regulation of intranasal pressure.

It has been proposed that the sinuses provide a source of environmentally uncontaminated mucus that is delivered to the midportions of the middle and superior meatus and dilutes the contamination of the mucus more directly exposured to the incoming air.

Pterygopalatine Fossa

This space, an elongated triangular area with apex laterally, lies between the rounded posterior border of the maxillary sinus and the pterygoid process. It is bounded medially by the perpendicular plate of the palatal bone and superiorly by the undersurface of the sphenoid bone.

Just medial to the sphenopalatine foramen, an opening in the perpendicular plate of the palatal bone opposite the posterior end of the middle turbinate, is the sphenoethmoid recess. Through the foramen, vessels and nerves pass to the nasal cavity. The pterygopalatine (sphenopalatine) ganglion is located just lateral to the foramen.

Also communicating with the pterygopalatine space are the foramen rotundum and the pterygomaxillary and inferior orbital fissures. A wire passed up the greater or lesser palatine canals enters the fossa from below. Within the fossa are also found the second division of the fifth cranial nerve, the third division of the internal maxillary artery, and the vidian nerve. Infection into the pterygopalatine space may spread from the molar teeth. Entrance for drainage

can be made through the alveobuccal sulcus above the third molar tooth or through a Caldwell-Luc operation with removal of the posterior wall of the antrum and direct exposure into the pterygopalatine fossa.

Nasal Mucous Membrane

The skin within the nasal vestibule is a tough, keratinized, squamous cell epithelium containing coarse hairs (the vibrissae) and sebaceous and sweat glands. As the turbinates are approached, the epithelium blends first into a cuboidal cell type and then into the respiratory type. As the nasopharynx is reached, the respiratory type blends to a moist, nonkeratinized, squamous cell mucous membrane similar to that in the oral cavity. The nasal mucous membrane is further considered later in this chapter.

The mucosa of the paranasal sinuses contains pseudostratified ciliated columnar to cuboidal cell epithelium, is thin, and contains a few glands. The basement membrane is thin. Cilia, somewhat more abundant near the sinus ostia, propel the overlying blanket of mucus through the ostium, where it joins that in the nose.[1]

Nerve Supply of the Nose

The nerve supply of the nose (olfaction is considered in Chapter 27) consists mainly of ophthalmic and maxillary divisions of cranial nerve V. The former gives rise to the nasociliary nerve, dividing into the anterior and posterior ethmoid and infratrochlear branches. The anterior ethmoid nerve passes over the anterior end of the cribriform plate (Figures 26–8 and 26–9) and enters with the anterior ethmoid artery by way of the anterior ethmoid foramen, dividing into medial and lateral branches.

The medial branch passes forward and downward on the nasal septum and the lateral branch a similar part of the lateral wall. An external branch exits at the distal end of the nasal bone to reach the external surface of the nose. The posterior ethmoid crosses the cribriform plate to enter the nose with the artery through the posterior ethmoid foramen to the nasal septum as well as the olfactory region.

Branches from the maxillary division give rise to the posterior superior nasal nerves that enter the nose by way of the sphenopalatine foramen and pass over the anterior face of the sphenoid bone to reach the nasal septum as the nasopalatine nerve (n. of Cotunnius) and finally reach the incisive canal. Laterally, a branch, the posterior inferior, passes downward and forward to distribute on the middle and inferior conchae.

The autonomic innervation of the nose consists of parasympathetic and sympathetic fibers. The former originates in the superior salivary nucleus and travels via the nervus intermedius to the geniculate ganglion, where it joins with the greater superficial petrosal nerve. As this nerve leaves the temporal bone, it is joined by the deep petrosal nerve to form the nerve of the pterygoid

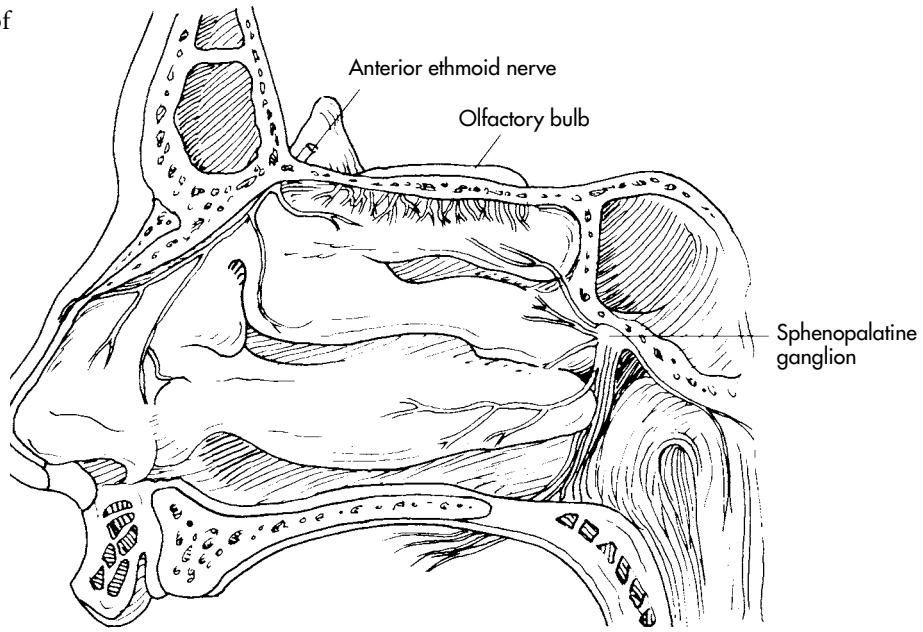


FIGURE 26–8. Nerve supply of the lateral wall of the nose.

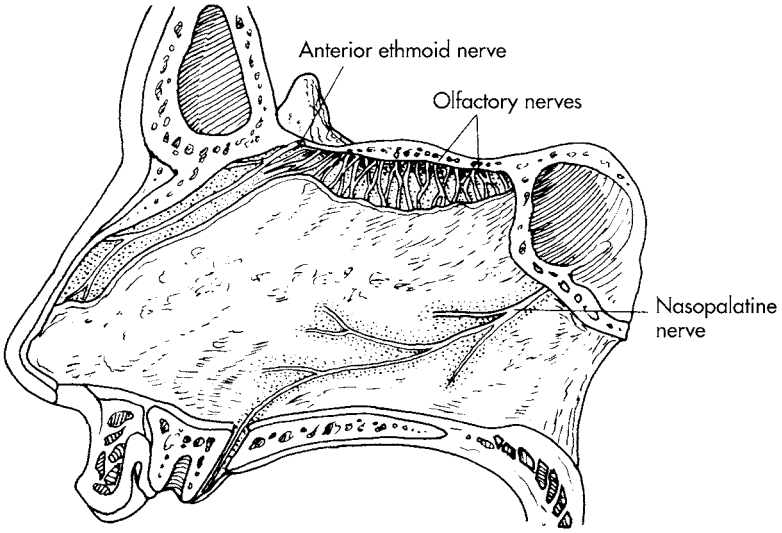


FIGURE 26–9. Nerve supply of the medial wall of the nose.

canal, or vidian nerve, to synapse in the sphenopalatine ganglion.

The postganglionic sympathetic fibers follow the internal carotid artery and form the deep petrosal nerve, which joins the greater superficial petrosal nerve to form the vidian nerve. Postganglionic sympathetic fibers pass through the sphenopalatine foramen without synapsing to innervate the nasal mucosa via the posterior nasal nerve.

Blood Supply of the Nose

The blood supply of the nasal interior comes from the anterior and posterior ethmoid branches (crossing the ethmoid plate) of the ophthalmic and sphenopalatine arteries. The latter is the terminal branch of the internal maxillary artery. The anterior and superior portions of the septum and the lateral wall of the nose receive blood via the anterior ethmoid arteries, the smaller posterior branch supplying only a small posterior region, including the olfactory area.

The internal maxillary artery, usually the penultimate branch of the external carotid artery, passes lateral to the pterygoid plate to enter the pterygoid fossa and continue as the sphenopalatine artery into the nasal cavity by way of the sphenopalatine foramen at the posterior end of the middle turbinate. Within the nose, the artery divides into posterior lateral nasal and posterior septal branches that accompany second and third divisions of the trigeminal nerve. There is an anastomosis between the lateral nasal arteries and the ethmoid; thus, bleeding can arise from either. Other branches of the sphenopalatine artery descend in the greater palatine canal to enter the oral cavity and spread over the undersurface of the palate.

Veins follow a course similar to that of the sphenopalatine artery and drain into the ophthalmic plexus and partly to the cavernous sinus. The nasal venous system is without valves and thus predisposes the spread of infection upward to the cavernous sinus.

Lymphatics

The lymphatics of the nasal vestibule drain forward toward the lip and external aspects of the nose. The lymphatics of the nasal fossa drain posteriorly, one collecting trunk in the olfactory region and a second below. They carry the lymph posteriorly to either the lateral retropharyngeal or subdigastric nodes.

PHYSIOLOGY OF THE NOSE

The nose is the entrance to the lower respiratory tract as well as to the olfactory epithelium (olfaction is discussed in Chapter 27). In its passage through the nose, the incoming air is prepared for reception by the alveolar sacs. In health, the nasal mucosa has the ability to cleanse itself and to condition the air.[2] The part played by the nose as a resonator is obvious to all who have suffered from the common cold.

Respiratory Mucosa

The pseudostratified respiratory mucosa consists of ciliated, intermediate, basal, and goblet cells. They rest on a well-defined basement membrane supported by a relatively deep, loose lamina propria

containing small blood vessels, venous plexuses, ducts of mucous and serous glands, sensory nerves, and blood cells (primarily lymphocytes). The blood capillaries and the venules are thin walled and possess a fenestrated endothelial lining and a porous basement membrane.

The tall (15 to 20 μm) columnar ciliated cell is the predominant cell and extends from the basement membrane to the luminal surface, where cilia admixed with microvilli are present. The microvilli are shorter than the cilia (3 μm × 0.1 μm versus 6 to 7 × 0.13), and some are branched. The microvilli contain bundles of microfilaments and display hair-like projections. The function of the microvilli is unknown, although it is clear that they greatly increase the cell surface area.

The ciliated cell cytoplasm forms complex interdigitations with adjacent cell membranes, presumably to permit intercellular exchange. Irregular intercellular spaces exist to accommodate edema fluid and inflammatory cells for implementation of the immune response.

Basal cells lie on the basement membrane and likely are progenitors of the columnar and goblet cells. Evidence suggests that the primary progenitor may be a nonciliated columnar cell that can form a ciliated cell. Goblet cells taper upward from the basement membrane to an expanded body at the lumen, where microvilli are found on their exposed surfaces. The nucleus is situated basally, and secretory granules that contain mucin are seen toward the lumen.

Columnar cells extend from a narrow base at the basement membrane to an expanded surface area covered by microvilli. These cells are related to adjacent cells by tight junctions apically and by interdigitations of the cell membrane. This cell may be the progenitor of the airway epithelium.

Airway

The nose provides a semirigid passageway for in- and outgoing air. On entering the nose, the air is directed upward by the nares (Figure 26–10). The air stream turns 80 to 90 degrees posteriorly as it reaches the nasal vault to traverse a mostly horizontal path until it impacts against the posterior wall of the nasopharynx. At this point, joined by the air stream from the other side, an 80- to 90-degree downward bend occurs. Each of these two bends, termed "impaction points," likely facilitates the removal of particulates contained in the incoming air. Impaction against the adenoid may enable the adenoid to respond immunologically by "sampling" the contaminants contained in the air. "Sniffing" draws the inspired air higher in the nasal cavities to reach the olfactory areas. The expiratory route is generally the reverse of the inspiratory and also may reach the olfactory area. In 1941, Proetz renewed interest in and advanced the study of nasal physiology.[3]

Air Stream

The anterior nasal valve, or ostium internum, is located at the limen nasi, some 1.5 to 2 cm posterior to the nares. At this point, the cross-section of the airway is 20 to 40 mm^2 on each side, is the narrowest part of the upper respiratory tract, and provides about 50% of the total airway resistance. Posterior to this in the horizontal section of the nasal airway, the cross-section widens while the air stream remains narrow and thus provides a large surface area in intimate contact with the air stream.

Evidence has accumulated that there is a 2½- to 4-hour cyclic alteration of nasal resistance from one side to the other.[4,5] A prolonged increase in nasal

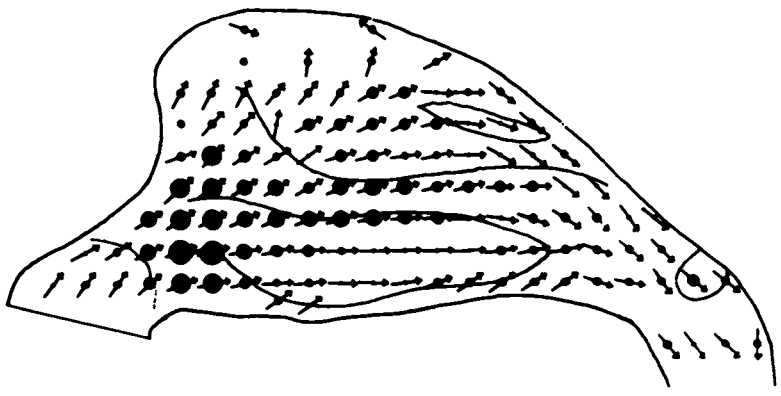


FIGURE 26–10. Diagram of the nasal airway and air speed, the size of the dot indicating the velocity. Reproduced with permission from Brau JD, Proctor DF, Reid LM, editors. Respiratory defense mechanisms. Part I. New York: Marcel Dekker; 1977.

resistance can lead to cor pulmonale, cardiomegaly, and pulmonary edema. The most common sequel to increased nasal resistance is mouth breathing, thus bypassing the air conditioning and cleansing functions of the nose.

Air Speed

The air speed is greatest at the anterior nasal valve, reaching 3.3 m/s at an inspiratory flow rate of 200 mL/s compared with 1 mm/sec in secondary bronchi (see Figure 26–10). Beyond the valve, the air speed slows down, thus enabling a longer contact between the incoming air and the nasal walls. At the choana, the stream again narrows.

Particle Removal

The aerodynamic equivalent diameter (AED) is the diameter of a unit density sphere with the same settling speed as the particle in question. Particles of approximately 5 μm AED or greater are 80 to 85% removed by the nose and nasopharynx. Smaller particles penetrate to varying degrees to the lower respiratory tract. Virus-containing droplets coalesce into diameters, frequently exceeding 5 to 6 μm, and thus are largely retained in the nose.

Air Conditioning

The air is heated (or cooled) by radiation from the mucosal blood vessels. Humidification occurs by evaporation from the mucous blanket. That this is an efficient mechanism is attested to by the fact that, in the nasopharynx, the inspired air is near normal body temperature and the relative humidity is near 100%. The mucosal blood vessels lie in two layers of more or less parallel rows. The more superficial layer sends capillaries into the epithelium, and the capillaries of the deeper layer near the basement membrane are fenestrated to facilitate fluid movement. The flow of blood is from posterior to anterior, opposite to the flow of inspired air and mucus. The mechanism of a "counter current" adds to the efficiency of the system.

The nasal mucous membrane is cooler by varying amounts than the expired air; thus, some condensation on and warming of the membrane occur —a so-called regenerative effect.

Respiratory Cilia

In humans, the respiratory cilia are found throughout the respiratory tract except for the nasal vestibule, the posterior oropharyngeal wall, portions of the larynx, and terminal ramifications of the bronchial tree. They are found in the eustachian tube, much of the middle ear, and the paranasal sinuses. Cilia in a modified form also occur in the maculae and cristae in the inner ear and in the eye as retinal rods.

Ciliary Ultrastructure Human cilia extend about 6 μm above the luminal surface of the cell and are about 0.3 μm in width. As many as 100 are found on each cell in the nose. Each cilium appears to be anchored to a basal body located just below the cell surface. The structure of the centriole of a dividing cell is similar to the basal body, the former giving rise to the latter and the latter to the cilium.

Each cilium is encased by an extension of the cell plasma membrane. Within the cilium is a sheaf of longitudinally arranged microtubules (or fibrils) termed an axoneme. The microtubules are, in fact, doublets or pairs with nine outer pairs arranged in a cartwheel pattern at the periphery of the axoneme. In addition, two single microtubules are located in

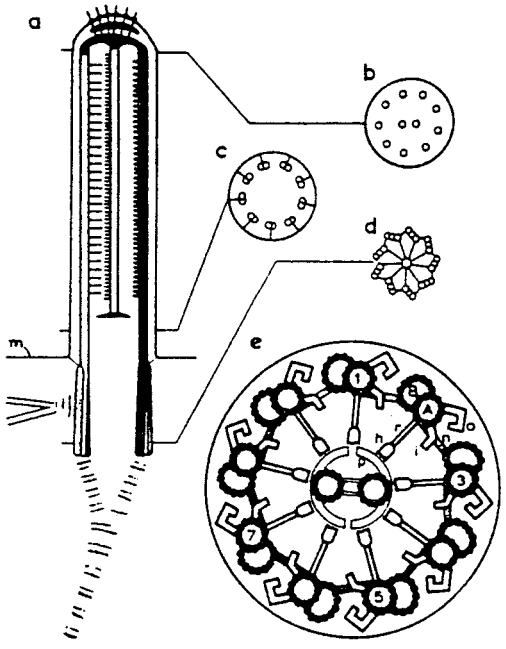


FIGURE 26–11. Diagram of the ultrastructural arrangement of ciliary tubules at various levels.

the center of the axoneme, creating the characteristic "9 plus 2" pattern (Figure 26–11).

At the tip of the cilium is a dense cap or crown from which three to seven "claws," 25 to 35 nm long, project. Below the cell membrane and the axoneme is a short, cylindrical basal body, and below this, the tubules (with a third added to form a triplet) seem to extend into the apical cytoplasm of the cell as a "rootlet." They converge into a cone-shaped form and acquire periodic striations. This structure, the basal foot, curves in the direction of the effective ciliary beat.

The basal foot has a cross-striated appearance resembling that of collagen fibers. In addition, other fine microtubules, which appear to be branched, attach to adjacent basal bodies, to each other, and ultimately to the junctional complex of the cell, constituting the terminal web.

Looking top-downward on the outer ring of doublets, each is composed of two juxtaposed microtubules—a slightly more centrally located subfiber, A, and a slightly more peripherally located subfiber, B (see Figure 26–11). Two regularly arranged arms extend from A toward B. They are composed of adenosine triphosphate and are called dynein arms. Also present are links attaching each subfiber A to B of the adjacent doublet. They are believed to be of an elastic material called nexin. Extending centrally from A to the central pair are radial spokes.

At the base of the cilium, the two central microtubules terminate, and each of the peripheral doublets continues downward to enter the basal body as a triplet with a subfiber, C, added.

Ciliary Beat The to-and-fro movement of the cilium is termed the ciliary beat (Figure 26–12). The forward beat is the more forceful, effective stroke in which the cilium is fully extended, and the claws at the tip penetrate the top layer of mucus bordering on the luminal surface and propel the mucus forward. The recovery stroke is less forceful and slower, and the shaft curls back on itself so that it does not reach the overlying flakes of mucus. Beating is metachronous and occurs 1,000 or more times per minute.

The motion of the cilium is caused by the sliding of one tubule against an adjacent one, thus creating a shearing force that induces the bending. The nature of the "nervous" connection between adjacent cilia is not clear; perhaps the "touch" of one cilium against the adjacent cilium is enough to initiate the coordinated, metachronous movement. Energy for this work is derived from the adenosine triphosphate found in the dynein arms. The "spokes" seem to detach and reattach several times during the bending process. The axis of motion of the cilium is defined by a line perpendicular to a plane connecting the central pair of tubules.

In health, particles resting on the mucous blanket are moved by active cilia at 3 to 255 mm per

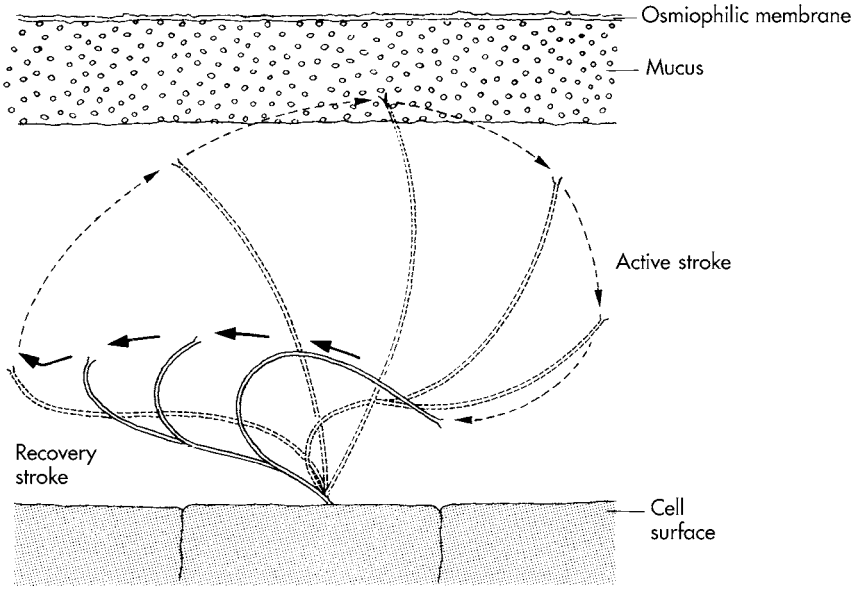


FIGURE 26–12. Diagram of the normal ciliary cycle.

minute, with an average of about 6 mm per minute. Apparent health exists, however, with speeds at some variance from this average. Dryness of the mucosa is quickly detrimental to ciliary activity. Other factors known to influence clearance speeds are the relative humidity and the pH of the fluids. β_2-Adrenergic agonists accelerate the wave frequency, whereas α_2-adrenergic activity retards the movement.

Numerous structural abnormalities of the axonemal tubules have been found (Figure 26–13).

Either a dye or saccharin can be used to measure the movement of material resting on the mucous blanket of the nose. Dye or saccharin is placed on the anterior aspect of the nasal mucosa and the time until the color of the dye is seen in the pharynx or the taste is reported by the subject is recorded. The use of tagged particles, developed by Quinlan et al, involves placing an anion exchange resin particle about 0.5 mm in diameter tagged with a technetium 99m ion on the anterior nasal ciliated mucosa, and its clearance can be followed accurately with a gamma camera or multicollimated detectors.[6]

Mucous Blanket

The bilayered mucous blanket, produced mostly by the serous and goblet glands, is a 12 to 15 μm thick, sticky, tenacious, adhesive sheet consisting of an upper, more tenacious mucus riding on the tips of the ciliary shafts and a deeper, thinner periciliary layer.[7] The blanket functions as a lubricant, protects against desiccation, and traps particulate matter and soluble gases. It amounts to 1 to 2 L per day.

In health, the pH is slightly acid. Its approximate composition is 2.5 to 3% glycoprotein, 1 to 2% salts, and 95% water. Immunoglobins comprise about 70% of the protein content. Mucus is found throughout the nose (except the vestibule), paranasal sinuses, middle ear, eustachian tube, and bronchial tree (and extending into the alveoli in the form of surfactant). The beating of the underlying cilia propels the blanket of mucus, along with trapped or dissolved material, in a more or less continuous movement toward the pharyngeal end of the esophagus.

Enveloping the shafts of the mucosal cilia is the thicker, less viscid, and deeper periciliary layer, and above this, interfacing with the luminal surface, is the more viscid layer of mucus riding on the periciliary fluid below. The mucus on the luminal surface may take the form of flakes. Soluble and insoluble matter caught in the mucus and on the mucous flakes are carried posteriorly by the movement of the mucous blanket to the upper end of the esophagus.

Mucociliary Transport

The mucociliary transport or clearance system is really two systems working in concert with one another. Its motive force is dependent on the actively beating cilia reaching the flakes of mucus at the luminal surface and propelling the flakes and surrounding mucus posteriorly to the esophagus. The mechanism by which the deeper periciliary, less viscid fluid, with its dissolved contaminants and likely viruses, also moves posteriorly is less well understood. More needs to be learned about the rheology and the effects of various viscoelastic properties, as well as the thickness of the mucous blanket under

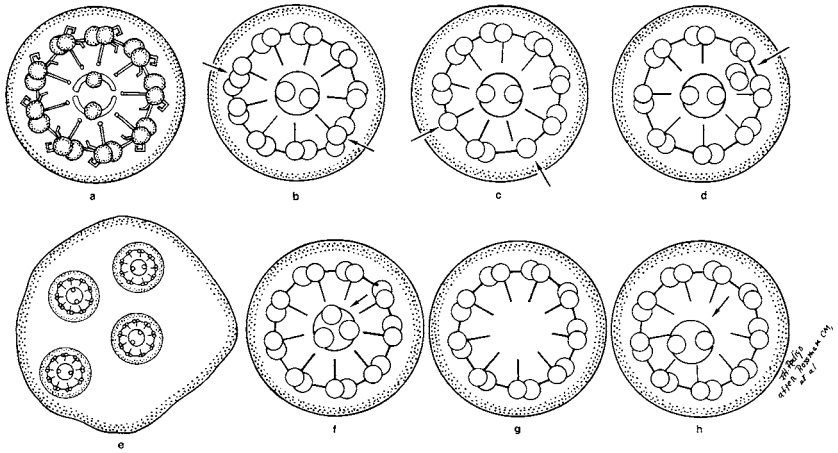


FIGURE 26–13. Some observed abnormal alterations of axonemal tubules.

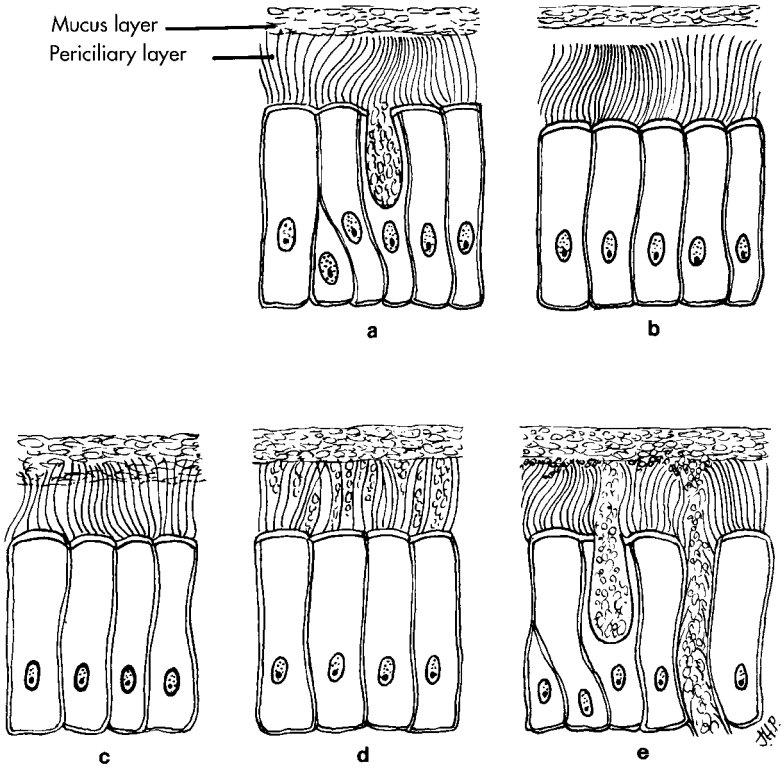


FIGURE 26–14. Various possible alterations of the mucous blanket.

varying conditions (Figure 26–14) and its relation to health and the cleansing function of the nose.

Although most bacteria seem to impede the metachronous ciliary cycle and efficient transport system little or not at all, *Bordetella pertussis*, *Mycoplasma pneumoniae*, and *Pseudomonas aeruginosa*, among others, do so. Some respiratory viruses and bacteria, notably influenza virus, rhinovirus, adenovirus, herpes simplex virus, and respiratory syncytial virus, seem to impede mucociliary transport by altering either the axonemal ultrastructure or the viscoelastic properties of the mucous blanket[8] (see Figure 26–13). The relationship of healthy mucociliary activity to disease is discussed further in Chapter 33.

REFERENCES

1. Baroody F, Naclerio RM. A review of the anatomy and physiology of the nose. Alexandria (VA): American Academy of Otolaryngology Head Neck Surgery; 1990.
2. Courtiss EH, Gargan TJ, Courtiss GB. Nasal physiology. Ann Plast Surg 1984;11:214–223.
3. Proetz AW. Essays on the applied physiology of the nose. St. Louis: Annals Publishing Co.; 1941.
4. Lund VJ. Objective assessment of nasal obstruction. Otolaryngol Clin North Am 1989;22:279–90.
5. Hasegawa M, Kern EB. The human nasal cycle. Mayo Clin Proc 1977;52:28–34.
6. Quinlan MF, Salman SD, Swift DL, et al. Measurement of mucociliary function in man. Ann Rev Respir Dis 1969;99:13–23.
7. Proctor DF. The mucociliary system. In: Proctor DF, Anerson IB, editors. The nose. New York: Elsevier Biomedical Press; 1982.
8. Grossman M. Clinical measurement of mucociliary clearance. In: Otolayrngology. Vol 2. Baltimore: Lippincott; 1994.

SUGGESTED READING

Alberti PW. Applied surgical anatomy of the maxillary sinus. Otolaryngol Clin North Am 1976;9:3–20.

Hollishead W. Anatomy for surgeons. Vol 1. Head and neck. 2nd ed. New York: Harper and Row; 1968.

Lang J. Clinical anatomy of the nose, nasal cavity, and paranasal sinuses. New York: Thieme; 1989.

Libersa C, Laude M, Libersa J. The pneumatization of the accessory sinuses of the nasal fossae during growth. Anat Clin 1981;2:265.

Peele J. Unusual anatomical variations of the sphenoid sinus. Laryngoscope 1957;67:208.

Schaeffer J. The clinical anatomy and development of the paranasal sinuses. Pa Med J 1936;39:395.

CHAPTER 27

Olfaction and Gustation

Richard L. Doty, PhD, Steven M. Bromley, MD

The chemical senses are important to humans as they determine the flavor of foods and beverages and provide a sensitive and early means for detecting dangerous environmental situations, including the presence of fire, spoiled food, and leaking natural gas. Unfortunately, taste and smell are typically ignored by otolaryngologists, even though (1) the stewardship of these senses falls within the purview of their specialty; (2) some otolaryngologic operative procedures compromise the functioning of these senses; (3) alterations in chemosensory function can be an early sign of a number of diseases, including Alzheimer's disease (AD) and idiopathic Parkinson's disease (PD); and (4) losses or distortions of chemosensation are of considerable personal and practical significance to the patient. The latter should not be underestimated and is particularly acute for patients whose lifestyle, livelihood, or immediate safety depends on smelling and tasting (eg, cooks, firemen, homemakers, plumbers, professional food and beverage tasters, employees of natural gas works, chemists, and numerous industrial workers).

This chapter provides an overview of the anatomy and physiology of the smell and taste systems, describes basic chemosensory pathology, and discusses up-to-date means for quantitatively assessing, managing, and treating disorders of these generally neglected sensory systems.

THE OLFACTORY SYSTEM

Anatomy and Physiology

Olfactory Neuroepithelium The olfactory neuroepithelium contains the olfactory receptors of cranial nerve I (CN I)* and lines the cribriform plate and sectors of the superior turbinate, middle turbinate, and septum. Although it reportedly comprises ~ 2 cm^2 of the upper recesses of each nasal chamber, it is not a homogenous structure, at least in the adult, as metaplastic islands of respiratory-like epithelium accumulate within its borders beginning early in life, presumably as a result of insults from viruses, bacterial agents, and toxins.[1] On the basis of immunohistochemical and anatomic criteria,[2] at least six major classes of cells can be identified in this neuroepithelium: bipolar sensory receptor cells, supporting or sustentacular cells, microvillar cells, Bowman's gland and duct cells, globose basal cells, and horizontal basal cells. The ~ 6 million *receptor cells* are derived embryologically from the olfactory placode and thus are of central nervous system (CNS) origin. The cilia of these cells, which extend into the mucus of the nasal lumen, harbor the seven-domain transmembrane olfactory receptors. The axons of these cells ultimately unite into bundles of 50 or so "fila" ensheathed by glial cells that traverse the cribriform plate to form the outermost layer of the olfactory bulb. The *sustentacular cells* insulate the receptor cells from one another and extend microvilli, rather than cilia, into the mucus. These cells contribute to the mucus of the region and may be involved to some degree in deactivating odorants and xenobiotic agents. The function of the ~ 600,000 *microvillar cells* located at the epithelial surface is unknown. Although earlier workers reported that an axon-like process projects from these cells to the olfactory bulb,[3] more recent studies have been unable to confirm such a projection.[4] *Bowman's*

*It should be noted that humans possess elements of two other intranasal neural systems in addition to CN I: the nervus terminalis or terminal nerve (CN O) and the trigeminal nerve (CN V). CN O, known for its high gonadotropin-releasing hormone content, is apparently unresponsive to volatile chemicals. Chemically induced CN V sensations include coolness, sharpness, and warmth. Although a rudimentary vomeronasal organ is present on each side of the base of the nasal septum in most persons, this organ is nonfunctional. Despite reports of chemically induced local potentials within this structure, it has no neural connection with the central nervous system, and no accessory olfactory bulb exists in humans.

glands are a major source of the mucus in the olfactory region, whereas the *globose and horizontal basal cells* are the progenitor cells of the other cell types[2] (Figure 27–1).

Olfactory Bulb and Olfactory Cortex The first processing station in the olfactory system, the olfactory bulb, is located directly over the cribriform plate. Its neural components are arranged in six concentric layers: the olfactory nerve, glomerular, external plexiform, mitral cell, internal plexiform, and granule cell (Figure 27–2). The receptor cell axons of the olfactory nerve layer enter the glomeruli within the second layer of the bulb, where they synapse with the dendrites of the mitral and tufted cells within the spherical glomeruli. These second-order cells, in turn, send collaterals that synapse within the periglomerular and external plexiform layers, resulting in "reverberating" circuits in which negative and positive feedback occur. Indeed, mitral cells modulate their own output by activating granule cells (which are inhibitory to them). Whereas the olfactory bulbs of younger persons have thousands of glomeruli arranged in single or double layers within the glomerular layer, older persons typically have far fewer numbers of glomeruli, reflecting the decrease in olfactory receptor cell numbers within the epithelium. After the age of 80 years, such structures are nearly absent.[5,6]

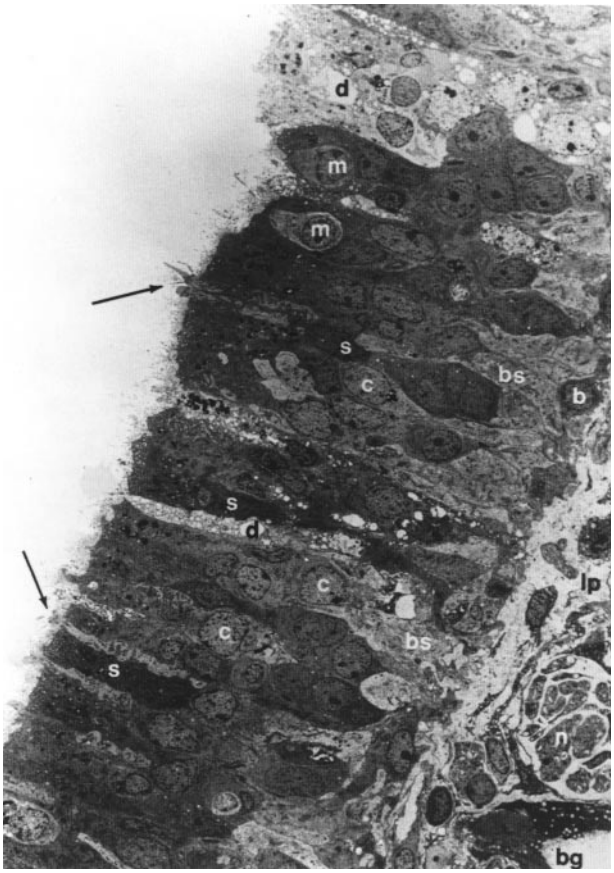


FIGURE 27–1. Low-power electron micrograph (original magnification × 670) of a longitudinal section through a biopsy specimen of human olfactory mucosa taken from the nasal septum. Four cell types are indicated: ciliated olfactory receptors (c), microvillar cells (m), supporting cells (s), and basal cells (b). The *arrows* point to ciliated olfactory knobs of the bipolar receptor cells. d = degenerating cells; bs = base of the supporting cells; lp = lamina propria; n = nerve bundle; bg = Bowman's gland. Courtesy of DT Moran.

The mitral and tufted cell axons project ipsilaterally to the primary olfactory cortex via the olfactory tract without an intervening thalamic synapse. The primary olfactory cortex is comprised of the anterior olfactory nucleus (AON), piriform cortex, olfactory tubercle, entorhinal area, periamygdaloid cortex, and corticomedial amygdala. Some projections occur, via the anterior commissure, from pyramidal cells of the AON to contralateral elements of the primary olfactory cortex. A number of projections from primary to secondary (ie, orbitofrontal) cortex are direct, whereas others relay within the thalamus. Recent functional imaging studies have found that odors reliably and significantly activate, in a concentration-dependent manner, posterolateral areas of the cerebellum, whereas sniffing alone tends to activate mainly anterior central cerebellar regions.[7] The cerebellar activity may reflect the circuits involved in modulating sniff size relative to the intensity of an odor or other movement-related behaviors.

Olfactory Transduction Humans can detect and discriminate among thousands of airborne odorants. Ten to 15% of the incoming air stream is shunted toward the olfactory cleft during inhalation (Figure 27–3). Some of the odorant molecules within this deflected air stream move from the air to the largely aqueous phase of the olfactory mucus, where they diffuse or are actively transported via "odorant binding proteins" to the olfactory receptors. Receptor activation then leads to transduction

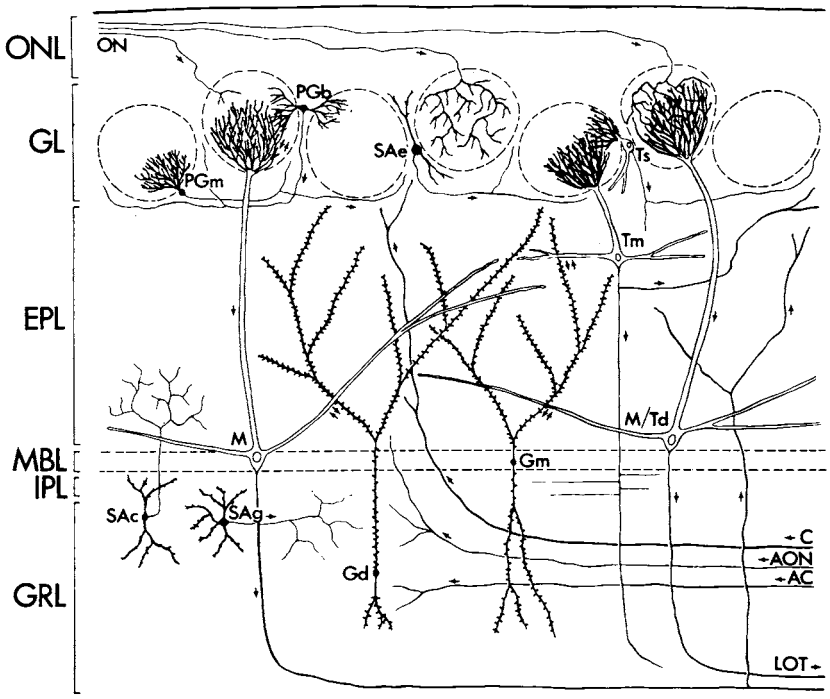


FIGURE 27–2. Diagram of major layers and types of olfactory bulb neurons in the mammalian olfactory bulb, as based on Golgi-stained material. Main layers are indicated on the left as follows: ONL = olfactory nerve layer; GL = glomerular layer; EPL = external plexiform layer; MBL = mitral cell body layer; IPL = internal plexiform layer; GRL = granule cell layer; ON = olfactory nerves; PGb = periglomerular cells with biglomerular dendrites; PGm = periglomerular cell with monoglomerular dendrites; SAe = short axon cell with extraglomerular dendrites; M = mitral cell; M/Td = displaced mitral or deep tufted cell; Tm = middle tufted cell; Ts = superficial tufted cell; Gm = granule cell with cell body in mitral cell body layer; Gd = granule cell with cell body in deep layers; SAc = short axon cell of Cajal; SAg = short axon cell of Golgi; C = centrifugal fibers; AON = fibers from anterior olfactory nucleus; AC = fibers from anterior commissure; LOT = lateral olfactory tract. Reproduced with permission from Shepherd GM.[8] Copyright © 1972, American Physiological Society.

cascades that produce action potentials within the olfactory receptor neurons.

Most olfactory receptors are representatives of a large (~ 1,000) multigene family of G protein–coupled seven-transmembrane receptors.[9] Each olfactory receptor neuron seems to express only one type of receptor, and neurons expressing the same gene appear randomly distributed within a few segregated strip-like "spatial zones" of the neuroepithelium, at least in the rodent. Each receptor binds a number of odorant molecules, although not all odorant molecules activate all receptors. Since the olfactory neurons that express a given receptor gene project to the same glomeruli of the olfactory bulb,[10] the glomeruli can be considered functional units. Data from a variety of sources suggest that it is the pattern across the activated receptors or glomeruli that serves as the proximal code for odorant quality.

The stimulatory guanine nucleotide-binding protein, G_{olf}, is activated by most olfactory receptor proteins.[11] In turn, G_{olf} induces production of the second messenger cyclic adenosine monophosphate by activating the enzyme adenyl cyclase. Cyclic adenosine monophosphate then diffuses through the cytoplasm, producing cellular depolarization by opening cyclic nucleotide–gated ionic channels and Ca^{2+}-dependent Cl^- or K^+ channels. Cyclic guanosine monophosphate is also activated by some odorants. It appears, among other things, to modulate the sensitivity of olfactory receptor neurons during adaptation.[12] G proteins other than G_{olf} (eg, G_{i2} and G_o) are present in olfactory receptor cells and aid in axonal

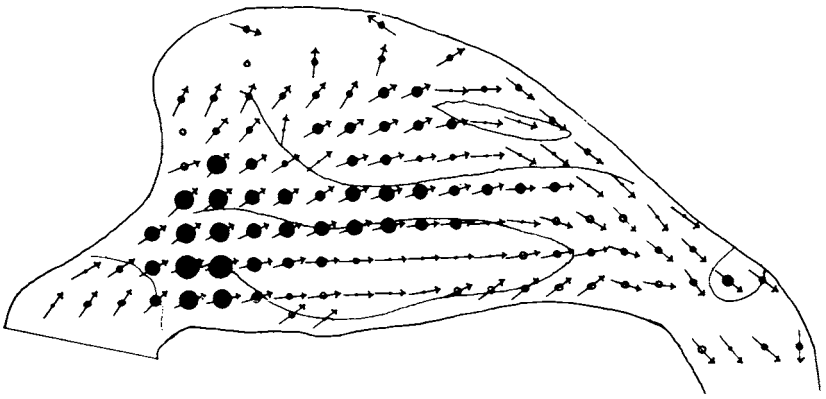


FIGURE 27–3. Pattern of the inspiratory nasal airflow as derived from studies in models. The external naris is to the left. *Arrows* indicate the direction of airflow, and *dot size* indicates the velocity. Reproduced with permission from Swift DL and Proctor DF.[13] Copyright © 1977 Marcel Dekker, Inc.

signal propagation, axon sorting, target innervation, and other such processes.[14]

Receptor Cell Regeneration The olfactory neuroepithelium has the ability to regenerate, although in cases in which significant damage to the basement membrane has occurred, regeneration is nonexistent or incomplete. Although, under normal circumstances, relatively continuous neurogenesis occurs within basal segments of the epithelium, many receptor cells are relatively long lived and appear to be replaced only after they are damaged.[15] Receptor cell death, as well as replenishment from progenitor cells, is determined by both endogenous and exogenous factors.[16] For example, differentiated neurons send regulatory signals that program the numbers of new neurons that need to be produced by the stem cells to maintain equilibrium in the cell population.[17] Apoptotic cell death occurs in cells representing all stages of regeneration, implying that biochemical regulation of neuronal numbers occurs at multiple stages of the neuronal lineage.[18]

Classification of Olfactory Disorders

Smell dysfunction can be reliably classified as follows: *anosmia*: inability to detect olfactory sensations (ie, absence of smell function); *partial anosmia*: ability to perceive some, but not all, such sensations; *hyposmia* or *microsmia*: decreased sensitivity to odors; *hyperosmia*: abnormally acute smell function; *dysosmia*: distorted or perverted smell perception to odor stimulation (sometimes termed cacosmia or parosmia, depending on the nature of the perversion); *phantosmia*: a dysosmic sensation perceived in the absence of an odor stimulus (ie, an olfactory hallucination); and *olfactory agnosia*: inability to recognize an odor sensation even though olfactory processing, language, and general intellectual functions are essentially intact, as in some stroke patients. Olfactory dysfunction can be either bilateral (binasal) or unilateral (uninasal), although, most commonly, such dysfunction is bilateral. Although presbyosmia is sometimes used to describe smell loss owing to aging, this term is less specific than those noted above (ie, it does not distinguish between anosmia and hyposmia) and is laden, by definition, with the notion that it is age per se that is causing the age-related deficit.

Clinical Evaluation of Olfactory Function

Three steps are involved in assessing a patient with chemosensory dysfunction: (1) obtaining a detailed clinical history, (2) testing olfactory function quantitatively, and (3) physically examining the head and neck.[19,20]

History Usually, precipitating antecedent events, such as head trauma, viral upper respiratory infections, allergies, toxic exposures, or iatrogenic (eg, operative) interventions, provide a basis for establishing causality, although in some patients a determination of the basis of the problem is complicated. When a clear-cut cause is not apparent, the clinician must explore a number of potential avenues to ascertain likely causal events or combinations of causal events. It is essential that potential underlying medical conditions known to have chemosensory consequences (eg, renal failure, liver disease, hypothyroidism, diabetes, or dementia) be identified

or ruled out. Questions regarding epistaxis, discharge (clear, purulent or bloody), nasal obstruction, allergies, and somatic symptoms, including headache or irritation, may have localizing value. A review of the history of tobacco, alcohol, or recreational drug use may also provide clues to etiology (eg, chronic alcoholism or Wernicke-Korsakoff syndrome). A detailed assessment of the medications being used prior to and during the onset of the dysfunction is important as many common medications (eg, antihypertensive and antilipid agents, antibiotics) can produce smell or taste disturbances. Importantly, medication-induced deficits may take some time to appear, so simply because a patient has been taking a drug for some time does not rule out its possible etiologic involvement. Likewise, some medication-induced alterations do not disappear immediately on drug cessation.

Specifics concerning the nature, timing of onset, and duration and pattern of fluctuations, if any, of the patient's chemosensory symptoms should be obtained. For example, *sudden olfactory loss* suggests the possibility of head trauma, infection, ischemia, or a psychogenic condition. *Gradual loss* can reflect the development of degenerative processes, progressive obstructive lesions or tumors within the olfactory receptor region or more central neural structures. *Intermittent loss* can be indicative of an intranasal inflammatory process.[20]

A family history of smell dysfunction may suggest a genetic basis. Delayed puberty in association with anosmia (with or without midline craniofacial abnormalities, deafness, and renal anomalies) suggests the possibility of Kallmann's syndrome or some variant thereof. Subtle signs of central tumors, dementia, parkinsonism, and seizure activity (eg, automatisms, occurrence of blackouts, auras, and déjà vu perceptions) should be sought in both the history and the physical examination.

Quantitative Olfactory Testing Many patients are inaccurate in describing their chemosensory function; some are unaware of a chemosensory deficit, whereas others overstate the nature of their problem. Hence, one should employ modern means for quantitatively assessing olfactory function in the office setting. Reliable quantitative testing is needed to (1) verify the validity of the patient's complaint, (2) characterize the exact nature and degree of the problem, (3) accurately monitor changes in function over time (including those resulting from therapeutic interventions), (4) detect malingering, and (5) obtain an objective basis for determining compensation for disability. In the past, many physicians have tested olfaction by simply asking the patient to identify several crude odorants, such as licorice, coffee grounds, or tobacco, placed under the nose. Unfortunately, such qualitative testing can lead to false positives (eg, patients have difficulty identifying odors without response alternatives), lacks reliability, has no normative reference, and is easily faked by malingerers.

During the last two decades, a number of standardized and practical psychophysical tests have been developed.[21] The most widely used test is the 40-item University of Pennsylvania Smell Identification Test (UPSIT; commercially known as the Smell Identification Test™; Figure 27–4).[22] This reliable (test–retest $r = .94$) test employs microencapsulated ("scratch and sniff") odorants and is available in English-, French-, German-, and Spanish-language versions. It can be self-administered in 10 to 15 minutes in the waiting room by most patients and scored in less than a minute by nonmedical personnel. In addition to providing a percentile rank of a patient's performance relative to age- and sex-matched controls, an absolute determination of normosmia, mild microsmia, moderate microsmia, severe microsmia, anosmia, or probable malingering can be made.

In some cases, unilateral testing is warranted, although usually olfactory problems are bilateral. To assess olfaction unilaterally, the naris contralateral to the tested side should be occluded without distorting the nasal valve region. This can be easily accomplished by sealing the contralateral naris using a piece of Microfoam™ tape (3M Corporation, Minneapolis, Minnesota) cut to fit the naris borders (Figure 27–5). The patient is instructed to sniff the stimulus normally and to exhale through the mouth. Such occlusion not only prevents air from entering the olfactory cleft from the contralateral naris (orthonasal stimulation), it also prevents active movement of odor-laden air into the occluded side from the nasopharynx on exhalation (retronasal stimulation).

The measurement of olfactory event-related potentials (OERPs) is available in only a few specialized centers. In essence, synchronized brain electroencephalographic (EEG) activity induced by repeated pulsatile presentations of an odorant is iso-

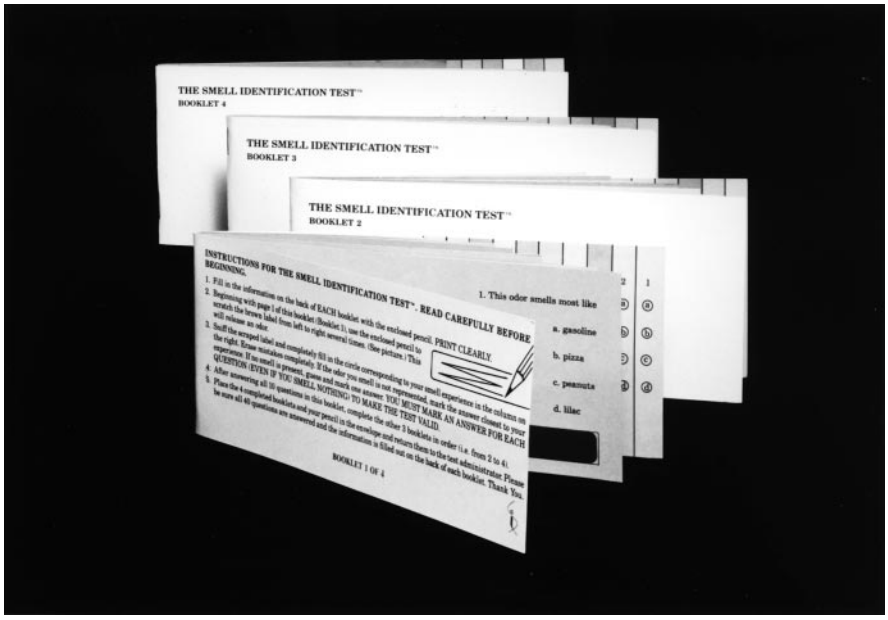


FIGURE 27–4. The four booklets of the 40-odorant University of Pennsylvania Smell Identification Test (UPSIT; commercially known as the Smell Identification Test™). Each page of each 10-page booklet contains a microencapsulated odorant that is released by means of a pencil tip and a multiple-choice question as to which of four possibilities smells most like the odorant. Forced-choice answers are recorded on columns on the last page of the test. Courtesy of Sensonics, Inc., Haddon Heights, New Jersey. Copyright © 2000, Sensonics, Inc.

lated from overall EEG activity. Averaging of responses from repetitive stimulation increases the signal-to-noise ratio. Although OERPs are relatively sensitive and useful in detecting malingering, they are presently unable to localize where in the olfactory pathway an anomaly exists, unlike their visual and auditory counterparts. Unfortunately, OERP testing requires specialized and expensive equipment capable of delivering odorant pulses with rapid rise times (~ 30 to 40 ms) to the olfactory neuroepithelium within a background of continuously flowing warmed and humidified air without inducing confounding somatosensory sensations.[23]

The electro-olfactogram (EOG) is another electrophysiologic measure of the olfactory system.[24] This surface potential, detected via an electrode placed on the surface of the olfactory neuroepithelium, reflects summated generator potentials mainly from olfactory receptor neurons. The recording of the EOG is, from a practical perspective, often more difficult than that of the OERP, and far fewer patients are amenable to such recording. The placement of the recording electrode is under endoscopic guidance, but since local anesthesia must be avoided, the placement of the electrode can be quite unpleasant, and sneezing and mucous discharge are common. Importantly, even after the electrode has been placed within the region of the olfactory epithelium, it cannot be recorded in many subjects. This may reflect the topographic distribution of specific olfactory receptors in combination with the relatively few odorants used or the presence of metaplasia of respiratory-like epithelium within the olfactory epithelium.

Physical Examination, Laboratory Tests, and Medical Imaging Careful otolaryngologic and neurologic assessment is warranted in patients complaining of olfactory dysfunction.[19,20] Visual field and acuity tests, as well as optic disk examinations, should be performed to determine whether intracranial mass lesions that produce increased intracranial pressure, papilledema, and optic atrophy are present (eg, Foster Kennedy syndrome, which consists of ipsilateral anosmia, ipsilateral optic atrophy, and contralateral papilledema owing to a meningioma of the ipsilateral optic nerve). With nasal endoscopy, employing both flexible and rigid endoscopes, specific attention should be paid to the olfactory meatal area. The nasal secretions should be carefully evaluated, as well as the color, swelling, surface texture, inflammation, erosion, ulceration, and atrophy of the nasal mucosa. Mucopus below the eustachian tube orifice suggests involvement of the osteomeatal complex, whereas mucopus above this orifice suggests posterior ethmoid and/or sphenoid sinus disease. Masses, polyps, and adhesions of the turbinates to the septum may all compromise airflow to the

olfactory receptor region. Unusual spaciousness, dryness, and crusting, as seen in atrophic rhinitis, suggest atrophy of the lamina propria. A pale mucous membrane can be indicative of allergy, usually resulting from edema within the lamina propria. Industrial or environmental pollutants, as well as excessive tobacco smoking, can produce metaplasia within the epithelium, in addition to swelling, inflammation, exudates, erosion, and ulceration.

Blood, serum, or other laboratory tests may help to identify or confirm underlying medical conditions that may relate to the dysfunction, including infection, nutritional deficiencies (eg, vitamins B_6 and B_{12}), allergy, diabetes mellitus, and thyroid, liver, and kidney disease. Although biopsies of the olfactory epithelium can be made, their interpretation is hindered by sampling issues and the fact that metaplasia of respiratory-like epithelium occurs throughout the olfactory epithelia even of persons with normal olfactory function.

Medical imaging can be invaluable in understanding the basis of a number of smell and taste disturbances. Magnetic resonance imaging (MRI) is the method of choice for evaluating soft tissue (eg, olfactory bulbs, tracts, and cortical parenchyma). Computed tomography (CT) is the most useful and cost-effective technique to assess sinonasal tract inflammatory disorders and is superior to MRI in the evaluation of bony structures adjacent to the olfactory pathways (eg, ethmoid, cribriform plate). Coronal CT scans are particularly useful in evaluating paranasal anatomy. Plain radiographs have substantial limitations and are rarely useful. Positron emission tomography, functional MRI (fMRI), and single-proton emission CT have limited usefulness at the present time outside of research institutions.

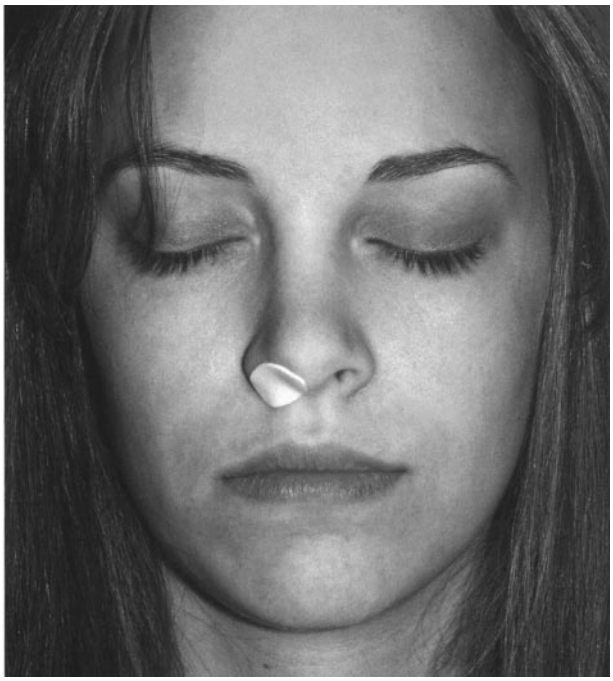


FIGURE 27–5. Picture of naris properly sealed with Microfoam™ tape for contralateral side testing. Photo courtesy of Sensonics, Inc., Haddon Heights, New Jersey. Copyright © 2000, Sensonics, Inc.

Detection of Malingering

Traditionally, it has been suggested that malingering can be detected reliably by having a patient inhale a strong CN V stimulant, such as ammonia, and asking the patient whether a smell is perceived. If denial occurs, the assumption is made that malingering is present. Unfortunately, this procedure is unreliable since usually CN V stimulants produce reflexive coughing, secretion from nasal mucous membrane, or other rejection reactions that the patient obviously cannot deny. Furthermore, CN V thresholds vary considerably, such that some patients experience little reaction to the ammonia and truthfully report perceiving no sensations.

A preferred method of the detection of malingering is to employ forced-choice psychophysical tests, such as the UPSIT. On such tests, malingering appears as the reporting of fewer correct responses than expected on the basis of chance (as would be expected in an anosmic). Since, in the case of the UPSIT, there are four response alternatives for each item and the patient must provide an answer even if no smell is perceived, 25% of the items, on average, should be correctly identified by chance alone (ie, 10 of 40). A sampling distribution exists around this expected probability, and empirical data are available on this point.[22,25] The theoretical probability of a true anosmic having an UPSIT score of 5 or less is less than 5 in 100. The theoretical probability of a true anosmic scoring 0 on the UPSIT is less than 1 in 100,000.[22] In general, if a patient scores within the probable malingering region of UPSIT scores, the UPSIT should be administered again to confirm the apparent avoidance of correct responses. Multiplication of the two probabilities is then used to establish

the statistical likelihood of malingering. Malingering is also suspected if the patient reports smell loss in the presence of a clear OERP, although olfactory agnosia in such cases cannot be ruled out.

It is noteworthy that malingering is sometimes discovered in head injury patients by their scores on forced-choice taste tests, rather than their scores on forced-choice smell tests, since bona fide smell loss is, in fact, present (negating the ability to avoid correct answers on the olfactory element of the examination). Such malingering implies relatively normal taste function and reflects the patient's naive attempt to embellish the "taste loss" that, in fact, results from lack of retronasal stimulation of the olfactory receptors.

Evidence for a general tendency to malinger can also be obtained using neuropsychological tests specifically designed for this purpose, for example, tests sensitive to head trauma patients trying to feign memory disturbances. Among those that are widely used is Rey's Memory Test (RMT), also known as Rey's 3 × 5 Test, and the Rey 15-item Memory Test.[26] The rationale behind this test is that malingerers typically fail at a memory task that all but the most developmentally disabled or severely brain-injured persons perform easily.

Causes of Olfactory Dysfunction

Olfactory dysfunction can result from three general causes: (a) *conductive* or *transport impairments* from nasal passage obstruction (eg, by chronic rhinosinusitis, polyposis, excessive mucus secretion, etc), (b) *sensorineural impairment* from injury to the olfactory neuroepithelium (eg, by viruses, airborne toxins, etc), and (c) *central olfactory neural impairment* from injury to CNS structures (eg, tumors, masses impinging on the olfactory tract, etc). These categories, however, are not mutually exclusive. For example, both blockage of airflow to the receptors and damage to the receptors and/or more central elements of the olfactory system can be simultaneously present or occur in stages. Thus, chronic rhinosinusitis can produce damage to the olfactory neuroepithelium in addition to blocking airflow, and altered neuroepithelium function can, over time, lead to degeneration within the olfactory bulb, a central structure.

There are numerous causes of olfactory disturbance (Table 27–1). Most cases of chronic anosmia or hyposmia are attributable to previous upper respiratory infections, head trauma, and nasal and paranasal sinus disease, reflecting long-lasting or permanent damage to the olfactory neuroepithelium.[27] Other causes include intranasal neoplasms (eg, inverting papillomas, hemangiomas, and esthesioneuroblastomas), intracranial tumors or lesions (eg, olfactory groove meningiomas, frontal lobe gliomas), neurologic diseases (eg, AD, idiopathic PD, multiple sclerosis [MS], schizophrenia), exposure to airborne toxins (including cigarette smoke), iatrogenic interventions (eg, septoplasty, rhinoplasty, turbinectomy, radiation therapy, medications), epilepsy, psychiatric disorders, and various endocrine and metabolic disorders.[†] The more common disorders or entities associated with smell loss are described in detail later in this section.

Most dysosmias reflect dynamic changes in the olfactory epithelium and remit over time.[27,28] Many anosmic patients report that prior to onset of their anosmia, they experienced a period of weeks or months when dysosmia was present. Usually, some smell function is present during the period of dysosmia. In rare instances, dysosmias present as aura-like hallucinations presumably associated with central (eg, temporal lobe) dysfunction. In many such cases, no seizure activity can be documented, and no evidence of CNS lesions or tumors is apparent. Nonetheless, low doses of anticonvulsant medication may be effective in mitigating the frequency and severity of some of these dysosmias. Dysosmias can also occur in a number of neurologic or psychiatric

[†]It has long been known that exogenous agents, including a number of neurotropic viruses, can enter the CNS from the nasal cavity via the olfactory nerve cells and that cauterization of the olfactory neuroepithelium can protect against infection from viruses instilled either intranasal or systemically.[29,30] In fact, otolaryngologists in the late 1930s prophylactically sprayed the noses of schoolchildren with zinc sulfate during poliomyelitis outbreaks.[31] The "olfactory vector hypothesis" has been proposed as an explanation for both the olfactory loss and the cause of several common neurologic diseases.[32–35]

TABLE 27–1. Agents, Diseases, Drugs, Interventions, and Other Etiologic Categories Associated in the Medical or Toxicologic Literature with Olfactory Dysfunction.*

- Air Pollutants and Industrial Dusts
  - Acetone
  - Acids (eg, sulfuric)
  - Ashes
  - Benzene
  - Benzol
  - Butyl acetate
  - Cadmium
  - Carbon disulfide
  - Cement
  - Chalk
  - Chlorine
  - Chromium
  - Coke/coal
  - Cotton
  - Cresol
  - Ethyl acetate
  - Ethyl and methyl acrylate
  - Flour
  - Formaldehyde
  - Grain
  - Hydrazine
  - Hydrogen selenide
  - Hydrogen sulfide
  - Iron carboxyl
  - Lead
  - Nickel
  - Nitrous gases
  - Paint solvents
  - Paper
  - Pepper
  - Peppermint oil
  - Phosphorus oxychloride
  - Potash
  - Silicone dioxide
  - Spices
  - Trichloroethylene
- Drugs
  - Adrenal steroids (chronic use)
  - Amino acids (excess)
    - Cysteine
    - Histidine
  - Analgesics
    - Antipyrine
  - Anesthetics, local
    - Cocaine HCl
    - Procaine HCl
    - Tetracaine HCl
  - Anticancer agents (eg, methotrexate)
  - Antihistamines (eg, chlorpheniramine malate)
  - Antimicrobials
    - Griseofulvin
    - Lincomycin
    - Macrolides
    - Neomycin
    - Pencillins
    - Streptomycin
    - Tetracyclines
    - Tyrothricin
  - Antirheumatics
    - Mercury/gold salts
    - D-Penicillamine
  - Antithyroids
    - Methimazole
    - Propylthiouracil
    - Thiouracil
  - Antivirals
  - Cardiovascular/antihypertensives
  - Gastric medications
    - Cimetidine
  - Hyperlipoproteinemia medications
    - Atorvastatin calcium (Lipitor)
    - Cholestyramine
    - Clofibrate
  - Intranasal saline solutions with
    - Acetylcholine
    - Acetyl-β-methacholine
    - Menthol
    - Strychnine
    - Zinc sulfate
  - Local vasoconstrictors
  - Opiates
    - Codeine
    - Hydromorphone HCl
    - Morphine
  - Sympathomimetics
    - Amphetamine sulfate
    - Fenbutrazate HCl
    - Phenmetrazine theoclate
- Endocrine/Metabolic
  - Addison's disease
  - Congenital adrenal hyperplasia

(continued on next page)

TABLE 27–1. *(continued)*

Endocrine/Metabolic *(continued)*
- Cushing's syndrome
- Diabetes mellitus
- Fröhlich's syndrome
- Gigantism
- Hypergonadotropic hypogonadism
- Hypothyroidism
- Kallmann's syndrome
- Pregnancy
- Panhypopituitarism
- Pseudohypoparathyroidism
- Sjögren's syndrome
- Turner's syndrome

Hallucinogenic Agents
- Lysergic acid diethylamide
- Psilocybin

Infections—Viral/Bacterial/Parasitic
- Acquired immune deficiency syndrome (AIDS)
- Acute viral rhinitis
- Bacterial rhinosinusitis
- Bronchiectasis
- Fungal
- Influenza
- Rickettsial
- Microfilarial

Lesions of the Nose/Airway Blockage
- Adenoid hypertrophy
- Allergic rhinitis
  - Perennial
  - Seasonal
- Atrophic rhinitis
- Chronic inflammatory rhinitis
- Hypertrophic rhinitis
- Nasal polyposis
- Rhinitis medicamentosa
- Structural abnormality
  - Deviated septum
  - Weakness of alae nasi
- Vasomotor rhinitis

Medical/Surgical Interventions
- Adrenalectomy
- Anesthesia
- Anterior craniotomy
- Arteriography
- Chemotherapy
- Frontal lobe resection
- Gastrectomy
- Hemodialysis
- Hypophysectomy
- Influenza vaccination
- Laryngectomy
- Oophorectomy
- Paranasal sinus exenteration
- Radiation therapy
- Rhinoplasty
- Temporal lobe resection
- Thyroidectomy

Neoplasms—Extranasal and Extracranial
- Breast
- Gastrointestinal tract
- Laryngeal
- Lung
- Ovary
- Testicular

Neoplasms—Intracranial
- Frontal lobe gliomas and other tumors
- Midline cranial tumors
  - Parasagittal meningiomas
  - Tumors of the corpus callosum
- Olfactory groove/cribriform plate meningiomas
- Osteomas
- Paraoptic chiasma tumors
  - Aneurysms
  - Craniopharyngioma
  - Pituitary tumors (especially adenomas)
  - Suprasellar cholesteatoma
  - Suprasellar meningioma
- Temporal lobe tumors

Neoplasms—Intranasal
- Neuro-olfactory tumors
  - Esthesioepithelioma
  - Esthesioneuroblastoma
  - Esthesioneurocytoma
  - Esthesioneuroepithelioma
- Other benign or malignant nasal tumors
  - Adenocarcinoma
  - Leukemic infiltration
  - Nasopharyngeal tumors with extension
  - Neurofibroma
  - Paranasal tumors with extension
  - Schwannoma

(continued on next page)

TABLE 27–1. *(continued)*

- Neurologic
  - Amyotrophic lateral sclerosis
  - Alzheimer's disease
  - Cerebral abscess (especially frontal or ethmoidal regions)
  - Down syndrome
  - Familial dysautonomia
  - Guam amyotrophic lateral sclerosis/Parkinson's disease/dementia
  - Head trauma
  - Huntington's disease
  - Hydrocephalus
  - Korsakoff's psychosis
  - Migraine
  - Meningitis
  - Multiple sclerosis
  - Myasthenia gravis
  - Paget's disease
  - Parkinson's disease
  - Refsum's disease
  - Restless legs syndrome
  - Syphilis
  - Syringomyelia
  - Temporal lobe epilepsy
    - Hamartomas
    - Mesial temporal sclerosis
    - Scars/previous infarcts
  - Vascular insufficiency/anoxia
    - Small multiple cerebrovascular accidents
    - Subclavian steal syndrome
    - Transient ischemic attacks
- Nutritional/Metabolic
  - Abetalipoproteinemia
  - Chronic alcoholism
  - Chronic renal failure
  - Cirrhosis of liver
  - Gout
  - Protein calorie malnutition
  - Total parenteral nutrition without adequate replacement
  - Trace metal deficiencies
    - Copper
    - Zinc
  - Whipple's disease
  - Vitamin deficiency
    - A
    - B_6
    - B_{12}
- Psychiatric
  - Anorexia nervosa (severe stage)
  - Attention-deficit disorder
  - Depressive disorders
  - Hysteria
  - Malingering
  - Olfactory reference syndrome
  - Schizophrenia
  - Seasonal affective disorder
- Pulmonary
  - Chronic obstructive pulmonary disease

*Categories are not mutually exclusive.
Adapted from Doty RL et al.[36]

disturbances, which are usually diagnosed on other grounds (eg, psychosis, MS). Infrequently, dysosmia may be owing to the perception of foul odors produced by the body, such as those from purulent nasal secretions in sinusitis or from exhalations in halitosis or uremia. Other disorders that may present as dysosmia include trimethylaminuria (fish odor syndrome) and cat odor syndrome, a pediatric neurologic disorder associated with a ß-methylcrotonyl-CoA carboxylase deficiency. Such rare disorders usually exist in the presence of a normally functioning olfactory system.

In contrast to cases of anosmia, hyposmia, and dysosmia, cases of hyperosmia are rare. Although untreated adrenal cortical insufficiency has been reported to produce hyperosmia in humans, this finding has yet to be confirmed, and no evidence for hypersensitivity following adrenalectomy has been found in animal studies.[37] There have been suggestions of hyperosmia in syndromes such as multiple chemical sensitivity, but the limited data available also fail to support this notion.[38] Hyperosmia reportedly occurs in some cases of epilepsy during the interictal period, although most patients with

long-term epilepsy and intractable seizure activity, such as candidates for temporal lobe resection, are hyposmic.[39]

Upper Respiratory Infections Upper respiratory system viruses (eg, influenza, colds) are the most common cause of *permanent* smell loss in adulthood.[27] Factors that predispose individuals to virus-induced smell dysfunction and the mechanisms underlying it remain unclear, although direct insult to an already compromised neuroepithelium is most likely, and, in rare cases, central structures may also become involved. If neural damage is present, topical or systemic corticosteroid treatment is ineffective. Reduced numbers of olfactory receptor cells and other epithelial abnormalities are commonly found in biopsies taken from the olfactory epithelia of patients with postviral anosmia or hyposmia.

Head Trauma Smell loss or distortion occurs in ~ 15% of patients with significant head trauma but can be present even in mild cases in which rapid acceleration/deceleration of the brain has occurred (coup contra coup injury) (Figure 27–6). In general, the severity of trauma is roughly correlated with the degree of loss. Occipital blows tend to produce more frequent and more severe olfactory deficits than frontal blows.[40] The physiologic mechanisms involved include shearing of the olfactory fila and contusions of the olfactory bulb and frontal and temporal poles.[41] Although the cribriform plate may become fractured in some cases, such fractures are not a prerequisite for smell loss.[40] In a recent study of 268 patients evaluated at the University of Pennsylvania Smell and Taste Center who had experienced head trauma, 66.8% had anosmia and 20.5% hyposmia. Of 66 patients retested 1 month to 13 years later, only 3, none of whom initially had anosmia, regained normal olfactory function. Dysosmia prevalence decreased from 41.1 to 15.4% over post-trauma periods averaging several years.[40]

Even though the loss of smell following head trauma is usually immediate, this is not always the case, and it may take the patient a while to recognize the presence of the loss. In some cases, delayed loss reflects delayed receptor cell death. Rodent research shows that intracranial hemorrhage and ischemia can lead to degeneration of the olfactory epithelium without transection of the olfactory nerves.[42]

Nasal and Sinus Disease It is now apparent that the olfactory loss associated with nasal or sinus disease is not solely caused by decreased conduction of airflow to the olfactory receptors. Although medical (eg, administration of topical or systemic corticosteroids) or surgical (e.g., excision of polyps) treatment can improve olfactory function in some cases, return to normal levels is not the norm.[43] In the case of rhinosinusitis, for example, factors other than, or in addition to, nasal airflow blockage are responsible for the loss. Chronic inflammation is, in fact, likely toxic to olfactory neurons. Hence, many cases of rhi-

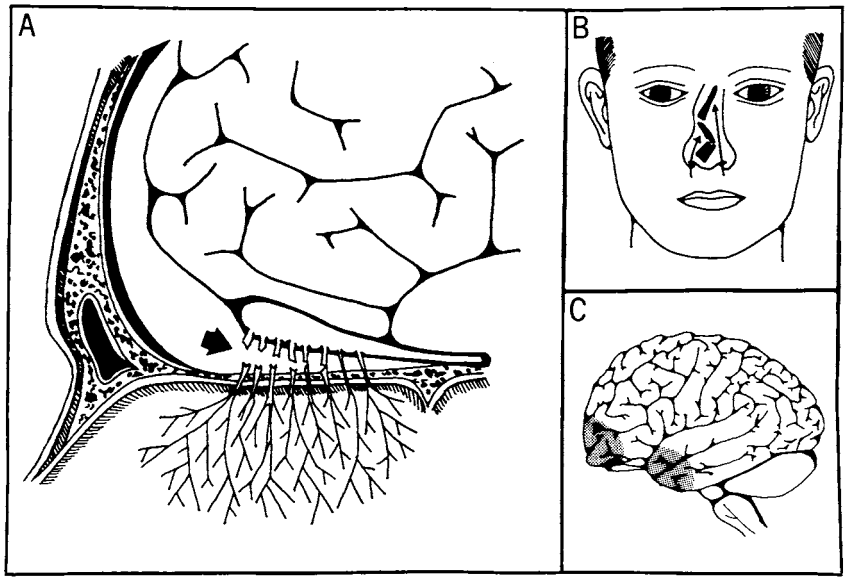


FIGURE 27–6. Mechanisms of post-traumatic olfactory dysfunction. *A*, Tearing of the olfactory fila. *B*, Injury to the sinonasal tract. *C*, Cortical contusions and brain hemorrhage. Reproduced with permission. Copyright © 1991, Lippincott Williams and Wilkins.

nosinusitis have a significant sensorineural component. The severity of histopathology within the olfactory epithelium of chronic rhinosinusitis patients is correlated with the degree of smell loss, as measured by the UPSIT.[44] Olfactory biopsies from patients with nasal disease are less likely to yield olfactory epithelial tissue than biopsies from controls.[45] The same is true for anosmic versus nonanosmic rhinosinusitis patients.[46]

Central Nervous System Neoplasms Olfactory disturbances can arise from tumors impinging on the olfactory bulbs or tracts, such as frontal lobe gliomas, olfactory groove meningiomas, and suprasellar ridge meningiomas arising from the dura mater of the cribriform plate, as well as from tumors on the floor of the third ventricle, pituitary tumors extending above the sella turcica, and tumors in the temporal lobe or uncinate convolution.[19] Foster Kennedy syndrome can result from tumors impinging on the bulb or tract.[47] *Pseudo* Foster Kennedy syndrome has been found in patients with increased intracranial pressure who had previous unilateral optic atrophy.[48]

Aging Age-related smell loss is well documented and occurs in most older people, including those who are healthy and taking no medications (Figure 27–7).[49] Such loss clearly impacts on the quality of life of the elderly, influencing nutrition, appetite, and, possibly in some cases, even general immunity and defensive responses to illnesses.[50,51] About one half of the population between 65 and 80 years of age experiences significant decrements in the ability to smell. Over the age of 80, this figure rises to nearly 75%. This is in marked contrast to persons under the age of 65 years, of whom only 1 to 2% suffers from major difficulty smelling.[49]

Despite the association with age, smell loss in the later years should not be attributed simply to age, per se, as often an accumulation of damage over the years is the culprit, and a single event, such as a bad cold, can be the precipitating factor. In general, the age-related changes in smell function reflect decrements in both olfactory receptors and the number of olfactory bulb glomeruli.[6] Interestingly, recent data suggest that age-related closure of cribriform plate foramina by ossification is common in skulls from older persons.[52] Whether other age-related factors increase the susceptibility of the epithelium to damage from exogenous agents is not clear.

Neurodegenerative and Other Neurologic Diseases An exciting chapter in the study of olfaction has been the discovery that a number of neurologic disorders are commonly accompanied by olfactory deficits, including AD, idiopathic PD, Huntington's disease, alcoholic Korsakoff's syndrome, Pick's disease, the parkinsonism dementia complex of Guam,

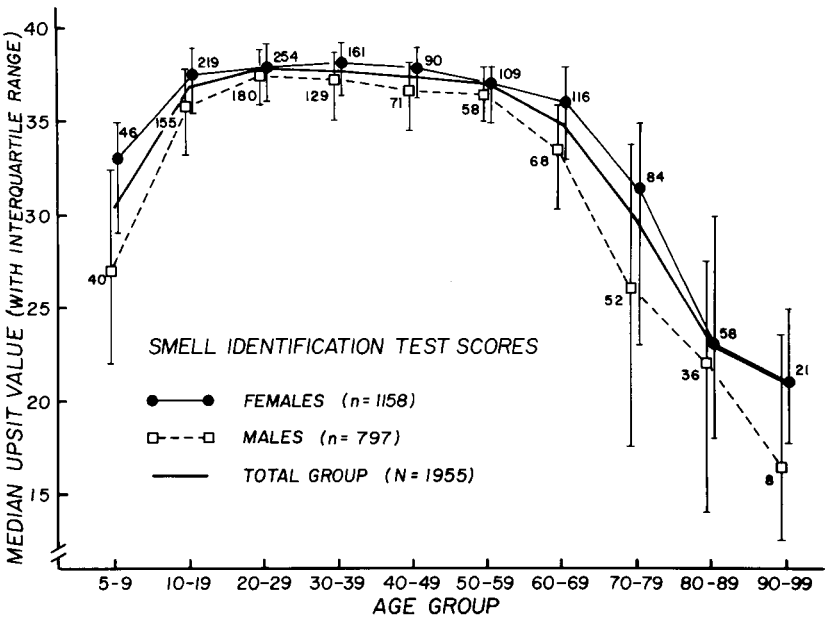


FIGURE 27–7. Scores on the University of Pennsylvania Smell Identification Test (UPSIT) as a function of age and gender in a large heterogeneous group of subjects. Numbers by data points indicate sample sizes. Reproduced with permission from Doty RL, Shaman P, Applebaum SL, et al. Smell identification ability: changes with age. Science 1984;226:1441–3. Copyright © 1984, American Association for the Advancement of Science.

amyotrophic lateral sclerosis, schizophrenia, and MS.[32,53–55] Indeed, olfactory dysfunction appears to be the first clinical sign of AD and idiopathic PD.[56,57] Although usually considered a neurodevelopmental disorder, smell loss is present in patients with schizophrenia and appears to be correlated with disease duration, suggesting a possible degenerative component to this disorder in olfaction-related pathways.[55] It has recently been shown that patients with schizophrenia have much smaller olfactory bulbs and tracts than those of matched controls.[58]

Several studies provide data suggesting that smell testing is useful in identifying persons at risk for later significant cognitive decline or AD. Graves et al administered a 12-item abbreviated version of the UPSIT (termed the Brief-Smell Identification Test™ or B-SIT) and several cognitive tests to 1,985 Japanese American people around the age of 60 and then readministered these tests to 1,604 of these people 2 years later.[59] Sixty-nine percent of the follow-up participants were genotyped for apolipoprotein E (apoE). Low B-SIT scores in the presence of one or more APOE-epsilon 4 alleles were associated with a very high risk of subsequent cognitive decline, and the B-SIT identified persons who later came to exhibit cognitive decline better than a global cognitive test did.

More recently, the UPSIT was administered by Devanand et al to 90 outpatients with mild cognitive impairment and to matched healthy controls at 6-month intervals over a 5-year period.[60] Patients with mild cognitive impairment had lower UPSIT scores than did controls. Most importantly, patients with low UPSIT scores (< 34) were more likely to develop AD than the other patients. Low UPSIT scores, combined with lack of awareness of olfactory deficits on the part of the patients, predicted the time to development of AD. UPSIT scores from 30 to 35 showed moderate to strong sensitivity and specificity for diagnosis of AD at follow-up.

The underlying cause for the olfactory deficit of AD is not yet clear, even though AD is associated with loss of neurons in the anterior olfactory nucleus, olfactory bulb, and layer II of the entorhinal cortex.[32,61] Disproportionate numbers of neurofibrillary plaques and tangles are found within limbic brain regions that receive olfactory bulb projections. Central cholinergic deficits may also be involved since (1) individuals with no history of cognitive loss who are in the early histopathologic stages of AD exhibit a cholinergic deficit within the inferior temporal lobe and (2) drugs that alter cholinergic function alter the ability to smell. For example, the cholinesterase inhibitor physostigmine improves odor discrimination performance, at least in rats,[62] whereas scopolamine, a muscarinic cholinergic antagonist, reportedly decreases olfactory sensitivity in humans.[63]

In idiopathic PD, bilateral olfactory deficits occur early on in the disease, such as at the onset of hemiparkinsonism, and occur at a higher frequency than some cardinal signs of this disorder (eg, tremor).[57] The olfactory impairment is unrelated to use of antiparkinsonian medications, duration of illness, and severity of the symptoms and signs, such as tremor, rigidity, bradykinesia, or gait disturbance.[64] Since smell loss is absent or infrequent in a number of other neurologic disorders that exhibit similar motoric signs, smell testing can aid in differential diagnosis in some cases.[32] For example, patients with essential tremor, progressive supranuclear palsy, multiple system atrophy, and parkinsonism induced by the proneurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine exhibit little or no olfactory dysfunction.[33] It is of interest that familial PD is also associated with olfactory impairment that occurs independently of the parkinsonian phenotype.[65]

It has generally been assumed that olfactory function in MS is normal since the axons of the primary olfactory receptors are unmyelinated. However, myelin is present in other segments of the olfactory pathway, and we have recently shown that, at any one time, about a third of MS patients exhibit some degree of smell loss.[66] Furthermore, we have found that the degree of olfactory loss is directly related to the number of MRI-determined plaques in central brain regions associated with olfactory processing (eg, inferior middle temporal lobe and periorbital frontal cortex) (Figure 27–8).[54,66] Indeed, a seemingly one-to-one longitudinal association is present between UPSIT scores and changes in plaque load over time (Figure 27–9),[67] implying that olfactory function waxes and wanes as the plaque numbers increase and decrease. In effect, knowledge of a patient's UPSIT score accurately predicts the plaque load in the olfaction-related regions and vice versa.

Exposure to Neurotoxic Agents A number of environmental and industrial chemicals have been linked to olfactory dysfunction, including acrylates,

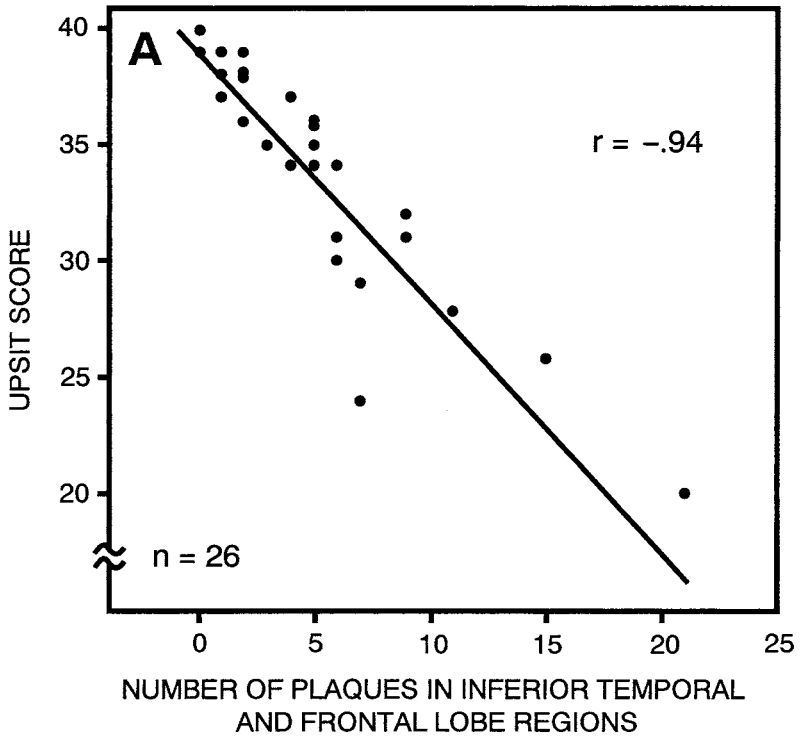


FIGURE 27–8. Relationship between the number of multiple sclerosis–related plaques in subtemporal and subfrontal lobes and scores on the University of Pennsylvania Smell Identification Test (UPSIT). No such relationship was present between UPSIT scores and plaques in brain regions outside of the primary olfactory cortical areas ($r = –.08$). Reproduced with permission from Doty RL, Li C, Mannon LJ, Yousem DM. Olfactory dysfunction in multiple sclerosis. N Engl J Med 1997;336:1918–9.

benzene, cadmium, cigarette smoke, formaldehyde, solvents, and nickel dust, among others, although few well-controlled studies exist in this area, and most reports are largely anecdotal.[68] In many cases, the decrements are specific to the exposed compounds and reflect long-term adaptation rather than damage to the olfactory receptors, being reversed after the worker is away from the workplace for a relatively short period of time. In the case of cigarette smoking, olfactory ability decreases as a function of cumulative smoking dose. Long-term cessation of smoking leads to a gradual improvement of olfactory function that is inversely related to the amount and duration of prior smoking activity.[69]

Other Causes As noted earlier in this chapter, medications commonly affect smell function and should be considered early in an evaluation, especially in the context of a new drug therapy (see Table 27–1). Olfactory hallucinations occur in mesial temporal lobe seizures and migraine, as well as in some other central brain lesion disorders. In epilepsy, the hallucinations are usually unpleasant and are rarely isolated events.[39]

Prognosis and Treatment of Olfactory Disorders

Prognosis seems to be better for patients with less severe deficits (ie, for patients with UPSIT scores > 25) than for those with severe microsmia or anosmia. This likely reflects, in cases having sensorineural damage, less extensive necrosis within the basal cell layer of the epithelia and possibly less fibrosis around the foramina of the cribriform plate that can prevent regeneration or appropriate migration of receptor axons to the bulb. It is important for many patients to know the true degree of olfactory loss, given that prognosis relates to the magnitude of the loss. Quantitative testing places the patient's problem into an overall perspective; thus, it can be very therapeutic for an older person to learn that although his or her smell function is not what it used to be, it still falls above the average of his or her peer group.

Treatments are available for some olfactory disorders owing to blockage of airflow to the olfactory neuroepithelium (ie, conductive disorders). Mechanical obstruction secondary to polyposis,

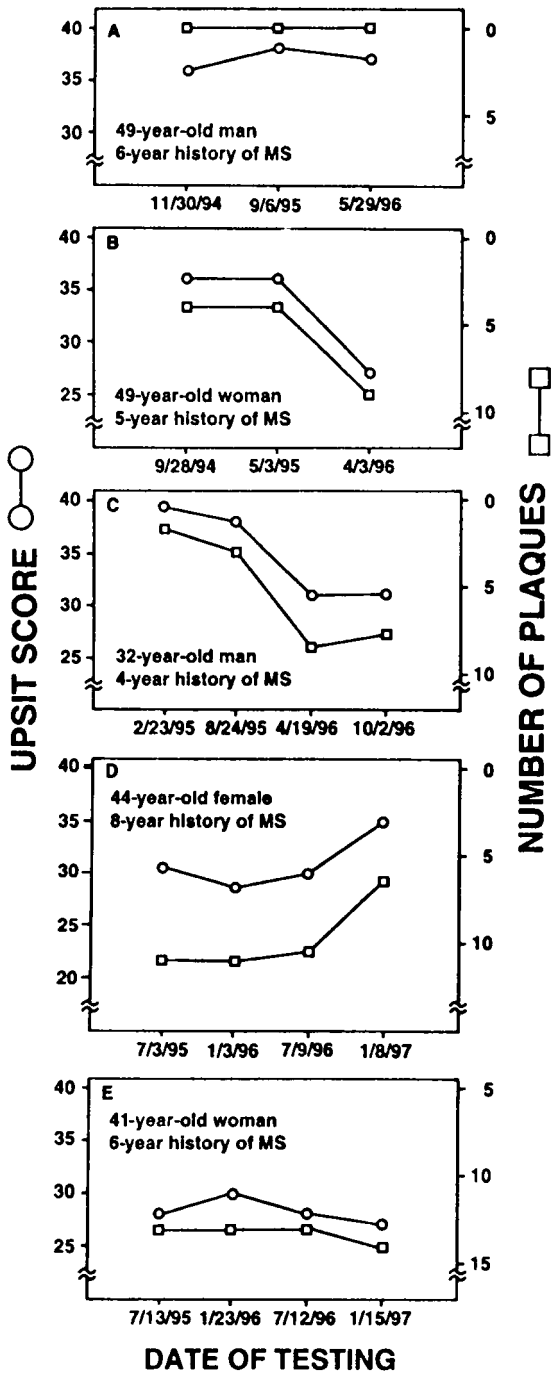


FIGURE 27–9. Longitudinal changes in University of Pennsylvania Smell Identification Test (UPSIT) scores and plaque numbers within the inferior frontal and inferior temporal lobe regions of five patients with multiple sclerosis (MS) (*A* to *E*). Note that the plaque number is inversely plotted. Also note the close association between the measures across time in all cases. Reproduced with permission from Doty RL et al.[67] Copyright © 1999, American Academy of Neurology.

intranasal tumors, and distorted intranasal architecture can often be corrected surgically, and inflammatory causes of nasal and sinus disease can often be addressed medically. However, quantitative olfactory testing is advised before and after the treatment to ensure an accurate assessment of the intervention and to assess relapse. Examples of treatments that have restored olfactory function in some patients include allergy management, topical and systemic corticosteroid therapies, antibiotic therapy, and various invasive interventions, including functional endoscopic sinus surgery. A brief course of systemic corticosteroid therapy can be useful in distinguishing between conductive and sensorineural olfactory loss as patients with the former will often respond positively to the treatment. Longer-term systemic corticosteroid therapy, however, is not advised. Topical nasal corticosteroids may be ineffectual in altering smell dysfunction because the corticosteroid fails to penetrate the higher recesses of the nose. Administering nasal drops or sprays in the head-down Moffett's position can sometimes increase efficacy.

Sensorineural causes of olfactory dysfunction are more difficult to manage. When spontaneous recovery occurs in head trauma patients whose loss reflects damage to the olfactory receptor cell axons, it typically does so within 3 or 4 months of the injury. Patients who quit smoking tobacco typically have dose-related improvement in olfactory function and flavor sensation over time,[69] although tobacco smoking by itself rarely causes complete loss of the sense of smell. Some *central olfactory neural impairments* (eg, tumors within the mesial temporal lobe or tumors that impinge on the olfactory bulbs or tracts) can, in some cases, be resected in a manner that allows for at least some restoration of olfactory function. Chitanondh reported successful treatment of seven patients with seizure disorder, olfactory hallucinations, and psychiatric problems by stereotactic amygdalotomy, concluding that "stereotactic amygdalotomy has a dramatic effect on olfactory seizures, auras, and hallucinations. It is a safe surgical procedure and can be done without neurological deficit."[70]

Some cases of extremely debilitating chronic dysosmia (usually of a number of years duration and often unilateral), in which weight loss is marked or daily functioning is markedly impaired, are amenable to surgical intervention (eg, ablation of regions of the olfactory epithelium or olfactory bulb

removal). Of the surgical approaches, intranasal ablation or stripping of tissue from the olfactory epithelium on the affected side is more conservative and less invasive than removal of the olfactory bulb and/or tract via a craniotomy.[71] Should the dysosmia reappear after intranasal intervention, additional or repeat ablations can be performed. In the majority of cases, demonstrable smell loss does not accompany the dysosmic condition, reiterating the requirement of at least some degree of olfactory system function for the dysosmic expression.

Discontinuance, dose changes, or substitution of other modes of therapy can be effective for some medications that induce distortions of olfaction. Although there are advocates for zinc and vitamin therapies, there is no compelling evidence that these therapies work, except in rare cases in which frank zinc or vitamin deficiencies are present.

THE GUSTATORY SYSTEM

Anatomy and Physiology

Such sensations as sweet, sour, bitter and salty, as well as possibly "metallic" (iron salts), "umami" (monosodium glutamate, disodium guanylate, disodium inosinate), and "chalky" (calcium salts), are mediated via the taste buds of the gustatory system. Unlike olfaction, taste sensations are carried by several cranial nerves (ie, CN VII, IX, X), as discussed in detail below. Because of this fact, complete loss of taste function is rare from peripheral insults or trauma (since all nerves would have to be involved) and is more likely owing to systemic or central causes. Intraoral CN V free nerve endings are also stimulated by some foods and beverages (eg, carbonated or spicy foods), contributing to the overall gestalt of flavor. Hence, a piece of milk chocolate in the mouth is not only sweet but has texture and temperature. The sensation of "chocolate," however, is dependent on retronasal stimulation of the olfactory receptors. Unfortunately, as noted earlier in this chapter, many patients and their physicians fail to distinguish between "taste" sensations mediated by the taste buds from CN I–mediated "taste" sensations (eg, strawberry, chocolate, meat sauce, etc).

Taste Buds The goblet-shaped taste buds are distributed over the dorsal surface of the tongue, margin of the tongue, base of the tongue, soft palate, pharynx, larynx, epiglottis, uvula, and first third of the esophagus.[72] The majority of taste buds are found on the lingual surface within the protruding papillae. Of the four types of papillae—fungiform, foliate, circumvallate, and filiform—only the first three harbor taste buds. Taste buds are continually bathed in secretions from the salivary glands and nearby lingual glands. Ebner's glands discharge into the troughs surrounding the vallate papillae, and lingual glands empty into the long fissures between the folds of the foliate papillae on the posterior aspect of the margin of the tongue. The saliva contains not only proteins such as amylase (which initiates starch breakdown of foodstuffs) but also growth factors important for wound healing and the maintenance of taste buds. Indeed, removal of the submandibular and sublingual salivary glands in rats leads to taste bud loss that can be prevented by supplying epidermal growth factor in the drinking water.[73]

The human tongue contains, on average, around 4,600 taste buds.[72] Each bud is comprised of 50 to 150 thin epithelial cells arranged much like the segments of an orange or grapefruit (Figure 27–10). The opening to the bud is termed the taste pore, and the excavation below the pore is termed the taste pit. Several types of cells are found within the taste bud, including cells that project microvilli into the taste pit and basal cells from which other cell types arise. On the basis of granule density in their apical regions, *light cells, dark cells*, and *intermediate cells* can be identified within each bud.[74] Like the cells of the olfactory neuroepithelium, taste bud cells have the propensity to replace themselves periodically, with the time course for at least some of this "turnover" being between 2 and 3 weeks.[75]

Gustatory Neural Transduction A tastant must be in solution to enter the taste pore. Hence, placing sugar or salt crystals on a dry lingual surface does not immediately lead to sweet or salty sensations, a common mistake made by some practitioners in attempting to assess taste function. After entering the taste pore, the tastant initiates the transduction process via one of two mechanisms: (1) activating receptors coupled to G proteins that, in turn, activate second-messenger systems (a process that probably occurs with sweet and bitter-tasting stimuli), and (2) directly gating apical ion channels on the microvillae within the taste bud pit (a process that probably occurs with sour and salty tasting stimuli). Gustducin, a specialized G protein specialized for taste

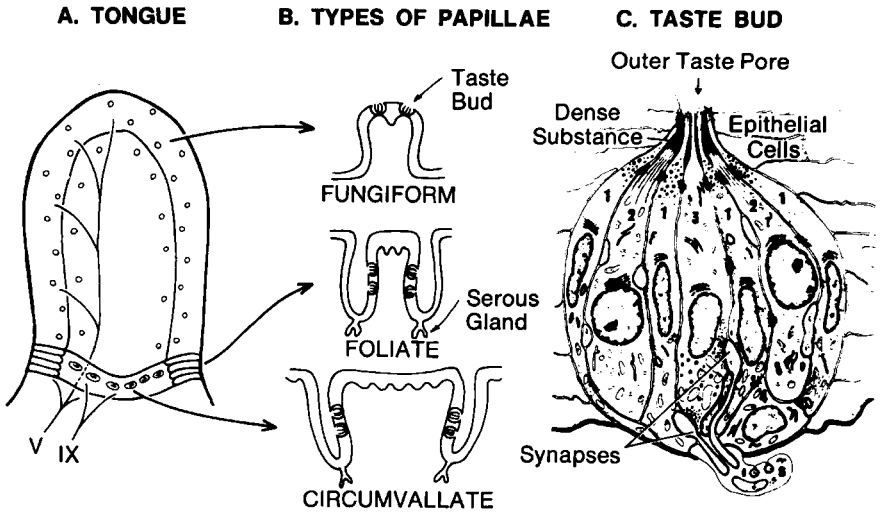


FIGURE 27–10. Lingual cranial nerve innervation (*A*), types of papillae (*B*) that harbor taste buds (*C*). In *C*, 1 and 2 are presumably supporting cells, which secrete materials into the lumen of the taste buds; 3 is a sensory receptor cell, and 4 is a basal cell from which the other types of cells arise. Adapted from Shepherd GM.[76] Copyright © 1983, Oxford University Press.

reception, has been identified in taste receptor cells of fungiform, foliate, and circumvallate papillae.[77]

Taste threshold sensitivity is directly related to the number of taste papillae and hence taste buds that are actively stimulated.[78] Furthermore, taste thresholds are inversely correlated with stimulus duration.[79] Interestingly, there are agents, when swished in the mouth, that selectively alter the quality of taste perceptions. For example, miraculin, a glycoprotein from the berry of the African shrub *Syncepalum dilcificum*, temporarily changes most sour-like sensations to sweet sensations. Gymnemic acid, an extract from the leaves of the Indian plant *Gymnema sylvestre*, can mitigate the perception of sweet sensation (and the corresponding electrophysiologic activity) without significantly changing the perception of the other taste qualities.[80]

Taste Afferent Nerves Different taste buds are innervated by different cranial nerves, depending on the region of the oral cavity in which they are located. Unlike CN I, such nerves are mixed motor and sensory nerves that transmit multiple forms of information. An understanding of this fact can be important when clinical syndromes that involve taste dysfunction are considered. The nerve fibers from each of the taste nerves enter the brainstem and synapse within the nucleus of the tractus solitarius (NTS), which extends from the rosterolateral medulla caudally along the ventral border of the vestibular nuclei (Figure 27–11).[81]

Taste sensations from the taste buds of the fungiform papillae on the anterior two-thirds of the tongue and the taste buds of the soft palate are conveyed by two respective branches of the facial nerve (CN VII): the chorda tympani nerve and the greater petrosal nerve via the lesser palatine nerve. Cranial nerve VII also supplies the salivary and lacrimal glands, the mucous membranes of the oral and nasal cavities, the muscles of facial expression, and the stapedius muscle. Cranial nerve VII taste fibers, whose cell bodies are located within the geniculate ganglion, share a common path with the lingual nerve (CN V_3) proximal to the tongue (see Figure 27–11).

All circumvallate and most, if not all, foliate taste buds within the posterior one-third of the tongue are innervated by the lingual-tonsillar branch of the glossopharyngeal nerve (CN IX). The taste buds in the region of the nasopharynx are supplied by the pharyngeal branch of this nerve. The nerve cell bodies of these gustatory afferent fibers are found within the petrous ganglion immediately outside the jugular foramen, where the fibers eventually pass to enter the cranium. Cranial nerve IX also innervates the stylopharyngeus muscle, the parotid gland, the baroreceptors of the carotid sinus, and the pharyngeal mucous membrane.

The taste buds on the epiglottis, aryepiglottal folds, and esophagus are innervated by the vagus nerve (CN X) via the internal portion of its superior laryngeal branch.[82] Nontaste sensory functions are mediated by this nerve from the external ear, external auditory canal, external surface of the tympanic membrane, and vocal cords. Cranial nerve X also mediates (1) visceral sensation from the larynx and the gut; (2) motor function to the smooth muscle of the pharynx, larynx, and viscera; and (3) motor

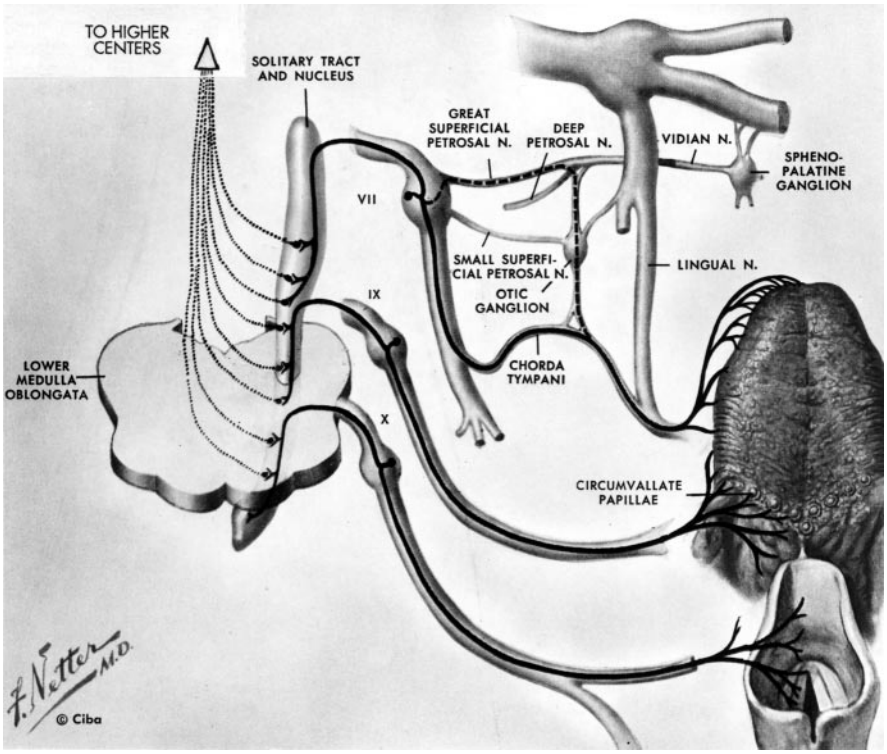


FIGURE 27–11. Distribution of cranial nerves to gustatory regions. Cranial nerve VII fibers from the geniculate ganglion innervate taste buds on the anterior portion of the tongue and on the soft palate. Cranial nerve IX fibers from cell bodies within the petrous ganglion innervate taste buds on foliate and vallate papillae of the tongue, as well as pharyngeal taste buds. Cranial nerve X fibers from cell bodies in the nodose ganglion innervate taste buds on the epiglottis, larynx, and esophagus. Adapted from Netter FH.[86] Copyright © Ciba Pharmaceutical Corporation, New York, 1964.

function to all striated muscles of the pharynx, larynx, and palate except the stylopharyngeus (CN IX) and tensor veli palatini (CN V_3). Afferent fibers from the superior laryngeal nerve project via their cell bodies in the inferior nodose ganglion.

Central Gustatory Regions: Functional Anatomy

The first central relay station of the taste system is the NTS.[82] The afferents from CN VII, IX, and X synapse, respectively, within the NTS in descending (and overlapping) order. Cells from the NTS also make reflexive connections, via the reticular formation, with cranial motor nuclei that control such taste-related behaviors as chewing, licking, salivation, swallowing, and preabsorptive insulin release, as well as the muscles of facial expression.[83] The major gustatory projections from the NST, however, are the tertiary ones that ultimately lead to activation of cortical gustatory structures. These occur via the thalamic taste nucleus (TTN) in primates, that is, the parvicellular division of the ventroposteromedial thalamic nucleus.[84] From the TTN, fibers project to the primary taste cortex deep in the parietal operculum and adjacent parainsular cortex.[85] Positron emission tomography and fMRI studies have found taste stimulus–induced activation largely within the insula and perisylvian regions, including the frontal operculum, superior temporal gyrus (opercular part), and inferior sectors of the pre- and postcentral gyrus.[87,88] A secondary cortical taste region is present within the caudomedial/caudolateral orbitofrontal cortex, several millimeters anterior to the primary taste cortex.[81]

The specific role of the gustatory cortical regions in the processing of taste information is not clear. These regions contain multimodal neurons responsive to touch and temperature, as well as taste.[89] Some structures within the right hemisphere seem to play a more important role in taste perception than their left-hemisphere counterparts. For example, citric acid presented to the whole mouth increases regional cerebral blood flow more in the right than in the left anteromedial temporal lobe and more in the right than in the left caudomedial orbitofrontal cortex, although bilateral activation is observed in the caudolateral orbitofrontal cortex.[90] Also, patients who have had right anterior temporal lobe resection for intractable seizure activity have higher citric acid recognition thresholds than both controls and left anterior temporal lobe resection patients. There is some evidence, however, that handedness may need to be controlled in such stud-

ies. Thus, whereas fMRI data suggest that the superior part of the insula is activated by tastants similarly in left- and right-handed subjects, this is not the case for the inferior parts of the insulae, where the left is relatively more activated in right-handed subjects and the right is relatively more activated in left-handed subjects.[87]

Classification of Gustatory Disorders

Gustatory disorders can be classified in a manner similar to olfactory disorders: *ageusia*: inability to detect qualitative gustatory sensations from all (total ageusia) or some (partial ageusia) tastants; *hypogeusia*: decreased sensitivity to tastants; *dysgeusia*: distortion in the perception of a normal taste (ie, an unpleasant taste induced by a stimulus that is normally associated with pleasant sensations) or the presence of a taste in the absence of a stimulus (sometimes termed phantogeusia); *gustatory agnosia*: inability to recognize a taste sensation, even though gustatory processing, language, and general intellectual functions are essentially intact. Some patients complain of oral sensations of burning or numbness, which may or may not have their genesis in gustatory afferents, for example, *burning mouth syndrome* in which the sensation of "burning" occurs within the mouth without obvious physical cause.

Total ageusia is rare and, when present, is usually produced by central (eg, ischemic) events since regeneration of taste buds can occur, and peripheral damage would have to involve multiple pathways to induce taste loss. Thus, whereas 433 of 585 patients (74%) studied at the University of Pennsylvania Smell and Taste Center who exhibited olfactory loss complained of both smell and taste disturbance, less than 4% had verifiable whole-mouth gustatory dysfunction, and even that was limited.[27] Regional deficits are much more common. For example, in one study, sensitivity to three concentrations of NaCl was measured on the tongue tip and 3 cm posterior to the tongue tip in 12 young (20 to 29 years of age) and 12 elderly (70 to 79 years of age) subjects. On average, the young subjects were more sensitive to NaCl on the tongue tip than on the more posterior stimulation site and exhibited, at both tongue loci, an increase in detection performance as the stimulus concentration increased. The elderly subjects, who would be expected to exhibit, at worse, moderate deficits on whole-mouth testing, performed at chance levels.[91]

Most patients are unaware of regional taste deficits. In fact, most patients can sustain loss of taste sensation on one half of the anterior part of the tongue following unilateral sectioning of the chorda tympani in middle ear surgery without noticing the problem. Such lack of awareness stems, in part, from the redundancy of the multiple taste nerves, as well as compensatory mechanisms.

Clinical Evaluation of Gustatory Function

History A history similar to that described earlier for patients complaining of olfactory disturbance should be obtained from patients complaining of gustatory disturbance. Specific consideration as to the type of stimuli that can or cannot be detected by the patient is essential to distinguish between retronasal CN I flavor loss and true taste bud–mediated gustatory loss. One should specifically inquire as to whether the patient can detect the saltiness of potato chips, pretzels, or salted nuts; the sourness of vinegar, pickles, or lemons; the sweetness of sugar, soda, cookies, or ice cream; and the bitterness of coffee, beer, or tonic water. If the patient indicates that there is a problem in such detection, the possibility of a true taste bud–mediated dysfunction exists.

Previous or current problems with salivation, chewing, swallowing, oral pain or burning, dryness of the mouth, periodontal disease, speech articulation, bruxism, or foul breath odor should be ascertained. Inquiry as to diet, oral habits, stomach problems, and possible problems with acid reflux is relevant, given that acid reflux into the oral cavity can irritate or damage taste buds. Recent dental work or exposure to radiation should be noted. Documentation of hearing or balance problems should be made since past or current ear infections or surgery can result in altered chorda tympani function and produce taste loss or distortions. A careful assessment of medication use is critical. As described in detail later in this chapter, many drugs, including lipid-reducing agents, antibiotics, and antihypertensives, can produce significant distortions or other alterations in taste function.

Physical Examination A thorough head and neck examination is essential for ascertaining the potential cause of a gustatory disorder.[20] Nongustatory deficits in CN VII, IX, and X can shed light on

whether gustatory dysfunction may be present (eg, abnormal facial motion, swallowing, salivation, gag reflex, voice production). Changes in epithelial color or signs of scarring, inflammation, or atrophy of lingual papillae should be noted. Neoplastic lesions deep in the tongue's musculature should be ruled out by palpation. Specific attention to the condition of the teeth and gums is important since exudates produced by gingivitis may produce or contribute to dysgeusic symptoms. An inspection of the nature and integrity of the fillings, bridges, and other dental work should also be made (eg, dissimilar metals can induce small electrical currents that, in turn, produce abnormal oral sensations). In cases for which an explanation of the taste problem is not clear, neuroimaging to rule out CNS tumors or lesions should be performed.

In some instances, it may be useful to evaluate biopsies of circumvallate or fungiform papillae to determine the presence of pathology. The tongue can also be stained with a dark food dye and can be photographed under high illumination and low-power magnification to allow for counting or better visualizing selected classes of papillae. In general, there is a high correlation between the number of fungiform papillae and the number of taste buds.[78]

Quantitative Gustatory Testing Quantitative taste testing is rarely performed in the clinic, largely because of issues of practicality (eg, time and expense of presenting and preparing limited shelf-life taste stimuli). Although a number of whole-mouth taste tests have been described in the literature,[92] regional taste testing is needed to establish the function of each of the nerves innervating multiple taste bud fields. As noted above, whole-mouth tests are insensitive to even complete dysfunction of one or several of the nerves that innervate the tongue.

Regional taste testing can be made using either chemical or electrical stimuli. The former requires applying known concentrations of liquid stimuli (eg, sucrose, citric acid, caffeine, and sodium chloride as prototypical representatives of sweet, sour, bitter, and salty taste qualities) to the tongue or oral cavity (in some cases in comparison with blank trials), rinsing the stimuli off between trials, and expectorating after each presentation. Such testing can be quite time consuming. For example, if responsiveness to each of the four basic taste qualities is to be made on left and right anterior (CN VII) and posterior (CN IX) tongue regions, 16 trials (4 tastants × 4 tongue regions) would be needed to present a single stimulus for each quality. Since multiple stimuli are required to produce reliable responses, the number of trials increases considerably.

In a regional chemical test used at our center, a single suprathreshold concentration of each of four stimuli (sucrose, citric acid, sodium chloride, and caffeine) is employed. The stimuli have been equated for perceived intensity (using a magnitude matching procedure) and are presented at the same volume (15 μL from an Eppendorff pipette; Brinkman Instruments, Westbury, New York) and kinematic viscosity (1.53 mm^2/s^2 using tasteless cellulose gum). Equating viscosity eliminates context effects and tactile cues and increases control of solution spread when placed on the tongue in microliter quantities. The stimuli are presented six times to each tongue region, resulting in a test consisting of 96 stimulus trials and 96 rinses. Testing time is between 1 and 1½ hours. The task of the subject is to report the presence of a sweet, sour, bitter, or salty sensation using a forced-choice procedure with these four alternatives available and to rate the perceived intensity of the stimulus on a standardized rating scale with logarithmic visual properties.

A more practical approach to taste testing is to employ an electrogustometer, a device that presents brief μA currents to small regions of the tongue for known durations (Figure 27–12).[93] No stimulus preparation or rinsing is required. Electrogustometric thresholds can be obtained easily, although their relationship to chemical thresholds is still not clear, and extreme care must be taken to apply the stimulator to the same region of the tongue on each trial as considerable regional variation in sensitivity is present. At low current ranges, electrogustometry has been demonstrated to activate only taste afferents, not trigeminal (CN V) afferents. Unfortunately, sound normative data based on forced-choice testing paradigms are still generally lacking for such devices.

Imaging Studies Imaging studies of the gustatory pathways can be useful in explaining the gustatory symptoms of some patients. In addition to detecting large central lesions and tumors, modern MRI techniques can detect discrete lesions (eg, infarcts) within brain structures that correlate both with

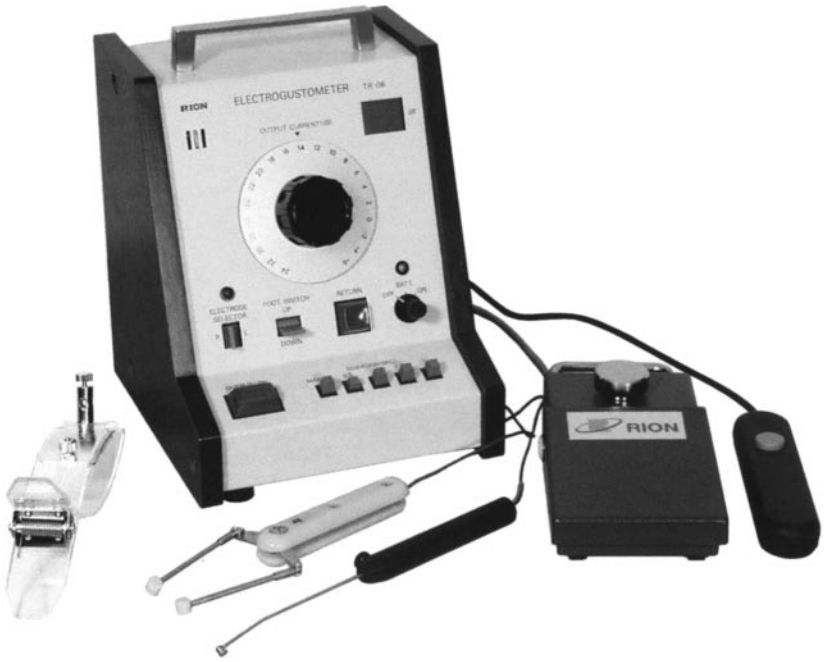


FIGURE 27–12. The Rion TR-06 electrogustometer, a practical device for quantitatively assessing regional taste function. This device stimulates the taste system using microvolt-level pulses of constant current of known duration. Photograph courtesy of Sensonics, Inc., Haddon Heights, New Jersey. Copyright © 2000, Sensonics, Inc.

patient complaints and with the results of sensory testing. As noted in detail in the next section, MRI-determined infarcts in the pons have been repeatedly associated with ageusia and dysgeusia.[94,95]

Causes of Gustatory Dysfunction

Although central or systemic factors are the most likely causes of ageusia, local factors can significantly alter taste perception. Proximal mechanisms include (1) the release of foul-tasting materials from oral medical conditions (eg, gingivitis, sialadenitis); (2) problems in movement of tastants to the taste buds (eg, damage to taste pores from a burn); (3) damage to the taste buds proper, as from caustic or allergic reactions to toxins or oral products; (4) damage to taste nerves (eg, postviral Bell's palsy, dental or surgical procedures); and (5) CNS damage (eg, from tumors, epilepsy, infarcts). One cause of dysgeusia can be the use of different metals in the mouth that set up subtle electrical currents within the oral cavity. The more common causes of taste disorders are discussed below.

Aging Although whole-mouth taste acuity declines with age, the perceptual decrease is not as marked as that seen for olfaction.[96] Compared with younger persons, the elderly tend to perceive tastes as being less intense. In conjunction with the loss seen with the sense of smell, decrements in taste function can be harmful to some of the elderly, leading to anorexia, weight loss, malnutrition, impaired immunity, and worsening of a medical illness.[50,97] As noted earlier, whole-mouth testing reveals moderate degrees of age-related taste loss, whereas regional taste testing reveals marked age-related decrements.[91]

Bell's Palsy Bell's palsy is the most frequent cause of facial nerve damage. Usually of viral origin (eg, herpes simplex), this disease affects both sexes of all ages. This disorder typically begins with pain in or behind the ipsilateral ear, followed by symptoms of unilateral facial weakness over the course of a few days (maximum paralysis within 48 to 72 hours). Hyperacusis occurs in many cases owing to a weakening of the stapedius muscle. Although commonly accompanied by ipsilateral taste loss over the anterior two-thirds of the tongue, patients are often unaware of the taste problem without formal testing. There is a suggestion that return of taste function by 2 weeks after onset is a positive prognostic indicator for complete and relatively rapid recovery from facial paresis, whereas longer-lasting taste impairment is associated with poor prognosis. The taste-salivary reflex arc, which extends from the taste

nerves via the NST to the parasympathetic fibers innervating the salivary glands, can also be compromised in this disorder.

Herpes zoster oticus, termed Ramsay Hunt syndrome, results from active infection of the geniculate ganglion and involves pain and vesicles in the external auditory canal or soft palate. Bartoshuk and Miller described a patient with Ramsay Hunt syndrome that involved both CN V and CN IX on the left side.[98] Intensity ratings were obtained for each of the four taste qualities at regular intervals across a nearly 600-day period for the front, back, and palate regions. About 3 months after the herpes attack, the entire left side of the patient was devoid of taste function. After 400 days, all taste qualities were perceived on the left rear and palate regions, but only sweetness could be perceived on the left front. Full taste recovery of the anterior part of the tongue was not present even by the end of the nearly 600-day-long testing period.

Burning Mouth Syndrome In some cases, burning mouth syndrome, also called glossodynia or glossalgia, is associated with salty or bitter dysgeusias. This poorly understood syndrome is characterized by idiopathic intense "burning" pain within the mouth without obvious physical cause. To what extent this disorder is mediated via gustatory or CN V afferents is not known. The burning sensation typically begins by late morning and continues throughout the day.[99] Suggested causes include (1) diabetes mellitus (possibly predisposing to oral candidiasis); (2) nutritional or hormone deficiencies (eg, iron, folic acid, B vitamins, estrogen, zinc); (3) denture allergy (including reactions to amalgam fillings); (4) mechanical irritation from dentures or oral devices; (5) parafunctional habits of the mouth (eg, tongue thrusting, teeth grinding, jaw clenching); (6) tongue ischemia as a result of temporal arteritis; (7) oral candidiasis; (8) periodontal disease; (9) reflux esophagitis; and (10) geographic tongue.[100] Anxiety and depression are common in the population with burning mouth syndrome.

Central Lesions and Tumors Taste dysfunction has been attributed to a variety of tumors and other lesions, including acoustic neuromas, tumors of the hypophysis with marked extrasellar growth, facial nerve schwannomas extending into the middle cranial fossa, and cerebellopontine angle lesions. Contralateral dysgeusia has been observed in patients with lateralized infarcts of the thalamus and in patients with infarcts in the corona radiata (reflecting the crossed taste pathways at this level of the nervous system).[101] Unilateral lesions above the brainstem do not usually cause complete loss of function because of the multiple areas involved in processing taste information.

Gustatory problems can arise from damage to the brainstem structures related to taste, often in conjunction with impairment of other cranial nerves or long tracts. Regions susceptible to damage include the NTS and the pontine tegmentum, which involves both gustatory lemnisci. Hemiageusia from an ipsilateral MS plaque at the midpontine tegmentum has been reported.[94] Similarly, ageusia to all taste qualities on the right side of the tongue was noted in a patient who had a small hemorrhage in the right tegmentum of the middle pons.[102] Recently, three cases of ipsilateral hemiageusia owing to focal ischemic lesions in the brainstem have been described.[103]

An example of unilateral hypogeusia resulting from a brainstem infarct is shown in Figure 27–13. In this case, ischemic activity is evident in a region of the upper medulla near the right NTS (note that, in this image, the patient's right side is on the left of the picture, and the left side is on the right of the picture). This individual complained of dysgeusia on the right side of the tongue and evidenced on testing a decrement in the ability to discern sweet, sour, bitter, and salty sensations that was greater on the right than on the left. Both CN VII and CN IX seemed to be involved.

Head and Neck Trauma Trauma-related taste loss is much less common than trauma-related smell loss, with fewer than 1% of persons with major head injury exhibiting ageusia to sweet, sour, salty, or bitter taste qualities.[27,104] Nonetheless, taste loss, as well as dysgeusia, can occur in some types of head trauma. For example, basilar temporal bone fractures and other injuries that impinge on the middle ear have the potential for impairing chorda tympani nerve–mediated taste function unilaterally, as well as for altering salivary secretion. Traumatic injury to the lingual nerve in and around the mouth and tongue can also occur in some trauma cases. This nerve, a branch of the mandibular division of CN V_3, is the most proximal pathway to the tongue for

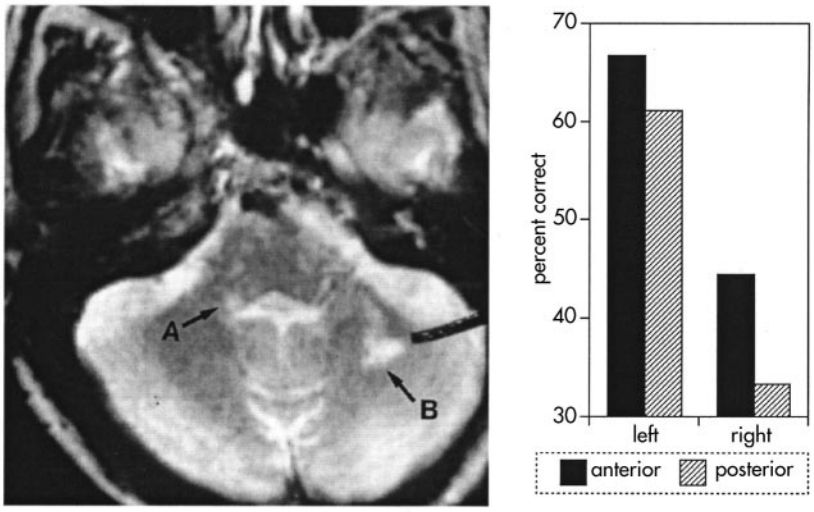


FIGURE 27–13. Left, Axial T_2-weighted (2500/90) magnetic resonance scan through the upper brainstem reveals a hyperintensive infarct (4 × 3 mm) in the right side of the medulla (*arrow A*). Note also the large infarct (15 × 8 mm) inside the white matter of the left side of the cerebellum (*arrow B*). *Right*, Taste identification scores showing relative decrement on the right side of the tongue. From a 65-year-old-woman with a history of ministrokes who developed a persistent salty/metallic dysgeusia and soreness on the right side of the tongue following a severe 2-day bout of emesis accompanied by marked dehydration and increased blood pressure but unaccompanied by fever. Reproduced with permission from Doty RL. Copyright © 2001, Richard L. Doty.

general somatic sensation (touch) and special visceral sensation of taste (owing to concurrent chorda tympani fibers).

Trauma-induced damage to the auriculotemporal nerve can produce a rare clinical syndrome of post-traumatic gustatory neuralgia, characterized by paroxysmal, lancinating facial pain in the cutaneous distribution of the auriculotemporal nerve following gustatory stimulation.[105,106] This condition, which can also be caused by viruses and by iatrogenic trauma (eg, that induced by parotid gland surgery, temporomandibular joint surgery, carotid endarterectomy, orthognathic surgery, and oncologic surgery), produces episodic, transient, electric shock-like pain in the preauricular region of the affected side.[105] Taste-mediated post-traumatic sweating and flushing in this same region are known as Frey's syndrome. Although usually provoked by taste stimuli, there is suggestion that such responses can be provoked, in some cases, by the smell of food or even emotional excitement.[105]

Iatrogenic Surgical Damage Numerous surgical interventions can induce taste dysfunction. The glossopharyngeal nerve is susceptible to damage during tonsillectomy, bronchoscopy, or laryngoscopy,[107–109] reflecting the close proximity of the lingual branch of this nerve to the muscle layer of the palatine tonsillar fossa.[110] Surgical treatment for snoring (eg, uvulopalatoplasty),[111] as well as surgery-related procedures such as endotracheal intubation (causing injury to the lingual nerve)[112] and the employment of a laryngeal mask,[113] has been associated with taste loss or alteration.

The chorda tympani nerve is at particular risk from surgical procedures that involve the middle ear, given its course between the malleus and the incus. The nerve is often stretched or sectioned during tympanoplasty, mastoidectomy, and stapedectomy, in some cases producing long-lasting symptoms. Bull, for example, found that 78% of patients with bilateral section and 32% of patients with unilateral section of the chorda tympani had persistent adverse gustatory symptoms.[114]

A recent study evaluated gustatory function in 17 patients before third molar surgery and at 1 and 7 months after surgery.[115] On average, an ~15% reduction of perceived intensity was observed 1 month after surgery for NaCl, citric acid, and quinine hydrochloride. The taste quality of NaCl was identified correctly less often after than before third molar extraction. Citric acid intensity perception

had not recovered at 6 months after the surgery. These data suggest that not only do gustatory deficits commonly occur after third molar extraction, they also can persist for at least 6 months after surgery and seem to be associated with the depth of impaction.

Medications Medications produce taste disturbances more frequently than olfactory disturbances. Such side effects can be very debilitating and, in rare instances, have contributed to extreme weight loss and even suicide. Among offending agents are antiproliferative drugs, lipid-reducing drugs, antihypertensive agents, diuretics, antifungal agents, antirheumatic drugs, antibiotics, and drugs with sulfhydryl groups, such as penicillamine and captopril.[36,116] Onset of taste dysfunction following the use of some agents can take weeks or even months. For example, in a well-controlled study of 87 patients experiencing taste loss as a result of the antifungal agent terbinafine, the average time period from the first intake of the drug and the experience of taste loss was 35 days. In most cases, recovery after drug cessation took several months.[117]

Taste function is reportedly altered by repetitive use of some oral topical agents, including hydrogen peroxide or corticosteroids. Some medications (eg, anticholinergics, antidepressants, antihistamines) reduce glandular secretion, causing hyperviscous saliva, a condition that, in time, can result in lessened taste acuity. It is noteworthy that approximately one-fourth of all cardiac medicines (including antilipemic agents, adrenergic blockers, angiotensin-converting enzyme inhibitors, angiotensin II antagonists, calcium channel blockers, vasodilators, anticoagulants, antiarrhymics, and various diuretics) are listed in the *Physicians' Desk Reference* as having potential side effects of "altered taste," "bad taste," "bitter taste," or "metallic taste."

Radiation Therapy Radiotherapy for head and neck cancer can induce taste loss and dysgeusia, salivary dysfunction, and conditioned taste aversions. Such problems can significantly alter quality of life and, on occasion, appetite to the degree that nutrition is compromised. Symptoms usually begin early in the course of treatment, and post-treatment recovery can take months and, in rare instances, years.[118] The cause of such disturbances can be quite varied, including direct radiation-induced damage to taste cells and buds, salivary glands, and taste nerve fibers (which normally provide trophic factors that maintain the integrity of the taste bud). The xerostomia secondary to salivary gland damage can influence food transport, protection from bacterial invasion, and salivary proteins potentially involved in taste transduction[119] and can aid in the promotion of opportunistic oral infections (eg, oral candidiasis).[120] Although typically idiosyncratic and transient, some taste aversions can be long lasting and can produce generalized anorexia and cachexia. A means for mitigating such aversions is to have the patient consume a novel food immediately before the first course of chemo- or radiotherapy. This simple maneuver somehow focuses the aversion primarily on the novel food and interferes with the formation of conditioned aversions to preferred dietary items.[121]

Other Causes Among other reported causes of taste dysfunction are hypothyroidism, renal disease, liver disease, myasthenia gravis, Guillain-Barré syndrome, numerous neoplasms, and familial dysautonomia (a genetic disorder with lack of taste buds and papillae). Idiopathic dysgeusia has been associated with blood transfusions.[122] Complaints of taste loss or distortion are reported in many carcinomas and mass lesions. For instance, whereas squamous cell carcinoma of the mucous membranes of the upper aerodigestive tract can interfere with taste by direct destruction of receptors, mass lesions along the course of CN VII, IX, and X may cause impairment through neural compression. Carcinoma-related malnutrition may also lead to ageusia.

A persistent and unpleasant sweet dysgeusia was recently described in three patients with small cell carcinoma of the lung.[123] The dysgeusia was the presenting symptom in all three cases, and hyponatremia secondary to the syndrome of inappropriate secretion of antidiuretic hormone was present in each case. Resolution of the dysgeusia paralleled an increase in serum sodium concentration after water restriction alone. The close association between the dysgeusia and the low serum sodium concentration suggests that hyponatremia is the causative factor, rather than the carcinoma, antiduretic hormone, medications, or chemotherapy.

Gustatory symptoms have been reported in association with epileptic seizures. Examples of taste sensations that have been reported in such cases include "peculiar," "rotten," "sweet," "like a cigarette," "like rotten apples," and "like vomitus."[39] Some of

these "tastes," however, likely represent smell sensations that are miscategorized as tastes by both the patients and their physicians.

Selective taste nerve damage or alterations may produce some forms of *hyper*geusia and dysgeusia. For example, one study found that anesthetizing a single chorda tympani nerve reportedly increases the perceived intensity of bitter substances, such as quinine, applied to taste fields innervated by the contralateral glossopharyngeal nerve.[124] In contrast, perceived intensity of NaCl applied to an area innervated by the ipsilateral glossopharyngeal nerve appeared decreased. When both chorda tympani nerves were anesthetized, the taste of quinine is intensified, and the taste of NaCl was diminished in areas innervated by the glossopharyngeal nerve on both sides of the tongue. In about 40% of the subjects, a phantom taste, usually localized to the posterior part of the tongue contralateral to the anesthesized area, appeared in the absence of stimulation. This phantom taste was eliminated when the region of origin was anesthetized. These authors suggest that such phantoms arise because of release of inhibition normally present between the central projection areas of the different taste nerves.

Treatment of Taste Dysfunction

As with other sensory systems, prognosis in cases of taste dysfunction is likely inversely related to the degree of neural or structural damage, although clinically such damage can rarely be assessed. Fortunately, the taste nerves and buds appear to be relatively resilient, as many cases of taste loss or distortion spontaneously resolve over time.[28] In some dysgeusic cases, antifungal and antibiotic treatments have been reported to be useful, although double-blind studies of the efficacy of such treatments are lacking, and some of these agents themselves can produce taste disturbance. Chlorhexidine employed in a mouthwash has been suggested as having possible efficacy for some salty or bitter dysgeusias, possibly as a result of its strong positive charge.[125] In the case of neural damage from viruses or other agents, presumably the damaged taste afferents regenerate. Thyroxine replacement therapy reportedly brings back taste sensitivity to normal levels in cases of taste loss secondary to hypothyroidism.[126] Taste disorders caused by medications can, in some instances, be reversed by discontinuing the offending drug, by employing alternative medications, or by changing drug dosage. It should be kept in mind, however, that a number of pharmacologic agents appear to induce long-term alterations in taste that may take months to disappear after drug discontinuance.

Since most patients with burning mouth syndrome are postmenopausal, dysgeusias associated with this disorder may, in some cases, respond to estrogen replacement therapy. Topical capsaicin may also be helpful in some cases. Given that burning mouth syndrome is often associated with anxiety and depression, tricyclic antidepressants (eg, amitriptyline, desipramine, nortriptyline) and benzodiazepines appear, in selected instances, to have some therapeutic efficacy.[127]

CONCLUSIONS

Disorders of the chemical senses of taste and smell are relatively common, particularly in the elderly, and can result from a broad array of causes. Such causes range from simple irritation of the receptive elements to serious neurologic disorders, including AD and idiopathic PD. In this chapter, a succinct overview of the anatomy and physiology of the chemical senses, as well as of the primary causes of chemosensory dysfunction, has been provided. Approaches to therapy have been discussed, with an emphasis on the need for quantitative psychophysical evaluation of patients before initiating surgical or medical interventions. Clearly, a number of disorders of the chemical senses can be approached with optimism as long as the physician establishes the exact nature of the problem and is aware of the available avenues of treatment and objective assessments of efficacy.

ACKNOWLEDGMENT

This work was supported in part by Grants PO1 DC 00161, RO1 DC 04278, RO1 DC 02974, and RO1 AG 27496 from the National Institutes of Health, Bethesda, Maryland (R. L. Doty, principal investigator).

REFERENCES

1. Nakashima T, Kimmelman CP, Snow JB. Structure of human fetal and adult olfactory neuroepithelium. Arch Otolaryngol 1984;110:641–6.

2. Huard JM, Youngentob SL, Goldstein BL, et al. Adult olfactory epithelium contains multipotent progenitors that give rise to neurons and non-neural cells. J Comp Neurol 1998;400:469–86.
3. Rowley JC, Moran DT, Jafek BW. Peroxidase backfills suggest the mammalian olfactory epithelium contains a second morphologically distinct class of bipolar sensory neuron: the olfactory microvillar cell. Brain Res 1989;502:387–400.
4. Carr VM, Farbman AI, Colletti LM, Morgan JI. Identification of a new non-neuronal cell type in rat olfactory epithelium. Neuroscience 1991;45:433–49.
5. Smith CG. Age incident of atrophy of olfactory nerves in man. J Comp Neurol 1942;77:589–94.
6. Meisami E, Mikhail L, Baim D, Bhatnagar KP. Human olfactory bulb: aging of glomeruli and mitral cells and a search for the accessory olfactory bulb. Ann N Y Acad Sci 1998;855:708–15.
7. Sobel N, Prabhakaran V, Desmond JE, et al. Sniffing and smelling: separate subsystems in the human olfactory cortex. Nature 1998;392:282–6.
8. Shepherd GM. Synaptic organization of the mammalian olfactory bulb. Physiol Rev 1972;52: 864–917.
9. Buck L, Axel R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell 1991;65:175–87.
10. Ngai J, Chess A, Dowling MM, et al. Coding of olfactory information: topography of odorant receptor expression in the catfish olfactory epithelium. Cell 1993;72:667–80.
11. Jones DT, Reed RR. G_{olf}: an olfactory neuron-specific G protein involved in odorant signal transduction. Science 1989;244:790–5.
12. Leinders-Zufall T, Shepherd GM, Zufall F. Modulation by cyclic GMP of the odour sensitivity of vertebrate olfactory receptor cells. Proc R Soc Lond B Biol Sci 1996;263:803–11.
13. Swift DL, Procter DF. Access of air to the respiratory tract. In: Brau JD, Procter DF, Reid LM, editors. Respiratory defense mechanisms. New York: Marcel Dekker; 1977. p. 63–93.
14. Wekesa KS, Anholt RRH. Differential expression of G proteins in the mouse olfactory system. Brain Res 1999;837:117–26.
15. Hinds JW, Hinds PL, McNelly NA. An autoradiographic study of the mouse olfactory epithelium: evidence for long-lived receptors. Anat Rec 1984; 210:375–83.
16. Mackay-Sim A, Kittel PW. On the life span of olfactory receptor neurons. Eur J Neurosci 1990;3:209–15.
17. Calof AL, Rim PC, Askins KJ, et al. Factors regulating neurogenesis and programmed cell death in mouse olfactory epithelium. Ann N Y Acad Sci 1998;30:226–9.
18. Holcomb JD, Mumm JS, Calof AL. Apoptosis in the neuronal lineage of the mouse olfactory epithelium: regulation in vivo and in vitro. Dev Biol 1995; 172:307–23.
19. Doty RL, Snow JB. Olfaction. In: Goldman JL, editor. The principles and practice of rhinology. New York: Wiley; 1987. p. 761–84.
20. Bromley SM. Smell and taste disorders: a primary care approach. Am Fam Physician 2000;61:427–36.
21. Doty RL. Olfaction. Ann Rev Psychol 2001;52:423–52.
22. Doty RL. The Smell Identification Test™ administration manual. 3rd ed. Haddon Heights (NJ): Sensonics, Inc.; 1995. p. 1–57.
23. Doty RL, Kobal G. Current trends in the measurement of olfactory function. In: Doty RL, editor. Handbook of olfaction and gustation. New York: Marcel Dekker; 1995. p. 191–225.
24. Ottoson D. Analysis of the electrical activity of the olfactory epithelium. Acta Physiol Scand 1956;35: 1–83.
25. Doty RL, Shaman P, Dann M. Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function. Physiol Behav 1984;32:489–502.
26. Lezak MD. Neuropsychological assessment. 3rd ed. New York: Oxford University Press; 1995.
27. Deems DA, Doty RL, Settle RG, et al. Smell and taste disorders, a study of 750 patients from the University of Pennsylvania Smell and Taste Center. Arch Otolaryngol Head Neck Surg 1991;117:519–28.
28. Deems DA, Yen DM, Kreshak A, Doty RL. Spontaneous resolution of dysgeusia. Arch Otolaryngol Head Neck Surg 1996;122:961–3.
29. Clark WEL. Anatomical investigation into the routes by which infections may pass from the nasal cavities into the brain. Rep Pub Health Med Sub No 54 1929;1–27.
30. Armstrong C, Harrison WT. Prevention of intranasally-inoculated poliomyelitis of monkeys by instillation of alum into the nostrils. Public Health Rep 1935;22:725–30.
31. Tisdall FF, Brown A, Defries RD, et al. Nasal spraying as preventive of poliomyelitis. Can J Public Health 1937;28:431–4.
32. Doty RL. Olfactory dysfunction in neurogenerative disorders. In: Getchell TV, Doty RL, Barto-

shuk LM, Snow JB, editors. Smell and taste in health and disease. New York: Raven Press; 1991. p. 735–51.
33. Doty RL, Singh A, Tetrude J, Langston JW. Lack of olfactory dysfunction in MPTP-induced parkinsonism. Ann Neurol 1992;32:97–100.
34. Roberts E. Alzheimer's disease may begin in the nose and may be caused by aluminosilicates. Neurobiol Aging 1986;7:561–7.
35. Ferreyra-Moyano H, Barragan E. The olfactory system and Alzheimer's disease. Int J Neurosci 1989;49:157–97.
36. Doty RL, Bartoshuk LM, Snow JB. Causes of olfactory and gustatory disorders. In: Getchell TV, Doty RL, Bartoshuk LM, Snow JB, editors. Smell and taste in health and disease. New York: Raven Press; 1991. p. 449–61.
37. Doty RL, Risser JM, Brosvic GM. Influence of adrenalectomy on the odor detection performance of rats. Physiol Behav 1991; 49:1273–7.
38. Doty RL, Deems DA, Frye RE, et al. Olfactory sensitivity, nasal resistance, and autonomic function in patients with multiple chemical sensitivities. Arch Otolaryngol Head Neck Surg 1988;114:1422–7.
39. West SE, Doty RL. Influence of epilepsy and temporal lobe resection on olfactory function. Epilepsia 1995;36:531–42.
40. Doty RL, Yousem DM, Pham LT, et al. Olfactory dysfunction in patients with head trauma. Arch Neurol 1997;54:1131–40.
41. Yousem DM, Geckle RJ, Bilker WB, et al. Posttraumatic olfactory dysfunction: MR and clinical evaluation. AJNR Am J Neuroradiol 1996;17:1171–9.
42. Nakashima T, Kimmelman CP, Snow JB. Progressive olfactory degeneration due to ischemia. Surg Forum 1983;34:566–8.
43. Doty RL, Mishra A. Olfaction and its alteration by nasal obstruction, rhinitis, and rhinosinusitis. Laryngoscope 2001;111:409–23.
44. Kern RC. Chronic sinusitis and anosmia: pathologic changes in the olfactory mucosa. Laryngoscope 2000;110:1071–7.
45. Feron F, Perry C, McGrath JJ, Mackay S. New techniques for biopsy and culture of human olfactory epithelial neurons. Arch Otolaryngol Head Neck Surg 1998;124:861–6.
46. Lee SH, Lim HH, Lee HM, et al. Olfactory mucosal findings in patients with persistent anosmia after endoscopic sinus surgery. Ann Otol Rhinol Laryngol 2000;109:720–5.
47. Watnick RL, Trobe JA. Bilateral optic nerve compression as a mechanism for the Foster Kennedy syndrome. Ophthalmology 1989;96:1793–8.
48. Schatz NJ, Smith JL. Non-tumor causes of the Foster Kennedy syndrome. J Neurosurg 1967;27:37–44.
49. Doty RL, Shaman P, Applebaum SL, et al. Smell identification ability: changes with age. Science 1984;226:1441–3.
50. Mattes RD, Cowart BJ. Dietary assessment of patients with chemosensory disorders. J Am Diet Assoc 1994;94:50–6.
51. Schiffman SS. Taste and smell losses in normal aging and disease. JAMA 1997;278:1357–62.
52. Kalmey JK, Thewissen JG, Dluzen DE. Age-related size reduction of foramina in the cribriform plate. Anat Rec 1998;251:326–9.
53. Mesholam RI, Moberg PJ, Mahr RN, Doty RL. Olfaction in neurodegenerative disease: a meta-analysis of olfactory functioning in Alzheimer's and Parkinson's diseases. Arch Neurol 1998;55:84–90.
54. Doty RL, Li C, Mannon LJ, Yousem DM. Olfactory dysfunction in multiple sclerosis: relation to plaque load in inferior frontal and temporal lobes. Ann N Y Acad Sci 1998;855:781–6.
55. Moberg PJ, Doty RL, Turetsky BI, et al. Olfactory identification deficits in schizophrenia: correlation with duration of illness. Am J Psychiatry 1997; 154:1016–8.
56. Doty RL, Reyes PF, Gregor T. Presence of both odor identification and detection deficits in Alzheimer's disease. Brain Res Bull 1987;18:597–600.
57. Doty RL, Deems DA, Stellar S. Olfactory dysfunction in parkinsonism: a general deficit unrelated to neurologic signs, disease stage, or disease duration. Neurology 1988;38:1237–44.
58. Turetsky BI, Moberg PJ, Yousem DM, et al. Reduced olfactory bulb volume in patients with schizophrenia. Am J Psychiatry 2000;157:828–30.
59. Graves AB, Bowen JD, Rajaram L, et al. Impaired olfaction as a marker for cognitive decline: interaction with apolipoprotein E epsilon4 status. Neurology 1999;53:1480–7.
60. Devanand DP, Michaels-Marston KS, Liu X, et al. Olfactory deficits in patients with mild cognitive impairment predict Alzheimer's disease at follow-up. Am J Psychiatry 2000;157:1399–405.
61. Gomez-Isla T, Price JL, McKeel DW, et al. Profound loss of layer II entorhinal cortex neurons occurs in very mild Alzheimer's disease. J Neurosci 1996;16: 4491–500.

62. Doty RL, Bagla R, Kim N. Physostigmine enhances performance on an odor mixture discrimination test. Physiol Behav 1999;65:801–4.
63. Serby M, Flicker C, Rypma B, et al. Scopolamine and olfactory function. Biol Psychiatry 1990;28:79–82.
64. Doty RL, Stern MB, Pfeiffer C, et al. Bilateral olfactory dysfunction in early stage treated and untreated idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 1992;55:138–42.
65. Markopoulou K, Larsen KW, Wszolek EK, et al. Olfactory dysfunction in familial parkinsonism. Neurology 1997;49:1262–7.
66. Doty RL, Li C, Mannon LJ, Yousem DM. Olfactory dysfunction in multiple sclerosis. N Engl J Med 1997;336:1918–9.
67. Doty RL, Li C, Mannon LJ, Yousem DM. Olfactory dysfunction in multiple sclerosis: relation to longitudinal changes in plaque numbers in central olfactory structures. Neurology 1999;53:880–2.
68. Doty RL, Hastings L. Neurotoxic exposure and olfactory impairment. Clin Occup Environ Med 2001; 1:1–29.
69. Frye RE, Schwartz BS, Doty RL. Dose-related effects of cigarette smoking on olfactory function. JAMA 1990;263:1233–6.
70. Chitanondh H. Stereotaxic amygdalotomy in the treatment of olfactory seizures and psychiatric disorders with olfactory hallucination. Confin Neurol 1966;27:181–96.
71. Leopold DA, Schwob JE, Youngentob SL, et al. Successful treatment of phantosmia with preservation of olfaction. Arch Otolaryngol Head Neck Surg 1991;117:1402–6.
72. Miller IJ. Anatomy of the peripheral taste system. In: Doty RL, editor. Handbook of olfaction and gustation. New York: Marcel Dekker; 1995. p. 521–47.
73. Morris-Witman J, Sego R, Brinkley L, Dolce C. The effects of sialoadenectomy and exogenous EGF on taste bud morphology and maintenance. Chem Senses 2000;25:9–19.
74. Farbman AI. Renewal of taste bud cells in rat circumvallate papillae. Cell Tissue Kinet 1980;13:349–57.
75. Beidler LM, Smallman RL. Renewal of cells within taste buds. J Cell Biol 1965;27:263–72.
76. Shepherd GM. Neurobiology. 2nd ed. New York: Oxford University Press; 1994.
77. Takami S, Getchell TV, McLaughlin SK, et al. Human taste cells express the G protein alpha-gustducin and neuron-specific enolase. Brain Res Mol Brain Res 1994;22:193–203.
78. Doty RL, Bagla R, Mogensen M, Mirza N. NaCl thresholds: relationship to anterior tongue locus, area of stimulation, and number of fungiform papillae. Physiol Behav 2001;72:373–8.
79. Bagla R, Klasky B, Doty RL. Influence of stimulus duration on a regional measure of NaCl taste sensitivity. Chem Senses 1997;22:171–5.
80. Kurihara K, Kurihara Y, Beidler LM. Isolation and mechanism of taste modifiers; taste modifying protein and gymnemic acids. In: Pfaffman C, editor. Olfaction and taste. New York: Rockefeller University Press; 1969. p. 450–62.
81. Rolls ET. Central taste anatomy and neurophysiology. In: Doty RL, editor. Handbook of olfaction and gustation. New York: Marcel Dekker; 1995. p. 549–73.
82. Brodal A. Neurological anatomy. New York: Oxford University Press; 1981.
83. Smith DV, Shipley MT. Anatomy and physiology of taste and smell. J Head Trauma Rehabil 1992;7:1–14.
84. Beckstead RM, Morse JR, Norgren R. The nucleus of the solitary tract in the monkey: projections to the thalamus and brain stem nuclei. J Comp Neurol 1980;190:259–82.
85. Pritchard TC, Hamilton RB, Morse JR, Norgren R. Projections of thalamic gustatory and lingual areas in the monkey, *Macaca fascicularis*. J Comp Neurol 1986;244:213–28.
86. Netter FH. The CIBA collection of medical illustrations. Vol 1. Nervous system. New York: Ciba Pharmaceutical Corporation; 1964.
87. Cerf B, Lebihan D, Van De Moortele PF, et al. Functional lateralization of human gustatory cortex related to handedness disclosed by fMRI study. Ann N Y Acad Sci 1998;855:575–8.
88. Faurion A, Cerf B, Le BD, Pillias AM. fMRI study of taste cortical areas in humans. Ann N Y Acad Sci 1998;855:535–45.
89. Pritchard TC. The primate gustatory system. In: Getchell TV, Doty RL, Bartoshuk LM, Snow JB, editors. Smell and taste in health and disease. New York: Raven Press; 1991. p. 109–25.
90. Small DM, Jones-Gotman M, Zatorre RJ, et al. A role for the right anterior temporal lobe in taste quality recognition. J Neurosci 1997;17:5136–42.
91. Matsuda T, Doty RL. Regional taste sensitivity to NaCl: relationship to subject age, tongue locus and area of stimulation. Chem Senses 1995;20: 283–90.
92. Frank ME, Hettinger TP, Clive JM. Current trends in measuring taste. In: Doty RL, editor. Handbook

of olfaction and gustation. New York: Marcel Dekker; 1995. p. 669–88.
93. Frank ME, Smith DV. Electrogustometry: a simple way to test taste. In: Getchell TV, Doty RL, Bartoshuk LM, Snow JB, editors. Smell and taste in health and disease. New York: Raven Press; 1991. p. 503–14.
94. Combarros O, Sanchez-Juan P, Berciano J, De Pablos C. Hemiageusia from an ipsilateral multiple sclerosis plaque at the midpontine tegmentum. J Neurol Neurosurg Psychiatry 2000; 68:796.
95. Fujikane M, Nakazawa M, Ogasawara M, et al. [Unilateral gustatory disturbance by pontine infarction] [Japanese]. Rinsho Shinkeigaku 1998;38:342–3.
96. Weiffenbach JM. Taste and smell perception in aging. Gerodontology 1984;3:137–46.
97. Schiffman SS. Taste and smell losses in normal aging and disease. JAMA 1997;278:1357–62.
98. Bartoshuk LM, Miller IJ Jr. Taste perception, taste bud distribution, and spatial relationships. In: Getchell TV, Doty RL, Bartoshuk LM, Snow JB, editors. Smell and taste in health and disease. New York: Raven Press; 1991. p. 205–33.
99. Gorsky M, Silverman SJ, Chinn H. Clinical characteristics and management outcome in the burning mouth syndrome. An open study of 130 patients. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1991;72:192–5.
100. Tourne LP, Fricton JR. Burning mouth syndrome. Critical review and proposed clinical management. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1992;74:158–67.
101. Fujikane M, Itoh M, Nakazawa M, et al. [Cerebral infarction accompanied by dysgeusia—a clinical study on the gustatory pathway in the CNS] [Japanese]. Rinsho Shinkeigaku 1999;39:771–4.
102. Kojima Y, Hirano T. [A case of gustatory disturbance caused by ipsilateral pontine hemorrhage] [Japanese]. Rinsho Shinkeigaku 1999;39: 979–81.
103. Lee BC, Hwang SH, Rison R, Chang GY. Central pathway of taste: clinical and MRI study. Eur Neurol 1998;39:200–3.
104. Sumner D. Post-traumatic ageusia. Brain 1967; 90:187–202.
105. Scrivani SJ, Keith DA, Kulich R, et al. Posttraumatic gustatory neuralgia: a clinical model of trigeminal neuropathic pain. J Orofac Pain 1998;12:287–92.
106. Helcer M, Schnarch A, Benoliel R, Sharav Y. Trigeminal neuralgic-type pain and vascular-type headache due to gustatory stimulus. Headache 1998;38:129–31.
107. Ohtuka K, Tomita H, Yamauchi Y, Kitagoh H. [Taste disturbance after tonsillectomy] [Japanese]. Nippon Jibiinkoka Gakkai Kaiho 1994;97:1079–88.
108. Donati F, Pfammatter JP, Mauderli M, Vassella F. Neurologische Komplikationen nach Tonsillektomie. Schweiz Med Wochenschr 1991;121:1612–7.
109. Arnhold-Schneider M, Bernemann D. Uber die Haufigkeit von Geschmacksstorungen nach Tonsillektomie. HNO 1987;35:195–8.
110. Ohtsuka K, Tomita H, Murakami G. [Anatomical study of the tonsillar bed: the topographical relationship between the palatine tonsil and the lingual branch of the glossopharyngeal nerve] [Japanese]. Nippon Jibiinkoka Gakkai Kaiho 1994; 97:1481–93.
111. Walker RP, Gopalsami C. Laser-assisted uvulopalatoplasty: postoperative complications. Laryngoscope 1996;106:834–8.
112. Evers KA, Eindhoven GB, Wierda JM. Transient nerve damage following intubation for trans-sphenoidal hypophysectomy. Can J Anaesth 1999;46: 1143–5.
113. Ostergaard M, Kristensen BB, Mogensen TS. [Reduced sense of taste as a complication of the laryngeal mask use] [Danish]. Ugeskr Laeger 1997; 159:6835–6.
114. Bull TR. Taste and the chorda tympani. J Laryngol Otol 1965;79:479–93.
115. Shafer DM, Frank ME, Gent JF, Fischer ME. Gustatory function after third molar extraction. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1999;87:419–28.
116. Ackerman BH, Kasbekar N. Disturbances of taste and smell induced by drugs. Pharmacotherapy 1997;17:482–96.
117. Stricker BH, Van RM, Sturkenboom MC, Ottervanger JP. Taste loss to terbinafine: a case-control study of potential risk factors. Br J Clin Pharmacol 1996;42:313–8.
118. Bartoshuk LM. Chemosensory alterations and cancer therapies. NCI Monogr 1990;179–84.
119. Della Fera MA, Mott AE, Frank ME. Iatrogenic causes of taste disorders: radiation therapy, surgery, and medication. In: Doty RL, editor. Handbook of olfaction and gustation. New York: Marcel Dekker; 1995. p. 785–91.

120. Conger AD. Loss and recovery of taste acuity in patients irradiated to the oral cavity. Radiat Res 1973;53:338–47.
121. Chambers KC, Bernstein IL. Conditioned flavor aversions. In: Doty RL, editor. Handbook of olfaction and gustation. New York: Marcel Dekker; 1995. p. 745–73.
122. Erick M. Idiopathic dysgeusia associated with blood transfusion: a case report [letter]. J Am Diet Assoc 1996;96:450.
123. Panayiotou H, Small SC, Hunter JH, Culpepper RM. Sweet taste (dysgeusia). The first symptom of hyponatremia in small cell carcinoma of the lung. Arch Intern Med 1995;155:1325–8.
124. Yanagisawa K, Bartoshuk LM, Catalanotto FA, et al. Anesthesia of the chorda tympani nerve and taste phantoms. Physiol Behav 1998;63:329–35.
125. Helms JA, Della-Fera MA, Mott AE, Frank ME. Effects of chlorhexidine on human taste perception. Arch Oral Biol 1995;40:913–20.
126. Mattes RD, Kare MR. Gustatory sequelae of alimentary disorders. Dig Dis 1986;4:129–38.
127. Grushka M, Epstein J, Mott A. An open-label, dose escalation pilot study of the effect of clonazepam in burning mouth syndrome. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1998; 86:557–61.

CHAPTER 28

Cellular Biology of the Immune System

Joel M. Bernstein, MD, PhD

Immunology, the study of the immune system, grew out of the common observation that human beings who recover from certain infectious diseases are thereafter immune to the disease, that is, they rarely develop the same disease again. Immunity is highly specific: an individual who recovers from measles is protected against the measles virus but not against other common viruses, such as mumps or chickenpox. Such specificity is a fundamental characteristic of immune responses. In other vertebrate systems, cells such as macrophages and neutrophils also play an important role in defending vertebrates against infection, but they are only one part of a much more complex and sophisticated defense strategy. This chapter is devoted to the specific immune response.

Many of the responses of the human immune system initiate the destruction and elimination of invading microorganisms and any toxic molecules produced by them. Because these immune reactions are destructive, it is essential that they be made only in response to molecules that are foreign to the host and not to those of the host itself, resulting in autoimmunity. This ability to distinguish foreign molecules from self-molecules, in general, is another fundamental feature of the human immune system. Occasionally, it fails to make this distinction and reacts destructively against the host's own molecules; such autoimmune diseases can be fatal.

Although the immune system evolved to protect vertebrates from infection by microorganisms and larger parasites, most of what we know about immunity has come from the studies of the responses of laboratory animals to injections of noninfectious substances, such as foreign proteins and polysaccharides. Almost any macromolecule, as long as it is foreign to the recipient, can induce an immune response; any substance capable of eliciting an immune response is referred to as an *antigen* (*anti*body *gen*erator). Remarkably, the immune system can distinguish between antigens that are very similar, such as between two proteins that differ in only a single amino acid or between two optical isomers of the same molecule.[1]

This chapter specifically addresses two broad classes of immune responses: antibody response and cell-mediated immune responses. *Antibody responses* involve the production of antibodies, which are proteins called *immunoglobulins* (Igs). The antibodies circulate in the bloodstream and permeate the other body fluids, where they bind specifically to the foreign antigen that induced them. Binding by antibody inactivates viruses and bacteria and bacterial toxins by blocking their ability to bind to receptors on host cells. Antibody binding also marks invading microorganisms for destruction, either by making it easier for a phagocytic cell to ingest them or by activating a system of blood proteins, collectively called *complement*, which kills the invaders. The complement system is addressed in this chapter as well.

Cell-mediated immune responses, the second class of immune responses, involve the production of specialized cells, which react with foreign antigens on the surface of other host cells. The reacting cell, for example, can kill a virus-infected host cell that has viral antigens on its surface, thereby eliminating the infected cell before the virus has replicated. In other cases, the reacting cell secretes chemical signals that activate macrophages to destroy the invading microorganisms. The main challenge in immunology has been to understand how the immune system specifically recognizes and reacts aggressively to a virtually unlimited number of different foreign macromolecules but avoids reacting against the tens of thousands of different self-macromolecules made by host cells. Thus, initially in this chapter, the functional and structural features of antibodies that enable them to recognize and destroy extracellular antigens are considered. After discussing how this antibody diversity is generated,

this chapter then considers the special features of cell-mediated immune responses that are crucial in the defense against intracellular microorganisms. Wherever possible, the fundamental principles of immunobiology are applied to diseases of the middle ear and sinuses and particular reference is made to immunobiology of the tonsils and adenoids. Finally, an overview of three classes of soluble mediators of inflammation is discussed using the nasal polyp as a model of chronic inflammation.

CELLULAR BASIS OF IMMUNITY

The Human Immune Response Is Composed of Trillions of Lymphocytes

The cells responsible for immune specificity belong to a class of white blood cells known as lymphocytes.[2] They are found in large numbers in the blood and the lymph and in specialized lymphoid organs such as the thymus, bone marrow, lymph nodes, spleen, appendix, tonsils, and adenoids (Figure 28–1).

There are approximately 2×10^{12} lymphocytes in the human body, which make the immune system comparable in cell mass to the liver or brain.

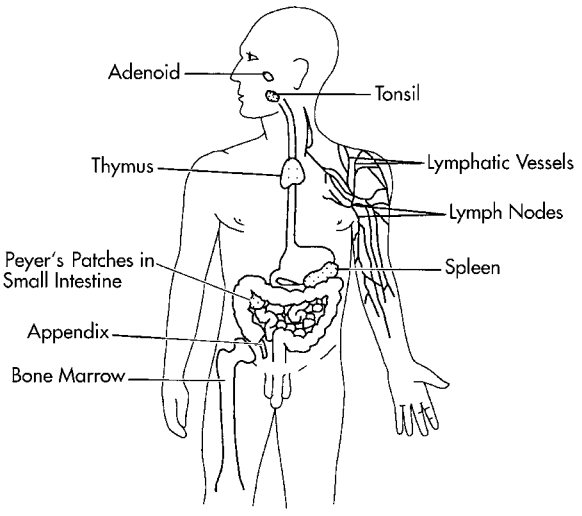


FIGURE 28–1. Human lymphoid organs. Lymphocytes develop in the thymus and bone marrow, which are therefore referred to as *primary* (or *central*) *lymphoid organs*. The newly formed lymphocytes migrate from these primary regions to *secondary* (or *peripheral*) *lymphoid organs*, where they can react with antigen. Only some of the secondary lymphoid organs are shown; many lymphocytes, for example, are found in the skin and lungs.

About 40 years ago, it was discovered that the two major classes of immune responses are mediated by different classes of lymphocytes: *T cells*, which develop in the thymus and are responsible for cell-mediated immunity, and *B cells*, which in mammals develop in the adult bone marrow or the fetal liver and produce antibodies.[3] This dichotomy of the lymphoid system was initially demonstrated in animals with experimentally induced immune deficiencies. These studies on immune-deficient animals demonstrated that individuals deficient in T cells were unable to make cell-mediated immune responses and had somewhat impaired antibody responses. We now know that there are two main classes of T cells; helper T cells and cytotoxic T cells.[4] Helper T cells enhance the response of other white blood cells, and some of these T cells help B cells to make antibody responses. Cytotoxic T cells, by contrast, kill infected cells; because they are involved directly in defense against infection, unlike helper cells, they, together with B cells, are sometimes referred to as *effector cells*.

Lymphocytes Develop in Primary Lymphoid Organs and React with Foreign Antigens in Secondary Lymphoid Organs

It is now known that lymphocytes develop from *pluripotent* hemopoietic *stem cells*, which give rise to all of the blood cells, including red blood cells, white blood cells, and platelets.[5] These stem cells are located primarily in the liver in fetuses and in the bone marrow of adults. T cells develop in the thymus from precursor cells that migrate in from the hemopoietic tissues via the blood in mammals (Figure 28–2). B cells develop from stem cells in the hemopoietic tissues themselves. Because they are sites in which lymphocytes develop from precursor cells, the thymus and the bone marrow are referred to as primary lymphoid organs. Other lymphocytes, however, mature and migrate via the blood to the secondary peripheral lymphoid organs such as the lymph nodes, spleen, and epithelium-associated lymphoid tissue found in the gastrointestinal tract, respiratory tract, skin, tonsils, and adenoids. It is chiefly in the secondary lymphoid organs that T and B cells react with incoming foreign antigens. The bone marrow in mammals continues to generate large numbers of new B cells, approximately 5×10^7/day in the mouse throughout life.[6]

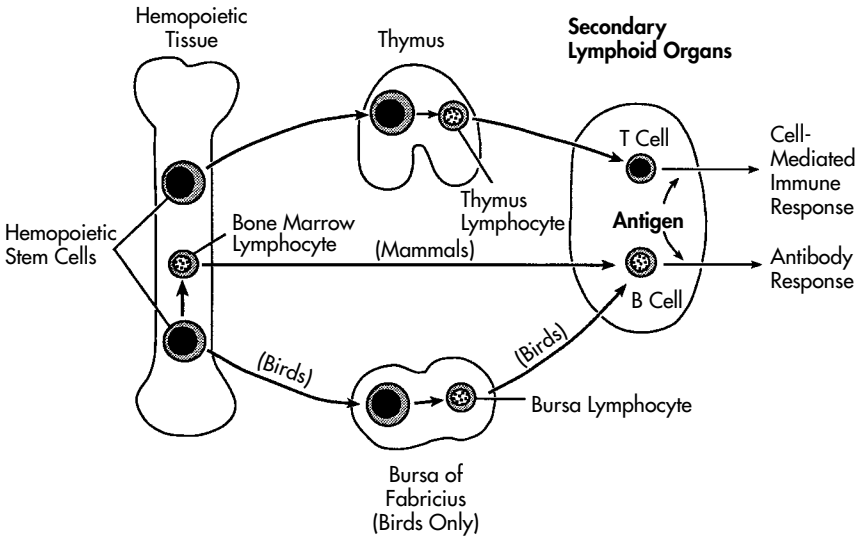


FIGURE 28–2. The development of T and B cells. The primary lymphoid organs are where lymphocytes develop from precursor cells. The bursa of Fabricius (birds only) is the lymphoid organ for B-cell development, whereas in mammals, the bone marrow is the predominant precursor of B cells. T cells in secondary lymphoid organs are primarily involved in cell-mediated immune response, whereas B cells are primarily involved in antibody responses.

Cell-surface markers make it possible to distinguish and separate T and B cells. T and B cells become morphologically distinguishable only after they have been stimulated by antigen. Both are activated by antigens to proliferate and mature further. Activated B cells develop into antibody-secreting cells, the most mature of which are *plasma cells* that are filled with an extensive rough endoplasmic reticulum. In contrast, activated T cells contain very little endoplasmic reticulum and do not secrete antibodies, although they do secrete a variety of mediators called *lymphokines*, *interleukins* (ILs), or *cytokines*. Since both T and B cells occur in all secondary lymphoid organs, it has been necessary to find ways to distinguish and separate the two cell types and their various subtypes to study their individual properties. Fortunately, there are many differences in the plasma membrane proteins of the different types of lymphocytes that can serve as distinguishing markers. Antibodies that react with the *Thy-1 protein*, for example, which is found on T cells but not on B cells in mice, are widely used to remove or purify T cells from a mixed population of mouse lymphocytes.[7] Similarly, antibodies against *CD4* and *CD8 proteins*, which are discussed later, are widely used to distinguish and separate helper T cells and cytotoxic T cells, respectively, in both mice and humans.[8]

The Immune System Works by Clonal Selection

The most remarkable feature of the immune system is that it can respond to millions of different foreign antigens in a highly specific way, for example, by making antibodies that react specifically with the antigen and induce their production. How this immune system produces such diversity of specific antibodies led to the emergence of the theory in the 1950s that came to be known as the *clonal selection theory*.[9] According to this theory, each animal first randomly generates a vast diversity of lymphocytes, and then those cells that react against the foreign antigens that the animal actually encounters are specifically selected for action. The theory is based on the proposition that during development, each lymphocyte becomes committed to react with a particular antigen before ever being exposed to it. Each cell expresses this commitment in the form of cell-surface receptor proteins that specifically fit the antigen. The binding of antigen to the receptors activates the cell, causing it to proliferate and mature. Therefore, a foreign antigen selectively stimulates those cells that express complementary antigen-specific receptors and are thus already committed to respond to it. This is what makes the immune response antigen specific.

The term "clonal" in clonal selection arises from the postulate that the immune system is composed of millions of different families, or *clones* of cells, each consisting of T or B cells descended from a common ancestor. Each ancestral cell is already committed to make one particular antigen-specific receptor protein, and all cells in a clone have the same antigen specificity. It is now known that the antigen-specific receptors on both T and B cells are encoded by genes that are assembled from a series of

gene segments by a unique form of genetic recombination that occurs early in the cell's development, beginning before it has encountered antigen.

Most antigens stimulate many different lymphocyte clones. Those parts of an antigen that combine with the antigen-binding site on the antibody molecule or on a lymphocyte receptor are called *antigenic determinants* or *epitopes*.[10] Most antigens have a variety of antigenic determinants that stimulate the production of antibodies or T-cell responses. Some determinants are more antigenic than others, so the reaction to them may dominate the overall response. Such determinants are said to be *immunodominant.* Even an antigen that activates many clones will stimulate only a tiny fraction of the total lymphocyte population. To ensure that these few lymphocytes are exposed to the antigen, antigens are generally collected by specialized *antigen-presenting cells* in secondary lymphoid organs, through which T and B cells continuously circulate. Antigens that enter through the gut are trapped by gut-associated lymphoid tissue, those that enter through the skin or respiratory tract are retained locally and/or are transported via the lymph to local lymph nodes, and those that enter the blood are filtered out in the spleen. The antigen-presenting cells of the tonsil, also called M cells, are in the crypts and present antigen to T lymphocytes and subsequently to B lymphocytes in the crypts.[11] However, in the germinal centers of the tonsil or adenoids, B cells may present antigen to helper T cells.[12] This will be discussed in a further section of this chapter.

The majority of T and B lymphocytes continuously recirculate between the blood and the secondary lymphoid tissue. In a lymph node, for example, lymphocytes leave the bloodstream, squeezing out between specialized endothelial cells, so-called high endothelial venules; after percolating through the node, they accumulate in small lymphatic vessels that leave the node and connect with other lymphatic vessels that pass through other lymph nodes downstream and finally pass into a large vessel called the thoracic duct, which carries them back to the blood. Lymphocyte recirculation depends on specific interactions between the lymphocyte cell surface and the surface of specialized endothelial cells lining small veins of the secondary lymphoid organs; because their endothelial cells are unusually tall, they are called postcapillary high endothelial venules (Figure 28–3). These migrations are guided by various *homing receptors* on lymphocytes and by the ligands for these receptors (often called *counter-receptors*) on endothelial cells.[13] Both receptors and counter-receptors have been identified by monoclonal antibodies that bind to the surface of either lymphocytes or the specialized high endothe-

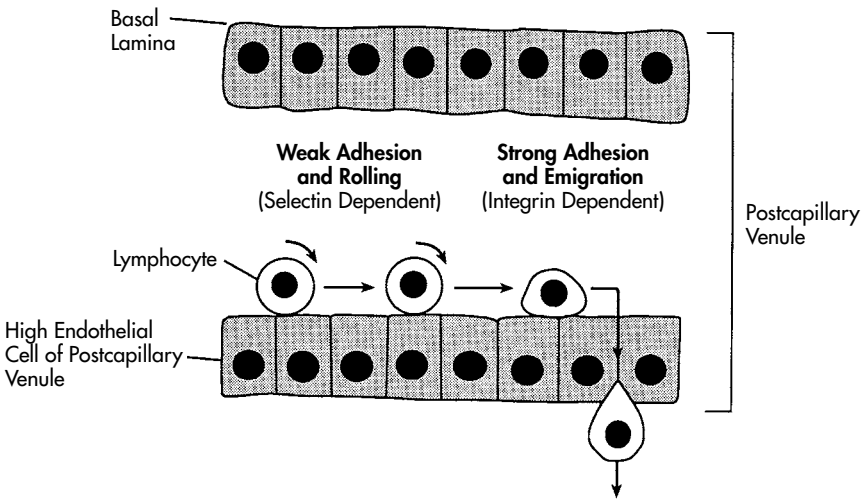


FIGURE 28–3. Migration of a lymphocyte out of the bloodstream into a lymph node. A circulating lymphocyte adheres weakly to the surface of the specialized high endothelial cells in a postcapillary venule in a lymph node. This initial adhesion, mediated by E-selectin on the lymphocyte surface, is sufficiently weak that it enables the lymphocyte to roll along the surface of the endothelial cells. The lymphocyte rapidly activates a stronger adhesion system mediated by an integrin, which enables the cell to stop rolling and migrate out of the venule between the endothelial cells.

lial cells and inhibit the ability of the lymphocytes both to bind the endothelial cells and tissue sections of secondary lymphoid organs and allows them to recirculate in vivo. Lymphocyte migration into lymph nodes, for example, depends on the cell adhesion protein called *E-selectin*, which belongs to the selectin family of cell-surface selectins, which are glycoproteins.[14] This homing receptor, which is present on most lymphocytes, binds to a specific sugar group on a highly glycosylated, mucin-like counter-receptor that is expressed exclusively on the surface of high endothelial cells lining postcapillary venules and lymph nodes. E-selectin binding causes the lymphocytes to adhere weakly to the endothelial cells and to roll slowly along their surface. The rolling continues until another, much stronger adhesion system is activated. This strong adhesion, which is mediated by a member of the *integrin* family of cell adhesion molecules on the lymphocyte surface, allows the lymphocytes to stop rolling and crawl out of the blood vessel into the lymph node.[15]

Other homing receptors on lymphocytes are found to be responsible for the subsequent segregation of T and B cells into distinct areas of the lymph node. For example, B cells are primarily restricted to the follicle and mantle zone of the tonsil and adenoids, and T cells are primarily located between the secondary follicles and in the reticular epithelium of these organs.[16] Once they are activated by antigen, most lymphocytes lose many of their homing receptors and acquire new ones: instead of migrating though lymphoid organs, they migrate through the nonlymphoid tissues to sites of inflammation. The migration of activated lymphocytes and other white blood cells into sites of inflammation is largely mediated by other combinations of selectins and integrins. For example, in nasal polyps, there is an up-regulation of eosinophils because of the integrin VLA-4 (very late activation antigen 4), which is increased on the surface of eosinophils, and a counter-receptor on the small blood vessels of venules in the nasal polyp called VCAM-1 (vascular cell adhesion molecule). Very late activation antigen 4 and VCAM-1 are receptor and counter-receptor and account for the significant increase of eosinophils in nasal polyps.[17]

The immune system, like the nervous system, can remember. This is why we develop lifelong immunity to many common viral diseases after our initial exposure to the virus, and this is why immunization works. If an animal is injected once with antigen A, its immune response, either antibody or cell mediated, will appear after a lag period of several days, rise rapidly and exponentially, and then more gradually fall again. This is characteristic of a *primary immune response* occurring on an animal's first exposure to an antigen. If some weeks, months, or even years are allowed to pass and the animal is reinjected with another antigen A, it would usually produce a *secondary immune response*, which is very different from the primary immune response. The lag period is shorter, and the response is greater. These differences indicate that the animal has "remembered" its first exposure to antigen A. Therefore, the secondary response reflects antigen-specific immunologic memory for antigen A (Figure 28–4).

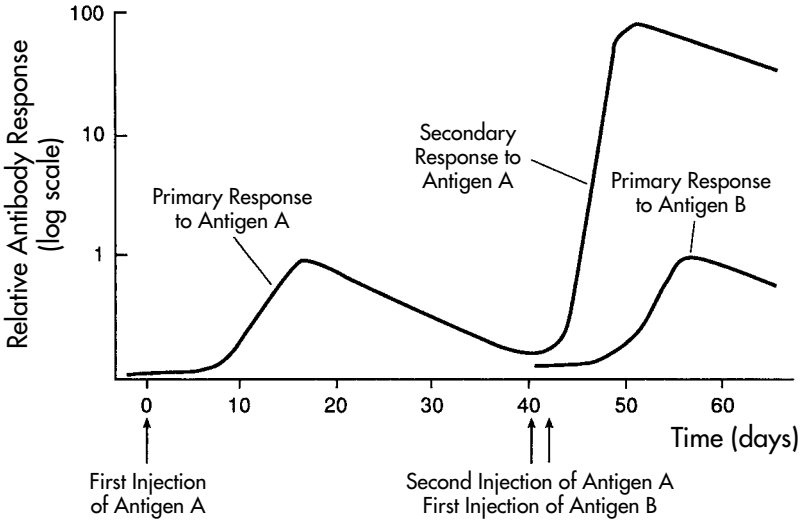


FIGURE 28–4. Primary and secondary antibody responses. The secondary response induced by a secondary exposure to antigen A is faster and greater than the primary response and is specific for A, indicating that the immune system has specifically remembered encountering antigen A before. Evidence for the same type of immunologic memory is obtained if T-cell-mediated responses rather than B-cell antibody responses are elicited.

The clonal selection theory provides a useful conceptual framework for understanding the cellular basis for immunologic memory. In an adult animal, the T and B cells in the secondary lymphoid organs are a mixture of cells in at least three stages of maturation, which can be designated *virgin* (or naive) cells, *memory* cells, and *activated* cells. When *virgin* cells encounter antigen for the first time, some of them are stimulated to multiply and become *activated cells*, which we define as cells that are actively engaged in making a response (activated T cells carry out cell-mediated responses, whereas activated B cells secrete antibodies). Other virgin cells are stimulated to multiply and mature instead into *memory cells*, which are not themselves making a response but are readily induced to become activated cells by a later encounter by the same antigen (Figure 28–5). One reason for the increased responsiveness of memory cells is that their receptors have a higher affinity for antigen. Memory T cells seem to respond to antigen more readily than virgin T cells, not only because they have higher-affinity receptors for antigen but also because they adhere more strongly to other cells and transduce extracellular signals more efficiently. Thus, immunologic memory is generated during the primary response in part because the proliferation of antigen-triggered virgin cells creates many memory cells, a process known as *clonal expansion* (see Figure 28–5).

The failure to respond to self-antigens is attributable to acquired immunologic tolerance.[18] How is the immune system able to distinguish foreign molecules from self-molecules? One possibility might be that an animal inherits genes that encode receptors for foreign antigens but not self-antigens, so its immune system is genetically constituted to respond only to foreign antigens. Alternatively, the immune system could be inherently capable of responding to both foreign and self-antigens but could "learn" during development not to respond to self-antigens. The latter explanation has been shown to be correct. The mechanisms of this acquired immunologic tolerance are critically important so that humans, in general, do not develop immunologic disease directed against the self, which results in autoimmunity, which could lead to death. The resulting state of antigen-specific immunologic unresponsiveness is known as *acquired immunologic tolerance*.

The learning process that leads to self-tolerance can involve either killing the self-reactive lym-

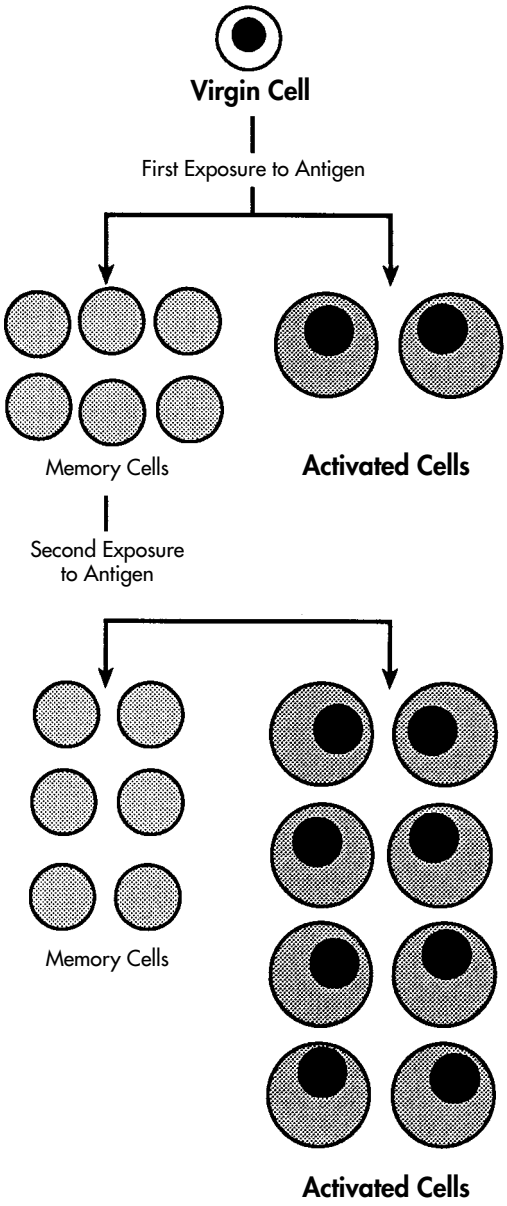


FIGURE 28–5. A model for the cellular basis of immunologic memory. When virgin T or B cells are stimulated by their specific antigens, they proliferate and mature; some become activated to make a response, whereas others become memory cells. During a subsequent exposure to the same antigen, the memory cells responded more readily than did the virgin cells: they proliferate and give rise to activated cells and to more memory cells. In the model shown, an individual virgin cell can give rise to either a memory cell or an activated cell, depending on the conditions. In an alternative model (not shown), the virgin cells that mature into memory cells are different from those that mature into activated cells. It is not known which of these models is correct.

phocyte (*clonal deletion*) or functionally inactivating cells and leaving them alive (*clonal anergy*). Tolerance to self-antigens sometimes breaks down, causing T or B cells or both to react against their own tissue antigens. Myasthenia gravis is an example of such an *autoimmune disease*.[19] Affected individuals make antibodies against the acetylcholine receptors on their own skeletal muscle cells; the antibodies interfere with the normal functioning of the receptors so that such patients become weak and can die because they cannot breathe.

In summary, the immune system evolved to defend vertebrates against infection. It is composed of millions of lymphocyte clones. The lymphocyte in each clone shares a unique cell-surface receptor that enables it to bind to a particular antigenic determinant consisting of a specific arrangement of atoms on a part of a molecule. There are two classes of lymphocytes: B cells, which are produced in the bone marrow and make antibodies, and T cells, which are produced in the thymus and make cell-mediated immune responses. Beginning early in lymphocyte development, many lymphocytes that would react against antigenic determinants on self-macromolecules are eliminated or inactivated; as a result, the immune system normally reacts only to foreign antigens. The binding of a foreign antigen to a lymphocyte initiates a process by the cell that helps to eliminate the antigen. As part of the response, some lymphocytes proliferate and mature into memory cells that are able to respond more readily to antigen than do virgin cells. Thus, the next time the same antigen is encountered, the immune response to it is much faster and stronger. The remainder of this chapter discusses the structure and property of antibodies that are the result of B-cell maturation and the development of T-cell immunity.

FUNCTIONAL PROPERTIES OF ANTIBODIES

Synthesized exclusively by B cells, antibodies are produced in millions of forms, each with a different amino acid sequence and a different binding site for antigen.[20] Collectively, these antibodies are called *Igs*. They are among the most abundant protein components in the blood, constituting about 20% of the total plasma protein by weight.[21] All antibody molecules made by an individual B cell have the same antigen-binding site. The antibodies made by a newly formed B cell are not secreted. Instead, they are inserted into the plasma membrane, where they serve as a receptor for antigen. Each B cell has approximately 10^5 such antibody molecules in its plasma membrane.[22] Each of these antibody molecules is noncovalently associated with an invariant set of transmembrane polypeptide chains that are involved in passing signals to the cell interior when the extracellular binding site of the antigen is occupied by antigen. Each B cell produces a single species of antibody with a unique antigen-binding site. When a virgin or memory B cell is activated by antigen, it proliferates and matures to become an antibody-secreting cell (plasma cell). The activated cells make and secrete large amounts of soluble rather than membrane-bound antibody, which has the same unique antigen-binding site as the cell-surface antibody that served earlier as the antigen receptor.[23] Plasma cells seem to have committed so much of their protein synthesizing machinery to making antibody that they are incapable of further growth and division, and most die after several days.

ANTIBODIES HAVE TWO IDENTICAL ANTIGEN-BINDING SITES

The simplest antibodies are Y-shaped molecules with two identical *antigen-binding sites*, one at the tip of each arm of the Y (Figure 28–6). Because of

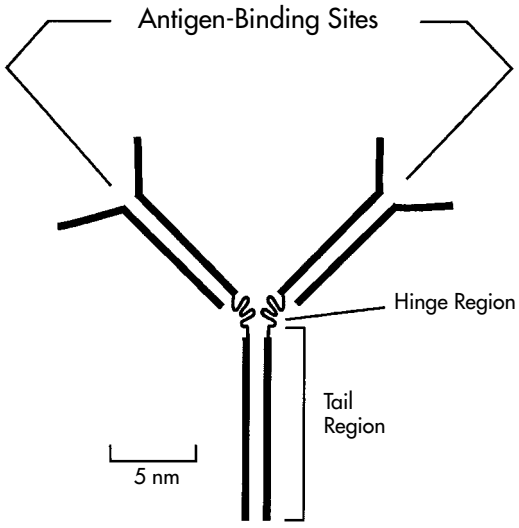


FIGURE 28–6. A simple representation of an antibody molecule. Note that its two antigen-binding sites are identical.

their two antigen-binding sites, they are said to be *bivalent*. As long as an antigen has three or more antigenic determinants, bivalent antibody molecules can cross-link it into a large lattice, which can be rapidly phagocytosed and degraded by macrophages (Figure 28–7).[24] The efficiency of antigen binding and cross-linking is greatly increased by a *flexible hinge region* in antibodies, which allows the distance between the two antigen-binding sites to vary.

The protective effect of antibodies is not due simply to their ability to bind antigen. They engage in a variety of activities that are mediated by the tail of the Y-shaped molecule. This part of the molecule determines what will happen to the antigen once it is bound to the antibody. One of several different tailed regions can confer on the antibody different functional properties such as the ability to activate complement or to bind to phagocytic cells. The basic structural unit of an antibody molecule consists of four polypeptide chains; two identical *light* (*L*) *chains*, each containing about 220 amino acids, and two identical *heavy* (*H*) *chains*, each usually containing about 440 amino acids.[25] The four chains are held together by a combination of noncovalent and covalent disulfide bonds. The molecules are composed of two identical halves, each with the same antigen-binding site, and both light and heavy chains usually cooperate to form the antigen-binding surface (Figure 28–8, A). There are five classes of heavy chains, each with different biologic properties. In higher vertebrates, there are five classes of antibodies, IgA, IgD, IgE, IgG, and IgM, each with its own class of heavy chain, α, δ, ε, γ, and μ, respectively.[26] Immunoglobulin A molecules have α chains, IgG molecules have γ chains, and so on. In addition, there are a number of subclasses of IgG and IgA; for example, there are four human IgG subclasses (IgG1, IgG2, IgG3, and IgG4), having $\gamma 1$, $\gamma 2$, $\gamma 3$, and $\gamma 4$ heavy chains, respectively.[27] The various heavy chains impart a distinctive confirmation to the hinge and tail regions of the body and give each class (and subclass) characteristic properties.

Immunoglobulin M, which has a μ heavy chain, is always the first class of antibody produced by developing B cells, although many of these cells eventually switch to making other classes of antibody, which are discussed below. The immediate precursor of the B cell, called a *pre-B cell*, initially makes μ chains, which associate with non–light-chain polypeptides, often referred to as surrogate light chains, which insert into the plasma membrane. As the synthesis of bona fide light chains increases,

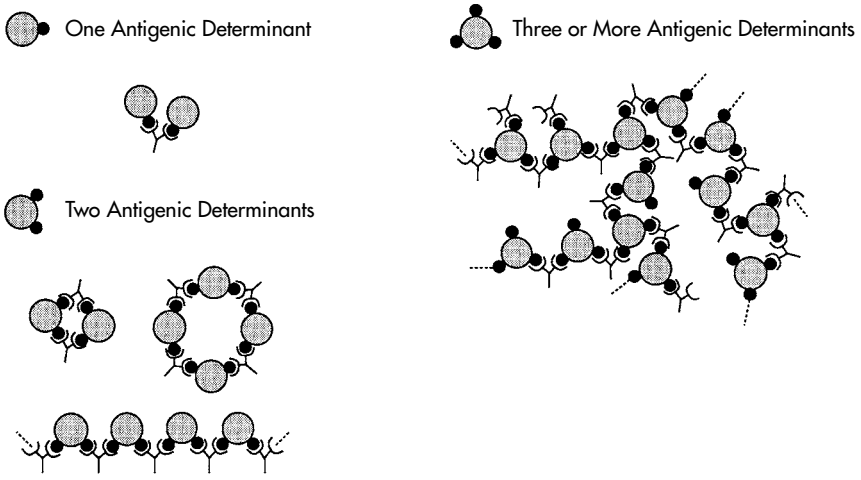


FIGURE 28–7. Antibody–antigen interactions. Because antibodies have two identical antigen-binding sites, they can cross-link antigens. The types of antibody-antigen complexes that form depend on the number of antigenic determinants on the antigen. Here a single species of antibody (a monoclonal antibody) is shown binding to antigens containing one, two, or three copies of a single type of antigenic determinant. Antigens with two antigenic determinants can form small cyclic complexes or linear chains with antibody, whereas antigens with three or more antigenic determinants can form large three-dimensional lattices that readily precipitate out of solution. Most antigens have many different antigenic determinants, and the different antibodies that recognize these different determinants can cooperate in cross-linking the antigen.

these combine with the μ chains, displacing the surrogate light chains, to form a four-chain IgM molecule (with two μ chains and two light chains), which inserts into the plasma membrane. The cell now has cell-surface receptors with which it can bind antigen, and, at this point, it is called a *virgin B cell*. Many virgin B cells soon start to produce cell-surface IgD molecules as well, with the same antigen-binding site as the IgM molecules. Finally, it is to be noted that IgM, IgD surface-positive B cells are found in the mantle zone in the palatine tonsil and nasopharyngeal tonsil.[28]

Immunoglobulin M is not only the first class of antibody to appear on the surface of the developing B cell, it is also the major class secreted into the blood in the early stages of a *primary* antibody response. In its secreted form, IgM is a pentamer composed of five four-chain units and thus has a total of 10 antigen-binding sites. Each pentamer contains one copy of another polypeptide chain called a *J chain* (joining chain).[29] The J chain is produced by IgM-secreting cells and is covalently inserted between two adjacent tails (Fc regions) (Figure 28–8, B). The binding of antigen to the Fab regions of the secreted pentameric IgM molecule induces the Fc regions to bind to and thereby activate the first component of the *complement system*.[30] When the antigen is on the surface of an invading microorganism, the resulting activation of complement unleashes a biochemical attack that kills the microorganism. Unlike IgM, IgD molecules are rarely secreted by an activated B cell, and their functions (other than receptors for antigen) are unknown.

The major class of Ig in the blood is *IgG*, which is produced in large quantities during *secondary* immune responses.[31] Besides activating the complement system, the Fc region of an IgG molecule binds to specific receptors on macrophages and neutrophils. Largely by means of such *Fc receptors*, these phagocytic cells bind, ingest, and destroy infecting microorganisms that have become coated with the IgG antibodies produced in response to the infection (Figure 28–9).

Immunoglobulin G molecules are the only antibodies that can pass from mother to fetus via the placenta. Immunoglobulin G is also secreted into the

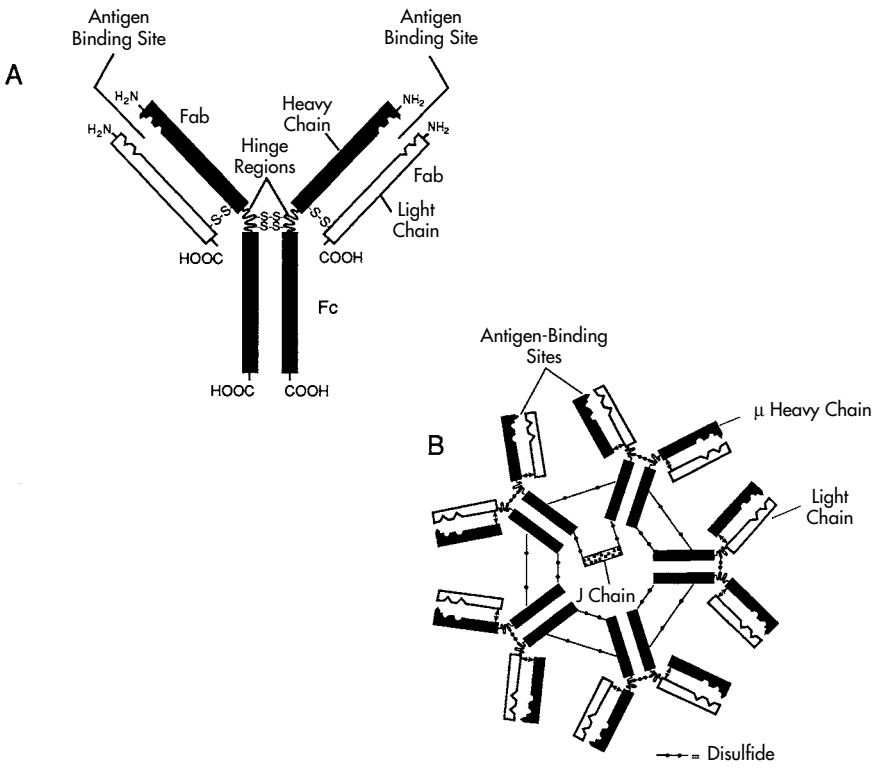


FIGURE 28–8. A, Schematic drawing of a typical antibody molecule. It is composed of two identical heavy chains and two identical light chains. Note that the antigen-binding sites are formed by a complex of the amino-terminal regions of both light and heavy chains, and the tail and hinge regions are formed by the heavy chains alone. *B*, A pentameric immunoglobulin (Ig) M molecule. The five subunits are held together by disulfide bonds. A single J chain, which has a structure similar to that of a single Ig domain, is disulfide bonded between two μ heavy chains. The J chain is required for the polymerization process. The addition of each successive four-chain IgM subunit requires a J chain, which is then discarded, except for the last one, which is retained.

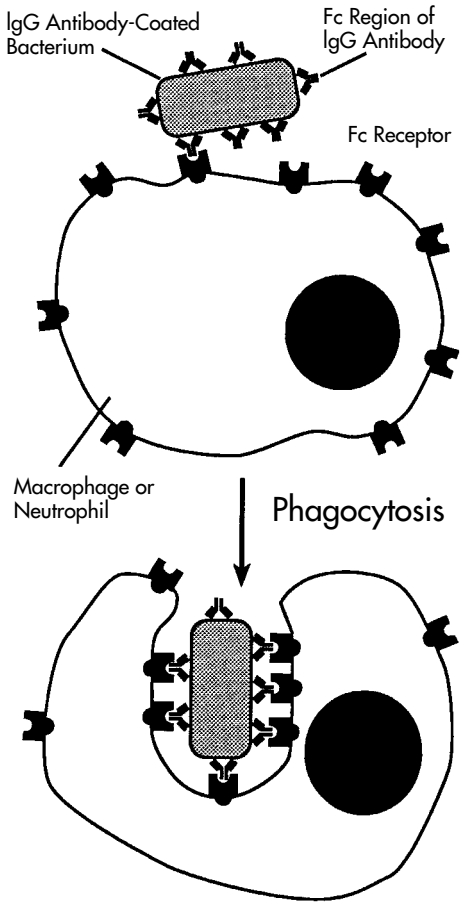


FIGURE 28–9. Antibody-activated phagocytosis. An immunoglobulin (Ig)G-antibody-coated bacterium is efficiently phagocytosed by a macrophage or neutrophil, which has cell-surface receptors able to bind the Fc region of IgG molecules. The binding of the antibody-covered bacterium to these Fc receptors activates the phagocytic process.

mother's milk and is taken up from the gut of the neonate into the blood.

Immunoglobulin A is the principal class of antibody in secretions (saliva, tears, milk, and respiratory and intestinal secretions).[32] Immunoglobulin A is transported through secretory epithelial cells from the extracellular fluid into the secreted fluid by another type of Fc receptor that is unique to secretory epithelia. A highly schematized diagram of dimeric IgA with J chain is shown in Figure 28–10, and the mechanism of transport of a dimeric IgA molecule across an epithelial surface such as, for example, the nasal mucosa is shown in Figure 28–11. The mechanism of transport of dimeric IgA into secretions of the mucosa-associated immune system is shown in Figure 28–11.

The Fc region of *IgE* molecules binds with unusually high affinity to yet another class of Fc receptors. These receptors are located on the surface of *mast cells* in tissues and on *basophils* in the blood, and the IgE molecules bound to them in turn serve as receptors for antigen.[33] Antigen binding triggers the cells to secrete a variety of biologically active amines, especially histamine. These amines cause dilation and increased permeability of blood vessels and are largely responsible for the clinical manifes-

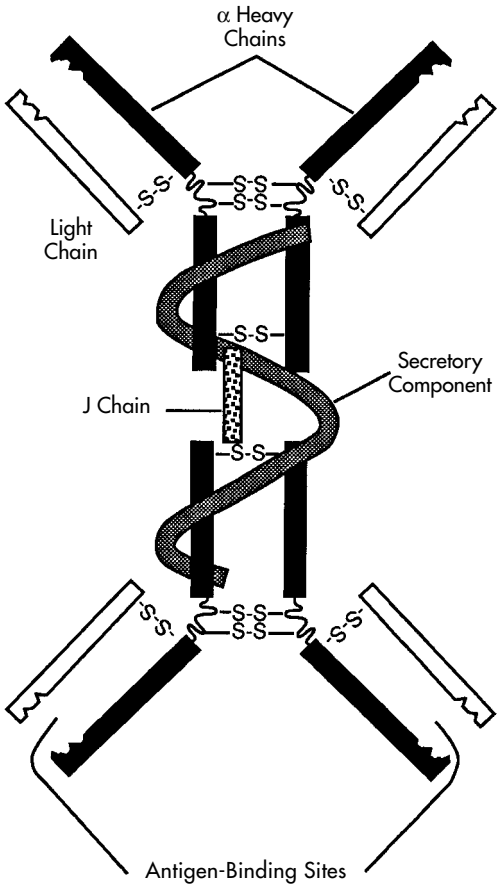


FIGURE 28–10. A highly schematized diagram of a dimeric immunoglobulin (Ig)A molecule found in secretions. In addition to the two IgA monomers, there is a single J chain and an additional polypeptide chain called the *secretory component*, which is thought to protect the IgA molecules from being digested by proteolytic enzymes in the secretions. The secretory component molecule is found on the basolateral surface of the epithelial cells and appears to be the receptor for the J chain.

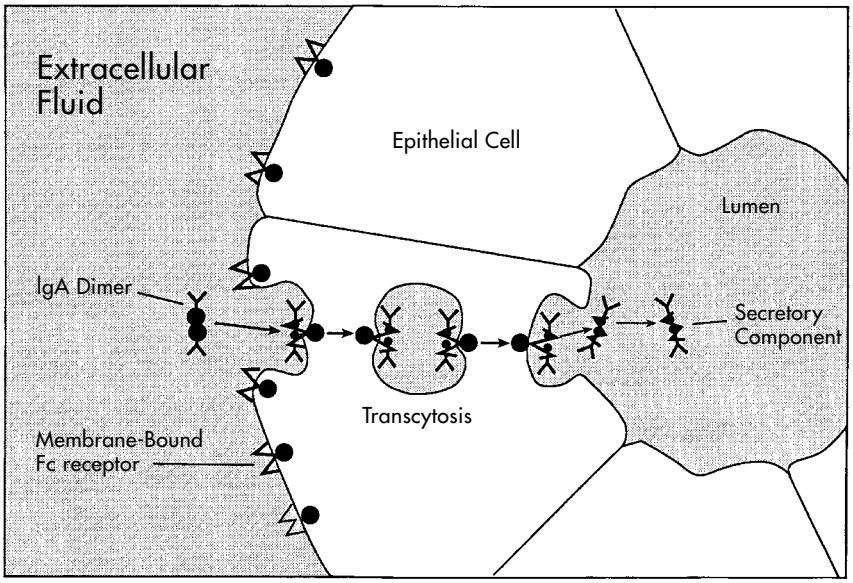


FIGURE 28–11. The mechanism of transport of a dimeric immunoglobulin (Ig)A molecule across an epithelial cell. The IgA molecule, as a J chain–containing dimer, binds to a specialized transmembrane Fc receptor protein on the nonluminal surface of the secretory epithelial cell. The surface of the receptor-IgA complexes are ingested by receptor-mediated endocytosis, transferred across the epithelial cell cytoplasm in vesicles, and then secreted into the lumen of the opposite side of the cell by exocytosis. When exposed to the lumen, the part of the Fc receptor protein that is bound to the IgA dimer (*the secretory component*) is cleaved from its transmembrane tail, thereby releasing the antibody in the form shown in Figure 28–10.

tations of such allergic reactions as hay fever, asthma, and hives. Mast cells also secrete factors that attract and activate a special class of white blood cells called eosinophils, which can kill various types of parasites, especially if the parasites are coated with IgE or IgA antibodies. The properties of the various classes of antibodies in humans are summarized in Table 28–1.

Antibodies Can Have Either κ or λ Light Chains But Not Both

In addition to the five classes of heavy chains, higher vertebrates have two types of light chains, κ and λ, either of which may be associated with any of the heavy chains.[34] An individual-antibody molecule always consists of identical light chains and identical heavy chains; therefore, its two antigen-binding sites are always identical. This symmetry is crucial for the cross-linking function of secreted antibodies. Consequently, an Ig molecule may have either κ or λ light chains but not both. No difference in the biologic function of these two types of light chains has yet been identified.

Antibodies Recruit Complement to Help Fight Bacterial Infections

Complement, so called because it *complements* and amplifies the action of antibody, is one of the principal means by which antibodies defend vertebrates against most bacterial infections. Individuals with a deficiency in one of the central complement components (called C3) are subject to repeated bacterial infections, just as are individuals deficient in antibodies themselves. The complement system consists of at least 20 interacting soluble proteins that are made mainly by the liver and circulate in the blood and extracellular fluid.[35] Most are inactive until they are triggered by an immune response or, more directly, by an invading microorganism itself. The ultimate consequence of complement activation is the assembly of the so-called late *complement components* into large protein complexes, called *membrane attack complexes*, that form holes in the membrane of a microorganism and thereby destroy the microorganism.[36]

Because one of its main functions is to attack the membrane of microbial cells, the activation of

TABLE 28–1. Properties of the Major Classes of Antibody in Human

	Class of Antibody				
Properties	*IgM*	*IgD*	*IgG*	*IgA*	*IgE*
Heavy chains	μ	δ	γ	α	ε
Light chains	κ or λ	κ or λ	κ or λ	κ or λ	κ
Number of four-chain units	5	1	1	1 or 2	1
Percent of total Ig in blood	10	< 1	75	15	< 1
Activates complement	++++	–	++	–	–
Crosses placenta	–	–	+	–	–
Binds to macrophages and neutrophils	–	–	+	–	–
Binds to mast cells and basophils		–	–	–	+

Ig = immunoglobulin.

complement is focused on the microbial cell membrane, where it is triggered either by antibody bound to the microorganism or by microbial envelope polysaccharides, both of which activate the early *complement components*. There are two sets of early components belonging to two distinct pathways of complement activation: *the classic pathway* and the *alternative pathway*. The early components of both pathways act locally to activate *C3*, which is the pivotal component of complement, whose cleavage leads not only to the assembly of membrane attack complexes but also to the recruitment of various white blood cells, particularly neutrophils (Figure 28–12).

The early components and C3 are proenzymes that are activated sequentially by a limited proteolytic cleavage: the cleavage of each proenzyme in the sequence activates the component to generate a serine protease, which cleaves the next proenzyme in the sequence, and so on. Since each activated enzyme cleaves many molecules of the next proenzyme in the chain, the activation of the early components consists of an amplifying *proteolytic cascade*. Thus, each molecule activated at the beginning of the sequence leads to the production of many activated components, including many membrane attack complexes.

Many of these cleavages liberate a small peptide fragment and thereby expose a membrane-binding site on the large fragment, which binds tightly to the target cell membrane and helps to carry out the next reaction in the sequence, eventually leading to the formation of membrane attack complexes. In this way, complement activation is confined largely to the particular cell surface where

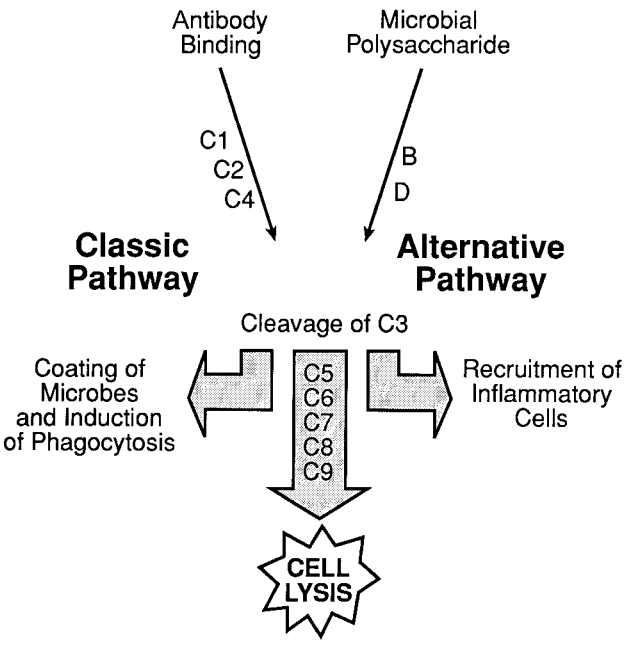


FIGURE 28–12. The principal stages in complement activation by the classic and alternative pathways. In both pathways, the reactions of complement activation usually take place on the surface of an invading microbe, such as a bacterium. C1–C9 and factors B and D are the reacting components of the complement system; various other components regulate the system. The early components are shown with *thin arrows*, whereas the late components are shown with *broad arrows*.

it began. The larger fragment of C3 is called C3b. It binds covalently to the surface of a target cell. There it not only acts as a protease to catalyze the subsequent steps in the complement cascade but also is recognized by specific receptor proteins on macrophages and neutrophils that enhance the ability of these cells to phagocytose the target cell. The smaller fragment of C3, C3a, acts independently as a diffusable signal that promotes an inflammatory response by encouraging white blood cells to migrate into the site of the infection.

The classic pathway is usually activated by clusters of IgG or IgM antibodies bound to antigens on the surface of a microorganism. The alternative pathway, in contrast, is activated by polysaccharides in the cell envelopes of the microorganisms even in the absence of antibodies. The alternative pathway therefore provides a first line of defense against infection before an immune response can be mounted and also amplifies the effects of the classic pathway once an immune response has begun.

The self-amplifying destructive properties of the complement cascade make it essential that key activated components be rapidly inactivated after they are generated to ensure that the attack does not spread to nearby normal host cells. Deactivation is achieved in at least two ways. First, specific inhibitor proteins in the blood terminate the cascade by either binding or cleaving certain components once they have been activated by proteolytic cleavage.[37] Second, many of the activated components in the cascade are unstable; unless they bind immediately to an appropriate component in the chain or nearby membrane, they rapidly become inactive. The assembly of the late complement components to form a membrane attack complex is illustrated in Figure 28–13.

In summary, a typical antibody molecule is a Y-shaped protein with two individual antigen-binding sites at the tips of the Y (the Fab regions), and binding sites for complement components and/or various cell-surface receptors are on the tail of the Y (the Fc region). Antibodies defend humans against infection by inactivating viruses and bacterial toxins and by recruiting the complement system in various cells to kill and ingest invading microorganisms. Multiple studies performed in our laboratory have shown that IgG is the principal Ig involved in defense against acute bacterial infections of the middle ear and sinuses.[38] On the other hand, IgA is the principal Ig covering the nasopharyngeal mucus and in preventing adherence of potentially pathogenic microorganisms.[39]

Each B-cell clone makes antibody molecules with a unique antigen-binding site. Initially, the molecules are inserted into the plasma membrane,

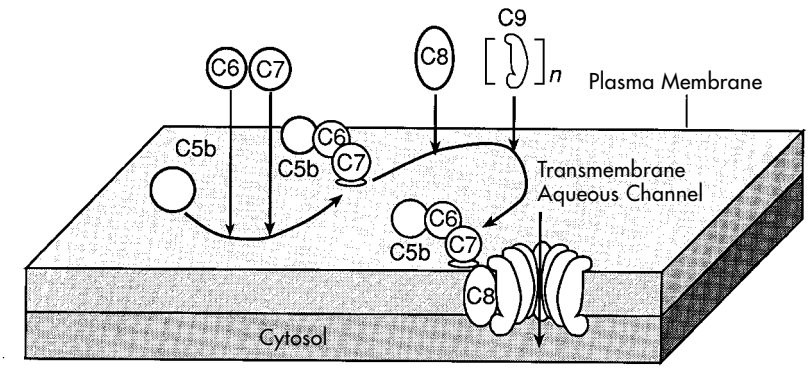


FIGURE 28–13. Assembly of the late complement components to form a membrane attack complex. When C3b is produced by either the classic or alternative pathway, it is mobilized on the membrane, where it causes the cleavage of a complement protein called C5 to produce C5a and C5b. C5b remains loosely bound to C3b (not shown) and rapidly assembles with C6 and C7 to form C5,6,7, which then binds firmly via C7 to the membrane, as illustrated. This complex adds one molecule of C8 to form C5,6,7,8. The binding of a molecule of C9 to C5,6,7,8 induces a confirmational change in the C9 that exposes a hydrophobic region and causes the C9 to insert into the lipid bilayer of the target cell next to C8. This starts a chain reaction in which the altered C9 binds a second molecule of C9, which undergoes a confirmational change and inserts into the bilayer, where it combines with another molecule of C9, and so on. In this way, a large transmembrane channel or pore is formed by a chain of C9 molecules.

where they serve as receptors for antigen. Antigen binding to these receptors activates the B cells (usually with the aid of helper T cells) to multiply and mature into either memory cells or antibody-secreting cells, which secrete antibodies with the same unique antigen-binding site as the membrane-bound antibodies. These B-cell clones are found in the tonsils and adenoids, and these two lymphoepithelial structures represent primarily B-cell tissues that are primarily responsible for a first line of defense against invading viruses and other microorganisms, possibly as well as food antigens.[40]

Each antibody molecule is composed of two identical heavy chains and two identical light chains. Typically, parts of both the heavy and light chains form the antigen-binding sites.[41] There are five classes of antibodies (IgA, IgD, IgE, IgG, and IgM), each with a distinctive heavy chain (α, δ, ε, γ, and μ, respectively). The heavy chains also form the Fc region of the antibody, which determines what other proteins will bind to the antibody and therefore what biologic properties the antibody class has. Either type of light chain (κ or λ) can be associated with any class of heavy chain, but the type of light chain does not seem to influence the properties of the antibody.

The complement system cooperates with antibodies to defend vertebrates against infection.[42] The early components are proenzymes that circulate in the blood and are sequentially activated in an amplifying series of limited proteolytic reactions. The most important complement component is the C3 protein, which is activated by proteolytic cleavage and binds to the membrane of the microbial cell, where it helps to initiate the local assembly of the late complement components and to induce the phagocytosis of the microbial cell. The late components form large membrane attack complexes in the microbial cell membrane and thereby kill the invading microorganism. Two final subjects on antibodies need to be discussed; the fine structure of antibodies and antibody diversity.

Light and Heavy Chains Each Contain Three Hypervariable Regions That Together Form the Antigen-Binding Site

Scrutiny of the amino acid sequences of a variety of Ig chains shows that the variability in the variable regions of both the light and heavy chains is, for the most part, restricted to three small hypervariable regions in each chain.[43] The remaining part of the variable region, known as the *framework regions*, is relatively constant. These findings led to the prediction that only 5 to 10 amino acids in each hypervariable region form an antigen-binding site. This prediction has since been confirmed by x-ray diffraction studies of antibody molecules. In agreement with the size of the antigen-binding site of an antibody molecule, the antigenic determinant that is specifically recognized by an antibody is generally comparatively small; it can consist of fewer than 25 amino acid residues on the surface of the globular protein and can be as small as a dinitrophenol group.[44]

Both light and heavy chains are made up of repeating segments, each about 110 amino acids and each containing one intrachain disulfide bind, that fold independently to form compact functional units or *domains*. As shown in Figure 28–14, a light

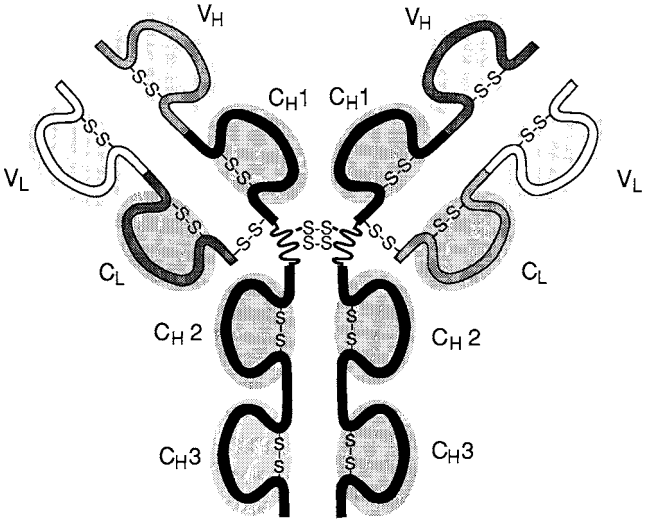


FIGURE 28–14. Immunoglobulin (Ig) domains. The light and heavy chains in an Ig molecule are each folded into repeating domains that are similar to one another. The variable domains of the light and heavy chains (V_L and V_H) make up the antigen-binding sites, whereas the constant domains of the heavy chains (mainly C_H2 and C_H3) determine the other biologic properties of the molecule. The heavy chains of IgM and IgE antibodies have an extra constant domain (C_H4). Hydrophobic interactions between domains on adjacent Ig chains play an important part in holding the chains together in the Ig molecule: C_L binds to C_H1, for example, and the C_H3 domains bind to each other.

chain consists of one variable (V_L) and one constant (C_L) domain, whereas most heavy chains consist of a variable domain (V_H) and three constant domains (C_H1, C_H2, C_H3).[45] The μ and ε chains each have one variable and four constant domains. The variable domains are responsible for antigen binding, whereas the constant domains of the heavy chains excluding C_H1 form the Fc region that determines the other biologic properties of the antibody. The similarity between their domains suggests that Ig chains arose during evolution by a series of gene duplications, beginning with the primordial gene cloning for a similar 110 amino acid domain of unknown function. This hypothesis is supported by the finding that each domain of the constant region of the heavy chain is encoded by a separate coding sequence (exon).[46] The amino acid sequence variation in the variable regions of both light and heavy chains is for the most part confined to several small hypervariable regions; they form protruding surface loops that come together to form the antigen-binding site.

Generation of Antibody Diversity

Finally, a brief comment of antibody diversity is necessary to complete our discussion of the B-cell repertoire and antibody function as part of the humoral immune system. It is estimated that even in the absence of antigen stimulation, a human makes at least 10^{15} different antibody molecules, its *preimmune antibody repertoire*.[47] Antibodies are proteins, and proteins are encoded by genes. Antibody diversity, therefore, poses a special genetic problem: how can an animal make more antibodies than there are genes in its genome? (The human genome, for example, is thought to contain fewer than 10^5 genes.) This problem is not quite as formidable as it might first appear. Because the variable regions of both the light and heavy chains contribute to an antigen-binding site, an animal with 1,000 genes encoding light chains and 1,000 genes encoding heavy chains could combine their products in 1,000 × 1,000 ways to make 10^6 different antigen-binding sites, assuming that any light chain can combine with any heavy chain to make an antigen-binding site. Nonetheless, the mammalian immune system has evolved unique genetic mechanisms that enable it to generate an almost unlimited number of different light and heavy chains in a remarkably economical way by joining separate *gene segments* together before they are transcribed.

Antibodies are produced from three pools of gene segments, encoding κ light chains, λ light chains, and heavy chains, respectively.[48] In each pool, separate gene segments that code for different parts of the variable regions of light and heavy chains are brought together by site-specific recombination during B-cell differentiation. The light chain pools contain one or more constant (C) gene segments and sets of variable (V) and joining (J) segments. The heavy chain pool contains a set of C gene segments and sets of V, diversity (D), and J gene segments. To make an antibody molecule, a V_L gene segment is recombined with a J_L gene segment to produce a DNA sequence coding for the V region of the light chain, and a V_H gene segment is combined with a D and a J_H gene segment to produce a DNA sequence coding for the V region for the heavy chain. Each of the assembled gene segments is then cotranscribed with the appropriate C-region sequence to produce a messenger ribonucleic acid molecule that codes for the complete polypeptide chain. By variously combining gene segments that code for V_L and V_H regions, mammals can make thousands of different light chains and thousands of different heavy chains. Since the antigen-binding site is formed where V_L and V_H come together in the final antibody, the heavy and light chains can pair to form antibodies with millions of different antigen-binding sites. This number is enormously increased by the loss and gain of nucleotides at the site of gene-segment joining, as well as by somatic mutations that occur with very high frequency in the assembled V-region coding sequences following antigen stimulation.

All B cells initially make IgM antibodies. Later, some make antibodies of other classes but with the same antigen-binding site as the original IgM antibodies. Such class switching allows the same antigen-binding sites to be distributed among antibodies with varied biologic properties. In regard to the tonsils and adenoids, such B class switching has been well documented. The mantle zone of the secondary follicles possesses primarily IgD and IgM. The follicular areas of the tonsils and adenoids possess IgM, IgA, and IgG. The reticular epithelium possesses mainly IgG followed by IgA, IgM, and IgD. Finally, only approximately 2% of the tonsillar lymphocytes are mature plasma cells that secrete Igs.[49]

T-CELL RECEPTORS AND SUBCLASSES

The diverse responses of T cells are collectively called *cell-mediated immune reactions.*[50] Like antibody responses, they are exquisitely antigen specific and are at least as important in defending vertebrates against infection. T cells differ from B cells, however, in several important ways.[51] First, they act only at short range, interacting directly with another cell in the body, which they either kill or signal in some way. These latter cells may be considered target cells such as cells infected with a virus or even tumor cells; B cells, by contrast, secrete antibodies that can act far away. Second, T cells are specialized to recognize foreign antigen only when it is displayed on the surface of a target cell. For this reason, the form of antigen recognized by T cells is different from that recognized by B cells: whereas B cells recognize intact antigens, T cells recognize peptide fragments of protein antigens that have been partially degraded inside the target cell and then carried to the cell surface and displayed there. In this way, T cells are able to detect the presence of microorganisms that proliferate inside cells, as well as foreign extracellular antigens that cells have ingested. There are two main classes of T cells: cytotoxic T cells and helper T cells.[52] *Cytotoxic T cells* directly kill cells that are infected with a virus or some other intracellular microorganism. *Helper T cells*, in contrast, help stimulate the response of other cells: they help activate macrophages and B cells, for example.

T-Cell Receptors Are Antibody-Like Heterodimers

Because T-cell responses depend on direct contact with the target cell, the antigen receptors made by T cells, unlike antibodies made by B cells, exist only in membrane-bound form and are not secreted. On both cytotoxic and helper T cells, receptors are composed of two disulfide-linked polypeptide chains called α and β, each of which contains two Ig-like domains and shares with antibodies the distinctive property of a variable amino-terminal region and a constant carboxyl-terminal region (Figure 28–15).[53]

The gene pools that encode the α and β chains are located on different chromosomes and contain, like antibody gene pools, separate V, D, J, and C gene segments, which are brought together by site-specific recombination during T-cell development in the thymus.

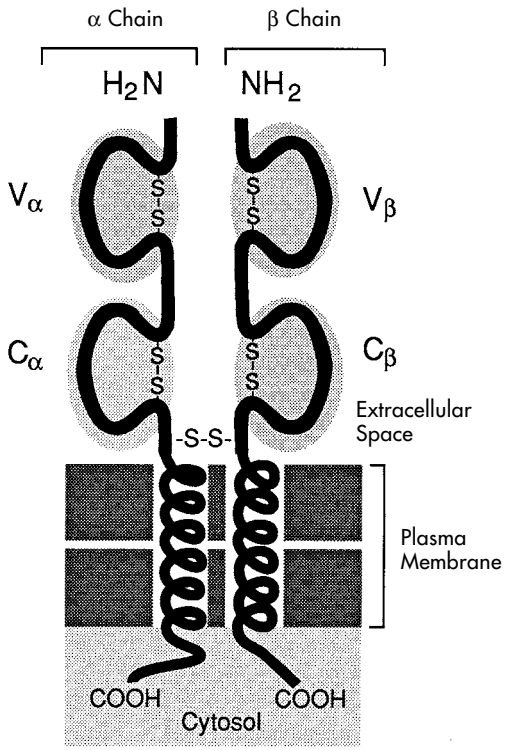


FIGURE 28–15. A T-cell receptor heterodimer. The receptor is composed of an α and a β polypeptide chain. Each chain is about 280 amino acids long and has a large extracellular part that is folded into two immunoglobulin-like domains, one variable (V) and one constant (C). It is thought that an antigen-binding site formed by a V_α and a V_β domain is similar in its overall dimensions in geometry to the antigen-binding site of an antibody molecule. Unlike antibodies, however, which have two binding sites for antigen, T cell receptors have only one. The α/β heterodimer shown is noncovalently associated with the large set of invariant proteins that help activate the T cell when the T-cell receptors bind to antigen. A typical T cell has about 20,000 such receptor complexes on its surface.

A small minority of T cells, instead of making α and β chains, make a different type of receptor heterodimer, composed of γ and δ chains.[54] These cells arise early in development and are found mainly in epithelia (eg, in the skin and gut) and also in the nasal mucosa. As is the case for antigen receptors on B cells, the T-cell receptors are tightly associated in the plasma membrane with a number of invariant proteins that are involved in passing a signal from an antigen-activated receptor to the cell interior.

Different T-Cell Responses Are Mediated by Distinct Classes of T Cells

The two major classes of T cells have very different functions. *Cytotoxic T cells* kill cells harboring harmful microbes, whereas *helper T cells* help activate the responses of other white blood cells, mainly by secreting a variety of local mediators collectively called lymphokines, ILs, or cytokines. These proteins are discussed later in this chapter. Thus, cytotoxic T cells provide protection against pathogenic microorganisms, such as viruses and some intracellular bacteria, which multiply in the host cytoplasm, where they are sheltered from attack by antibodies. The most efficient way of preventing such microorganisms from spreading to other cells is to kill the infected cell before the microorganism can proliferate. Helper T cells, in contrast, are crucial for stimulating responses to extracellular microorganisms and their toxic products. There are at least two types of helper T cells: *T_H1 cells*, which activate macrophages to destroy microorganisms that they have ingested, and *T_H2 cells*, which stimulate B cells to proliferate and secrete antibodies, particularly of the IgE and IgG classes.[55]

Both cytotoxic T cells and helper T cells recognize antigen in the form of peptide fragments that are generated by degradation of foreign protein antigens inside the target cell; therefore, both depend on the presence in the target cell of special proteins that bind these fragments, carry them to the cell surface, and present them there to the T cells. These special proteins are called *MHC molecules* because they are encoded by a complex of genes called the major histocompatibility complex (MHC).[56] There are two structurally and functionally distinct classes of MHC molecules: *class I MHC molecules*, which present foreign peptides to cytotoxic cells, and *class II MHC molecules*, which present foreign peptides to helper cells.[57] A close look at the MHC molecules themselves must now be undertaken to understand T-cell immunity.

Major Histocompatibility Complex Molecules and Antigen Presentation to T Cells

Major histocompatibility complex molecules were recognized long before their normal function was understood. They were initially defined as the main target antigens in *transplantation reactions*. When organ grafts are exchanged between adult individuals, either of the same species (allografts) or of different species (xenografts), they are usually rejected. This graft rejection is an immune response to antigens on the surface of the graft itself. It has been shown that these reactions are mediated mainly by T cells and that they are directed against "genetically foreign" versions of cell-surface proteins called *histocompatibility molecules*. The MHC family of proteins encoded by the cluster genes of the MHC is by far the most important of these. Major histocompatibility complex molecules are expressed on the cells of all higher vertebrates. They were first demonstrated in mice and were called *H-2 antigens*. In humans, they are called HLA antigens (human leukocyte–associated antigens) because they were first demonstrated on leukocytes (white blood cells). Three remarkable properties of MHC molecules baffled immunologists for a long time. First, MHC molecules are overwhelmingly the preferred target antigens for T-cell-mediated transplantation reactions. Second, an unusually large fraction of T cells is able to recognize foreign antigen MHC molecules: whereas fewer than 0.001% of an individual's T cells respond to a typical viral antigen, more than 0.1% of them respond to a single foreign MHC antigen. Third, many of the loci that code for MHC molecules are the most *polymorphic* known in higher vertebrates, that is, within a species, there is an extraordinarily large number of alleles (alternate forms of the same genes)[58] at each locus (in some cases as many as 100), each allele being present at a relatively high frequency in the population. For this reason, and because each individual has five or more loci encoding MHC molecules, it is very rare for two individuals to have an identical set of MHC proteins, making it very difficult to match donor and recipient for organ transplantation in humans (except in the case of genetically identical twins). Why is there this apparent obsession of one's own T cells to recognize foreign MHC molecules when a vertebrate does not usually need to be protected against invasion by foreign vertebrate cells? This puzzle has been solved since it has been discovered that the MHC molecules serve to focus T cells on those host cells that have foreign antigen on their surface and that the T cells respond to foreign MHC molecules in the same way as to self-MHC molecules that have foreign antigen bound to them.[59]

Thus, in summary, foreign protein molecules are presented to the T-cell receptor only in the context of their association with self-MHC molecules. In a sense, this is a form of physiologic autoimmunity in that a foreign molecule can be recognized by a T cell only when it is combined with a self-MHC molecule.

There Are Two Principal Classes of MHC Molecules

Class I and class II MHC proteins have very similar overall structures, as shown in Figure 28–16. They are both transmembrane heterodimers whose extracellular amino-terminal domains bind antigen for presentation to T cells. Each *class I MHC gene* encodes a single transmembrane polypeptide chain (called α), most of which is folded into three extracellular globular domains (α_1, α_2, α_3). Each α chain is noncovalently associated with a small extracellular protein called β_2-*microglobulin,* which does not span the membrane and is encoded by a gene that does not lie in the MHC cluster.[60] β_2-Microglobulin and the α_3 domain, which are closest to the membrane, are both homologous to an Ig domain. The two amino-terminal domains of the α chain, which are farthest from the membrane, bind antigen and contain the polymorphic (variable) amino acids that are recognized by T cells in transplantation reactions. Like class I MHC molecules, class II MHC molecules are heterodimers with two conserved Ig-like domains close to the membrane and two antigen-binding polymorphic (variable) amino-terminal domains farthest from the membrane. In these molecules, however, both chains (α and β) are encoded within the MHC, and both span the membrane. The presence of Ig-like domains in class I and class II proteins suggests that MHC molecules and antibodies have a common evolutionary history. The locations of the genes that encode class I and class II MHC proteins in mice and humans are shown in Figure 28–17.[61]

Class I and Class II MHC Molecules Have Different Functions

Class I MHC molecules are expressed on virtually all nucleated cells, presumably because cytotoxic T cells must be able to focus on any cell in the body that happens to become infected with an intracellular

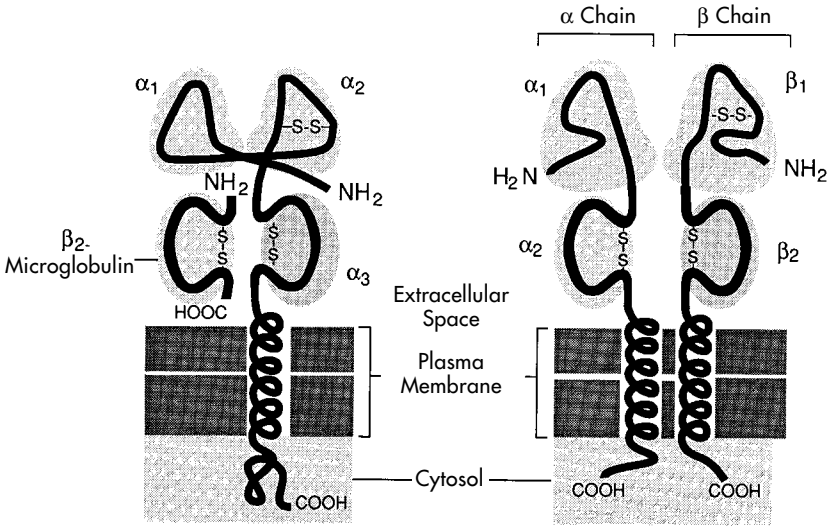


FIGURE 28–16. Class I and class II major histocompatibility complex (MHC) proteins. On the left, the α chain of the class I molecule has three extracellular domains, α_1, α_2, and α_3, encoded by separate exons. It is noncovalently associated with a smaller polypeptide chain, β_2-microglobulin, which is not encoded within the MHC. The α_3 domain and β_2-microglobulin are immunoglobulin (Ig)-like. Whereas β_2-microglobulin is invariant, the α chain is extremely polymorphic, mainly in the α_1 and α_2 domains. On the right in class II MHC molecules, both chains are polymorphic (β more than α), mainly in the α_1 and β_1 domains; the α_2 and β_2 domains are Ig-like. Thus, there are striking similarities between class I and class II MHC proteins. In both, the two outermost domains interact to form a groove that binds foreign antigen and presents it to T cells. All of the chains are glycosylated except for β_2-microglobulin.

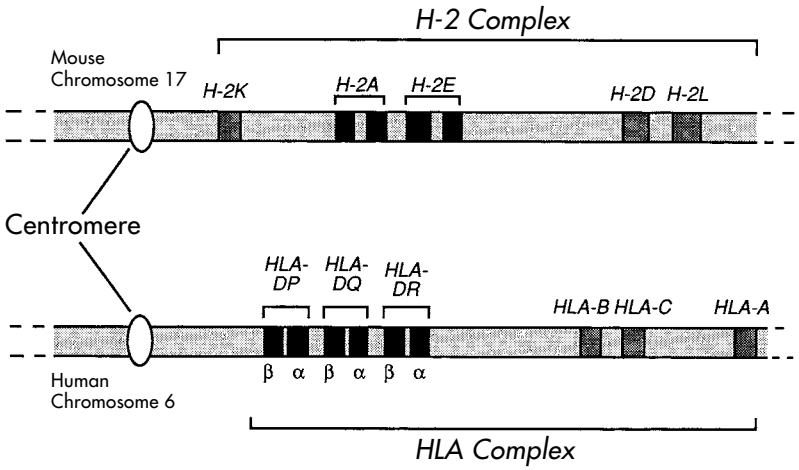


FIGURE 28–17. The H-2 and human leukocyte antigen (HLA) gene complexes. This simplified schematic drawing shows the location of the genetic loci that encode the transmembrane subunits of *class I* and *class II* major histocompatibility complex (MHC) proteins. There are three types of class I proteins (H-2K, H-2D, and H-2L in the mouse, and HLA-A, HLA-B, and HLA-C in the human). There are two class II MHC loci in the mouse, H-2A and H-2E, and more than three in humans, of which only three are shown, HLA-DP, HLA-DQ, and HLA-DR. Each class II locus encodes at least one α chain and at least one β chain, but some encode more than one α or β chain.

microbe such as a virus. Class II molecules, in contrast, are normally confined to specialized cells, such as B cells, macrophages, and other antigen-presenting cells that take up foreign antigens from the extracellular fluid and interact with helper T cells (Figure 28–18). The principal features of the two classes of MHC proteins are summarized in Table 28–2.

It is important that cytotoxic T cells focus their attack on cells that *make* foreign antigens, whereas helper T cells focus their help on cells that take up foreign antigens from the extracellular fluid. The former type of target cell is a menace, but the latter type is essential for the body's immune defenses. The immune system must be able to dispose of extracellular foreign antigens as many bacteria multiply outside cells, and some secrete protein toxins, such as tetanus toxin and botulinum toxin, that can be lethal. Helper T cells help eliminate such pathogens, both by helping B cells make antibodies against microbes and their toxins and by activating macrophages to destroy ingested microbes.

To ensure that there is no misdirection of cytotoxic and helper functions, each of the two major classes of T cells, in addition to the antigen receptor that recognizes a peptide-MHC complex, also expresses a *coreceptor* that recognizes a nonpolymorphic part of the appropriate class of MHC molecule. *CD4 and CD8 proteins act as MHC-binding coreceptors on helper and cytotoxic T cells, respectively.*[62] The affinity of T-cell receptors for peptide-MHC complexes on the target cell is usually too low to mediate a functional interaction between the two cells. *Accessory receptors* are normally required to help stabilize the interaction by increasing the overall strength of the cell-to-cell adhesion; when they also have a direct role in activating the T cell by generating their own intercellular signals, they are called *coreceptors*. Unlike T-cell receptors or MHC molecules, the accessory receptors do not bind antigen and are invariant and nonpolymorphic. The most important and best understood of the coreceptors on T cells are the CD4 and CD8 proteins, both of which are single-pass transmembrane proteins with extracellular Ig-like domains. Like T-cell receptors, they recognize MHC proteins, but unlike T-cell receptors, they bind to nonvariable parts of the protein, far away from the peptide-binding groove. *CD4* is expressed on helper T cells and binds to class II MHC molecules, whereas CD8 is expressed on cytotoxic T cells and binds to class I MHC molecules (Figure 28–19). The CD4 and CD8 proteins are not only required to increase the strength of cell-to-cell

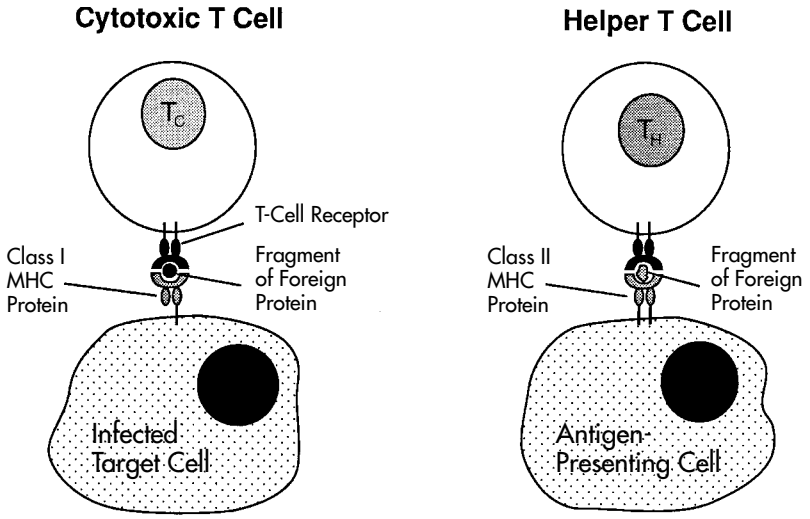


FIGURE 28–18. Cytotoxic and helper T cells recognize different major histocompatibility complex (MHC) molecules. Cytotoxic T cells recognize foreign antigens in association with class I MHC proteins on the surface of any infected host cell, whereas helper T cells recognize foreign antigens in association with class II MHC proteins on the surface of an antigen-presenting cell, such as a macrophage or a B cell. The foreign antigen bound to a class I MHC molecule is synthesized within the target cell, whereas the foreign antigen bound to a class II MHC molecule has been taken up by the cell by endocytosis and processed before it is presented on the cell surface. In transplantation reactions as well, helper T cells react against foreign class II MHC proteins, whereas cytotoxic T cells react against foreign class I MHC proteins.

adhesion and to help activate the T cell, they are also needed for T-cell development; if the genes that encode CD4 and CD8 are inactivated by targeted genetic recombination, then either helper T cells or cytotoxic T cells, respectively, do not develop. Ironically, CD4 also functions as a receptor for the acquired immune deficiency syndrome (AIDS) human immunodeficiency virus, allowing the virus to infect helper T cells.

CYTOTOXIC T CELLS

Cytotoxic T cells recognize fragments of viral proteins on the surface of virus-infected cells. As men-

TABLE 28–2. Properties of Class I and Class II MHC Molecules

	Class I	*Class II*
Genetic loci	H-2K, H-2D, H-2L in mice; HLA-A, HLA-B, HLA-C in humans	H-2A and H-2E clusters in mice; DP, DQ, DR, and several others in humans
Chain structure	α chain +β_2-microglobulin	α chain + β chain
Cell distribution	Most nucleated cells	Antigen-presenting cells (including B cells), thymus epithelial cells, some others
Involved in presenting antigen	Cytotoxic T cells	Helper T cells
Source of peptide fragments	Proteins made in cytosol	Endocytosed plasma membrane and extracellular proteins
Polymorphic domains	$\alpha_1 + \alpha_2$	$\alpha_1 + \beta_1$

MHC = major histocompatibility complex; HLA = human leukocyte antigen.

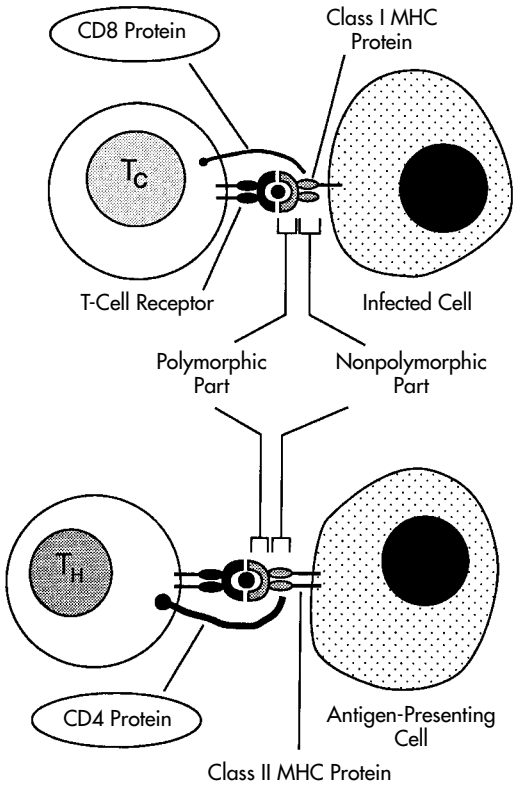


FIGURE 28–19. CD4 and CD8 coreceptors on the surface of T cells. Note that these proteins bind to the nonvariable part of the same major histocompability complex (MHC) molecule that the T-cell receptor has engaged, so that they are bought together with T-cell receptors during the cell activation process. Antibodies against CD4 and CD8 are widely used to distinguish helper and cytotoxic T cells, respectively.

tioned above, T cells of any individual that recognize a specific antigen will do so only when that antigen is associated with the allelic forms of MHC molecules expressed by that individual, a phenomenon known as *MHC restriction*. It has been shown that cytotoxic T cells, when activated by, for example, influenza virus, specifically recognize internal proteins of that virus that would not be accessible in the intact virus particle. Subsequent evidence has suggested that the T cells have recognized degraded fragments of the internal viral proteins. Because a T cell can recognize tiny amounts of antigen (only a few hundred molecules), only a small fraction of the fragments generated from viral proteins has to get to the cell surface to attract an attack by a cytotoxic T cell.[63] Proteins destined for the cell surface usually begin their journey by crossing from the cytosol into the lumen of the endoplasmic reticulum (ER). Once inside ER, the peptides bind to MHC molecules that they encounter there and are then carried to the cell surface with them. When T cells are activated by antigen, they secrete various signaling protein molecules, including *interferon-gamma (IFN-γ)*, which greatly enhances antiviral responses.[64] The IFN-γ induces the expression of many genes within the MHC chromosomal region, including those in class I and class II MHC proteins, the two specialized proteasome subunits, and the two subunits of the peptide pump located in the ER. Thus, all of the machinery required for presenting viral antigens to cytotoxic T cells is coordinately called into action by IFN-γ, creating a positive feedback that amplifies the immune response and culminates in the death of the infected cell.

Once a cytotoxic T cell recognizes a viral peptide bound to a class I MHC molecule on the surface of a target cell, its job is to destroy the cell before the virus that gave rise to the peptide can produce new viral particles that can escape from the infected cell. The mechanism by which cytotoxic T cells kill their targets is not known. They seem to employ at least two strategies, both of which are thought to operate by inducing the target cell to undergo *programmed cell death* (also called *apoptosis*).[65] In one strategy, binding to a target cell stimulates the cytotoxic T cells to release a pore-forming protein called *perforin*, which is homologous to the complement component C9 and polymerizes in the target cell plasma membrane to form transmembrane channels.[66] Perforin is stored in secretory vesicles and is released by local exocytosis at the point of contact with the target cell. The secretory vesicle also contains serine proteases and other proteins, which are also thought to play a part in killing the target cell, perhaps by entering the target cell through the perforin channels and inducing programmed cell death. The second strategy, by contrast, involves the cytotoxic T cell activating a receptor on the surface of the target cell, thereby signaling the target cell to undergo programmed cell death.[67]

HELPER T CELLS AND T-CELL ACTIVATION

Unlike cytotoxic T cells, *helper T cells* do not act directly to kill infected cells or to eliminate microorganisms. Instead, they stimulate macrophages to be more effective in destroying pathogens, and, most importantly, they help other types of lymphocytes to

respond to antigen. Helper T cells recognize fragments of endocytosed foreign protein antigens in association with class II MHC proteins.

Like the viral antigens presented to cytotoxic T cells, the antigen presented to helper T cells on antigen-presenting cells is degraded fragments of foreign protein that are bound to class II MHC molecules in much the same way that virus-derived peptides are bound to class I MHC molecules. But both the source of the peptide fragments presented and the route they take to find the MHC molecules are different from those of peptide fragments presented by the class I MHC molecules to cytotoxic T cells. Rather than being derived from foreign proteins synthesized inside the target cell, the peptides presented to helper T cells are derived from extracellular microbes or their products, which have been ingested by antigen-presenting cells and degraded in the acidic environment of endosomes.[68] These peptides do not have to be pumped across the membrane because they do not originate in the cytosol; they are generated in the compartment that is topographically equivalent to the extracellular space. They never enter the lumen of the ER, where the class II MHC molecules are synthesized and assembled, but instead bind to preassembled class II heterodimers in a late endosomal compartment. Once the peptide binds, the class II MHC molecule alters its confirmation, trapping the peptide in the binding groove for presentation at the cell surface to helper T cells.

Most of the class I and class II MHC molecules on the surface of the target cell have peptides derived from self-proteins in their binding groove: for class I molecules, fragments of degraded cytosolic and nuclear proteins; for class II molecules, fragments of degraded membrane and serum proteins that pass through the endosome-lysosome system. Only a small fraction of the class II MHC molecules on the surface of an antigen-presenting cell will have foreign peptides bound to them. This, however, suffices to initiate an immune response because only a few hundred such molecules are required to activate a helper T cell, just as only a few hundred peptide–class I MHC complexes on a target cell are required to activate a cytotoxic T cell.

Just as a B cell must be activated to proliferate and differentiate into an antibody-secreting cell before it can function, so, too, must a T cell be activated to proliferate and differentiate before it can kill an infected target cell or help a macrophage or B cell. The initial activation of a T cell usually occurs when it recognizes foreign peptide bound to an MHC molecule on the surface of an antigen target cell. For a helper T cell, the appropriate target is an *antigen-presenting cell*.

Antigen-presenting cells are derived from the bone marrow and comprise a heterogeneous set of cells, including *interdigitating dendritic cells* in lymphoid organs and *Langerhans' cells in skin*, as well as the B cells and macrophages that will subsequently be the target of T-cell help. Together with thymus epithelial cells, which have a special role in T-cell development and activate T cells in some mammals, these specialized antigen-presenting cells are the only cell types that normally express class II MHC molecules. In addition to class II MHC molecules, antigen-presenting cells also express a second cell-surface molecule, called B7, that plays a crucial part in activating T cells.

Two Simultaneous Signals Are Required for Helper T-Cell Activation

To activate a helper T cell, an antigen-presenting cell must provide at least two signals. *Signal 1* has already been discussed; it is provided by a foreign peptide bound to a class II MHC molecule on the surface of the presenting cell that activates the T-cell receptor complex. Depending on the type of helper T cell, a *signal 2* is provided by either a secreted chemical signal such as IL-1, or by the plasma membrane–bound signaling molecule *B7* on the surface of the antigen-presenting cell. B7 is recognized by a coreceptor protein called CD28, which is present on the surface of the helper T cell and is a member of the Ig superfamily.[69] If helper T cells receive both signals, they are activated to proliferate and secrete a variety of cytokines. In contrast, if they receive signal 1 without signal 2, they are altered so that they can no longer be activated even when they receive both signals. This has been suggested to be one mechanism whereby T cells become *tolerant*. Some accessory proteins present on the surface of T cells are summarized in Table 28–3.

Helper T Cells, Once Activated, Stimulate Themselves and Other T Cells to Proliferate by Secreting Interleukin-2

The combined action of signal 1 and signal 2 provokes helper T-cell proliferation by a curiously indi-

TABLE 28–3. Some Accessory Proteins on the Surface of T Cells

Protein	*Approximate Molecular Weight*	*Super-family*	*Expressed on*	*Ligand on Target Cell*	*Functions*
CD3	γ chain = 25,000 δ chain = 20,000 ε chain = 20,000 ζ chain = 16,000	Ig Ig Ig —	All T cells	—	Helps transduce signal when antigen-MHC complex binds to T-cell receptors
CD4	55,000	Ig	Helper T cells	Class II MHC	Promotes adhesion to antigen-presenting cells and to B cells; signals T cell
CD8	70,000 (homodimer or heterodimer)	Ig	Many helper and cytotoxic T cells	Class I MHC	Promotes adhesion to infected target cells; signals T cell
CD28	80,000 (homodimer)	Ig	Cytotoxic T cells	B7	Provides signal 2 to some T cells
LFA-1	a_1 chain = 190,000 b_2 chain = 95,000	Integrin	Most white blood cells, including T cells	ICAM-1	Promotes cell-to-cell adhesion

Ig = immunoglobulin; MHC = major histocompatibility complex; ICAM = intercellular adhesion molecule.

rect mechanism. It causes the T cells to stimulate their own proliferation by simultaneously secreting a growth factor called *IL-2* and the synthesized cell-surface receptors that bind it. The binding of IL-2 to these IL-2 receptors then directly stimulates the T cells to proliferate.[70] By this *autocrine mechanism*, helper T cells can continue to proliferate after they have left the surface of the antigen-presenting cell (Figure 28–20). The helper T cell can also stimulate the proliferation of any other nearby T cells, including cytotoxic T cells, which have first been induced by antigen to express IL-2 receptors. Because the expression of IL-2 receptors is strictly dependent on antigen stimulation, however, IL-2 causes the proliferation of only T cells that have encountered that specific antigen.

Helper T Cells Are Required for Most B Cells to Respond to Antigen

Although there are some antigens, including microbial polysaccharides, that can stimulate B cells to proliferate and differentiate into antibody-secreting cells without T-cell help, the activation of B cells by helper T cells is mediated by both membrane-bound and secreted signals. Once activated, the helper T cell can help activate a B cell that specifically displays the same complex of foreign antigen and class II MHC protein on its surface. The display of antigen on the B-cell surface reflects the selectivity with which it takes up foreign molecules from the extracellular fluid. The specific membrane-bound antibodies on the surface of the B cell are ingested and then degraded and recycled to the cell surface in the form of peptides bound to class II MHC proteins. Thus, the helper cell activates those B cells that make membrane-bound antibodies that specifically recognize the antigen that initially activated the T cell. In secondary antibody responses, memory B cells themselves may act as antigen-presenting cells and activate helper T cells, as well as be the subsequent targets of the helper T cells. The mutually reinforcing actions of T and B cells lead to an immune response that is both intense and highly selective. The membrane-bound signaling molecule is a transmembrane protein called *CD40 ligand*, which is expressed on the surface of activated, but not resting, helper T cells. It is recognized by the *CD40* transmembrane protein on the B cell.[71] The interaction between CD40 ligand and CD40 is required for helper T cells to activate B cells

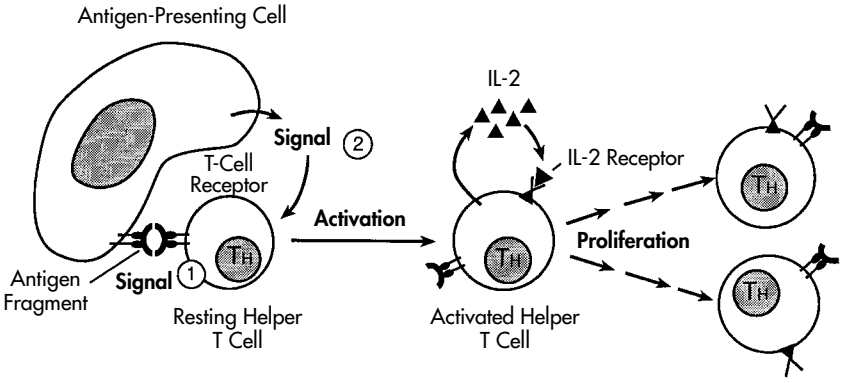


FIGURE 28–20. The stimulation of T-cell proliferation by interleukin (IL)-2. Signals 1 and 2 activate the helper T cells to make IL-2 receptors and to secrete IL-2. The binding of IL-2 to its receptors stimulates the cell to grow and divide. When the antigen is eliminated, the T cells eventually stop producing IL-2 and IL-2 receptors, so cell proliferation stops. Some helper T cells do not make IL-2; their proliferation, like that of cytotoxic cells, is stimulated by IL-2 made by neighboring helper T cells (signals 1 and 2 are discussed in the text).

to proliferate and mature into memory and antibody-secreting cells. This interaction is critical for T-cell help: individuals whose T cells lack the CD40 ligand because of the mutation in the gene encoding the protein can make only IgM antibodies and are severely immune deficient, being susceptible to the same infections that affect AIDS patients, whose helper T cells have been destroyed.

Secreted signals from the T cells provide additional help by activating B cells to proliferate, mature, and, in some cases, switch the class antibody they produce. *Interleukin-4* is one such signal.[72] It stimulates B-cell proliferation and maturation and promotes the switching to IgE and IgG1 antibody production: if the IL-4 gene is inactivated by targeted recombination in a mouse, the mouse is unable to make IgE and makes very little IgG1.

Thus, most B cells, like T cells, require multiple signals for activation, one provided by antigen binding to the membrane-bound Ig molecules and the others provided by the helper T cells; as in the case of T cells, if a B cell receives the first signal only, it may be functionally inactivated. The comparison of the signals required to activate helper T cell and a B cell are outlined schematically in Figure 28–21.

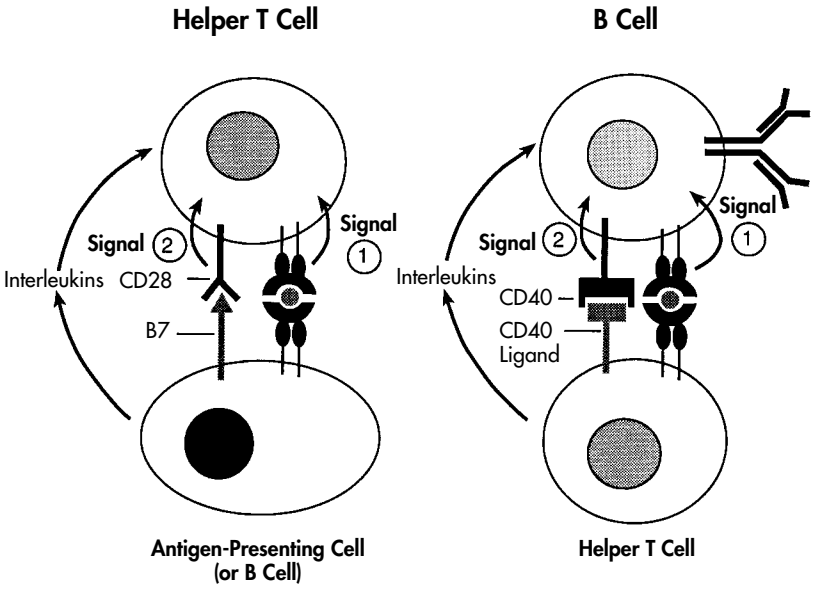


FIGURE 28–21. Comparison of the signals required to activate a helper T cell and B cell. Note that signal 2 can be provided either by a secreted signaling molecule (an interleukin) or by a cell-to-cell contact interaction.

Some Helper T Cells Help Activate Cytotoxic T Cells and Macrophages by Secreting Interleukins

There are at least two functionally distinct subclasses of helper T cells that can be distinguished by the cytokines that they secrete. T_H1 cells secrete IL-2 and IFN-γ and are concerned mainly with helping cytotoxic T cells and macrophages.[73] T_H2 cells secrete IL-4 and IL-5 and are concerned mainly with helping B cells and eosinophils.[74] The properties of some ILs are outlined in Table 28–4. The antigen-triggered secretion of ILs underlies the familiar tuberculin skin test. If tuberculin is injected into the skin of an individual who had been immunized against tuberculosis, a characteristic immune response occurs in the skin. It is initiated at the site of injection by the secretion of ILs by memory helper T cells that react to the tuberculin. The ILs attract macrophages and lymphocytes into the site, thereby causing the characteristic swelling of a positive reaction to tuberculin.

Another important effect of IFN-γ is to induce the expression of class II MHC proteins on the surface of some cells such as endothelial cells that do not normally express them, thereby enabling these cells to present antigen to helper T cells.[75]

In summary, helper T cells help activate B cells and macrophages. They are themselves initially activated when they recognize peptide fragments derived from foreign extracellular proteins that are endocytosed by specialized antigen-presenting cells. The ingested proteins are degraded in endosomes, and some of the resulting peptide fragments bind to class II MHC molecules, forming complexes that are carried to the cell surface, where the helper T cells recognize them. The activation of the helper T cell requires at least two signals: signal 1 is provided by the MHC-peptide complex, whereas signal 2 is provided by either the B7 protein on the surface of an antigen-presenting cell or a signal secreted by this cell. Once activated, helper T cells stimulate their own proliferation by secreting IL-2 and activate their target cells by a combination of membrane-bound and secreting signaling molecules. Developing T cells that recognize peptides in association with self-MHC molecules are positively selected in the thymus, whereas developing T cells that react strongly with self-peptides bound to self-MHC molecules are eliminated in the thymus. Thus, those immature T cells that will be capable of recognizing foreign peptides presented by self-MHC molecules are selected to survive, whereas the remainder, which would be

TABLE 28–4. Properties of Some Interleukins

Cytokines (IL)	*Approximate Molecular Weight*	*Source*	*Target*	*Action*
IL-1	15,000	Antigen-presenting cells	Helper T cells	Helps activate
IL-2	15,000	Some helper T cells	All activated T cells and B cells	Stimulates proliferation
IL-3	25,000	Some helper T cells	Various hemopoietic cells	Stimulates proliferation
IL-4	20,000	Some helper T cells	B cells	Stimulates proliferation, maturation, and class switching to IgE and IgG1
IL-5	20,000	Some helper T cells that make IL-4	B cells, eosinophils	Promotes proliferation and maturation
IL-6	25,000	Some helper T cells and macrophages	Activated B cells, T cells	Promotes B-cell maturation to Ig-secreting cells; helps activate T cells
γ-Interferon	25,000 (dimer)	Some helper T cells that make IL-2	B cells, macrophages, endothelial cells	Activates various MHC genes and macrophages

Ig = immunoglobulin; MHC = major histocompatibility complex.

of no use to the animal, die. Thus, MHC restriction is a required property of the immune system that emerges as T cells develop in the thymus. Positive selection still leaves a large problem to be solved. If developing T cells with receptors that recognize self-peptides associated with self-MHC molecules were to mature in the thymus and migrate to peripheral lymphoid tissues, they would wreak havoc. A second, *negative selection* process in the thymus is required to avoid this disaster. A fundamental feature of the immune system is that it can distinguish self from nonself and normally does not react against self-molecules. This state of *immunologic self-tolerance* is acquired mainly during T-cell development. It is not enough, therefore, for the thymus to select *for* T cells that recognize self-MHC molecules; it must also select *against* T cells that recognize self-MHC molecules complex with self-peptides. In other words, it must pick out for survival just those T cells that would be capable of recognizing self-MHC molecules complexed with foreign peptides, even though these peptides are not present in the developing thymus. Thus, the required goal can be achieved by (1) ensuring the death of T cells that bind strongly to the peptide-MHC complexes in the thymus while (2) promoting the survival of those that bind weakly and (3) permitting the death of those that do not bind at all. Process 2 is the positive selection that has been discussed above. Process 1 is called *negative selection*.

In summary, the T-cell repertoire is shaped mainly by a combination of positive and negative selection processes that operate during T-cell development in the thymus. These processes ensure that only T cells with potentially useful receptors survive and mature, whereas the others undergo programmed cell death: T cells that will be able to recognize foreign peptide complex with self-MHC molecules are positively selected, whereas T cells that react strongly with self-peptide complex with self-MHC molecules are eliminated.

THE IMMUNE SYSTEM AND THE INFLAMMATORY RESPONSE

The final part of this chapter briefly reviews certain surface molecules that are up-regulated in inflammatory tissues with specific emphasis on nasal polyposis as a model of chronic inflammation. A knowledge of these inflammatory cytokines as well as integrins that are heterodimeric proteins that promote cell to cell-surface interaction and counter-receptors on the surface of endothelial cells of blood vessels is emphasized because the cell-to-cell interaction is responsible for the migration of inflammatory cells into the tissue following an immune response.[76] A great deal of molecular biologic information has accumulated on the pathogenesis of nasal polyposis. It appears that eosinophils and lymphocytes are the predominant inflammatory cell in nasal polyp tissue. The events leading up to the extravasation of eosinophils into the lamina propria of nasal polyps are regulated by the proinflammatory cytokines, tumor necrosis factor alpha (TNF-α) and IL-1β.[77] These cytokines up-regulate VLA-4 on the surface of eosinophils and VCAM-1 on the surface of the endothelial cells of blood vessels.[78] Chemokines such as RANTES (regulated upon activation, normal T cell expressed and secreted) and eotaxin are responsible for the movement of eosinophils into the lamina propria of the nasal polyp.[79] The release of major basic protein from eosinophils has an effect on alteration of the epithelial architecture and on the sodium and chloride flux into and out of the apical epithelial cell of the tissue.[80] Therefore, it is incumbent on the student of chronic inflammation to understand how the immune process, which has been thoroughly discussed in this chapter, can effect the inflammatory response. Therefore, to conclude this chapter on the immune response and its effect on inflammation, a brief review of cytokines, integrins, and endothelial cell receptors is undertaken.

CYTOKINES

Cytokines are a group of signaling molecules involved in communication between cells, including those of the immune system.[81] Cytokine-mediated events occur during the initiation and effector stages of immune responses and the development of hematopoietic cells. Within the last decade, there has been an explosion of knowledge of cytokine and cytokine receptor structures and genes along with a clarification of the roles of individual cytokines. This section focuses on two proinflammatory cytokines, namely TNF-α and IL-1β. However, there are over 20 cytokines that have been identified, as well as their receptors.[82] A schematic overview of the role of various cytokines on T_H1- and T_H2-dependent

immune reactions is summarized in Figure 28–22. Tumor necrosis factor α can be derived from macrophages, T cells, thymocytes, B cells, and natural killer cells. The target of TNF-α may be tumor cells, transformed cell lines, fibroblasts, macrophages, osteoclasts, neutrophils, adipocytes, eosinophils, endothelial cells, chondrocytes, and hepatocytes. Tumor necrosis factor α is known to be made by activated macrophage and other cells and to have wide-ranging proinflammatory effects. It is purified from natural sources and has a molecular weight of 17 kDa. It has a vast array of biologic activities that include promoting bone resorption by osteoclasts, stimulation of T-cell production, enhancement of B-cell proliferation and Ig secretion, the activation of vascular endothelial cells, increasing eosinophilic toxicity for parasites, and the stimulation and proliferation of fibroblasts, among other activities. It is apparent that many of the biologic activities of TNF-α are important in both immunity and inflammation. They act in a very wide range of cellular targets and cause enormous

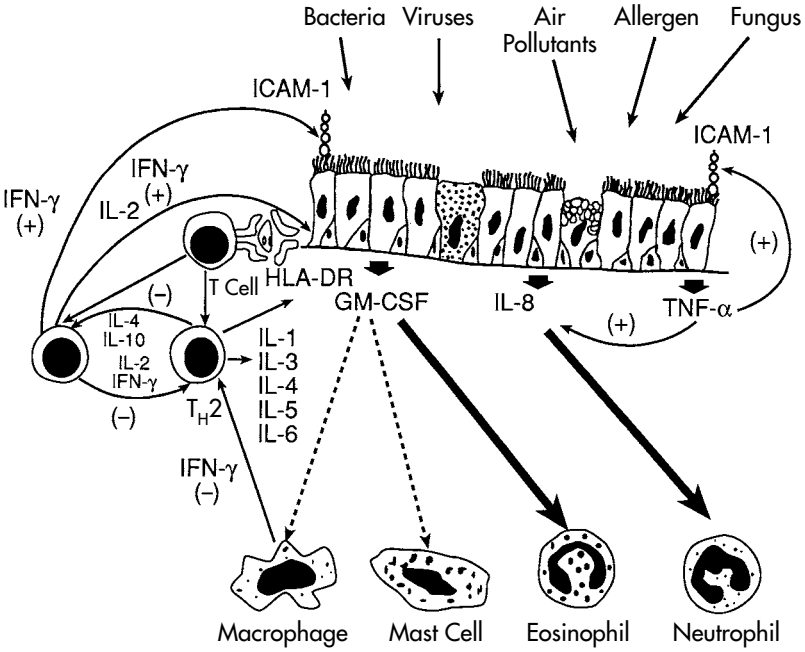


FIGURE 28–22. Diagram of the effect of multiple irritants on the epithelial cell in the lateral wall of the nose. Bacteria, virus, air pollution, allergy, or fungus can have an effect of up-regulating various cytokines such as granulocyte-macrophage colony stimulating factor (GM-CSF), interleukin (IL)-8, or tumor necrosis factor (TNF)-α directly by changing the genetic potential of the respiratory epithelial cell. These cytokines can mobilize from the microvasculature of the nasal polyp various inflammatory cells including neutrophils, macrophages, mast cells, and, particularly, eosinophils. Tumor necrosis factor-α and IL-1β that may be released by the stimulated epithelial cells can up-regulate very late activation antigen-4 and vascular cell adhesion molecule (VCAM)-1 that are receptors on eosinophils and counter-receptors on endothelial cells, respectively, and, in this way, specifically direct eosinophils into the tissue of the nasal polyp. Finally, interferon (IFN)-γ may be released by the epithelial cell and produce intercellular adhesion molecule (ICAM)-1 on the epithelial cell, which is a receptor for the rhinovirus as well as for neutrophils. The five stimulators of the epithelial cell as seen at the top of the figure can also induce human leukocyte antigen (HLA)-DR on the basolateral surface of the cell. This class II antigen can, in turn, up-regulate the immune response for both T_H1 and T_H2 lymphocytes, with the final outcome of the production of the typical cytokines for the T_H1 lymphocytes (IFN-γ and IL-2) and for the T_H2 lymphocyte (IL-4 and IL-5), which, in turn, can up-regulate the eosinophilia in the nasal polyp. It is thus suggested that the lateral wall of the nose, either by mucosal contacts or more likely by specific aerodynamic airflow, can allow various stimulants to activate the epithelial or constitutive cells of the lateral wall of the nose to produce a heightened inflammatory response in a genetically predisposed individual for the development of nasal polyposis.

numbers of cellular changes. For the purposes of this discussion, TNF-α, along with IL-1β, which is discussed briefly below, has the effect of increasing the surface receptors on eosinophils, particularly VLA-4, and of up-regulating VCAM-1 on the endothelial cells of blood vessels of nasal polyps.

Interleukin-1β is produced by macrophages, endothelial cells, T and B cells, fibroblasts, epithelial cells, and osteoblasts. The cellular target for IL-1β includes thymocytes, neutrophils, hepatocytes, chondrocytes, muscle cells, endothelial cells, epidermal cells, osteocytes, macrophages, T and B cells, and fibroblasts. The biologic activities of IL-1β parallel those of TNF-α. In regard to chronic inflammation, IL-1β also has a powerful effect on the up-regulation of VLA-4 on eosinophils and VCAM-1 on nasal polyp vascular endothelial cells. It is emphasized that both of these cytokines (TNF-α and IL-1β) can be synthesized by constitutive cells of the nasal mucosa, including the epithelial cells and the endothelial cells. Thus, any kind of irritation caused by inflammatory mediators or external environmental irritants can stimulate the constitutive cells of the nasal mucosa to produce proinflammatory cytokines, which, in turn, cause a whole array of inflammatory effects in the lateral wall of the nose in the development of nasal polyposis. It is to be emphasized that the cellular immunology as described in this chapter following stimulation of helper T cells, cytotoxic cells, and B cells will give rise to the synthesis of cytokines that produce chronic inflammation.

Integrins

The integrins are the major family of cell-surface receptors that mediate adhesion to the extracellular matrix and sometimes cell-to-cell adhesive interactions.[83] These integrin-mediated adhesive interactions are involved in the regulation of many cellular functions, including embryonic development, tumor cell growth and metastasis, prolonged cell death, homeostasis, inflammation, immune reaction, and bone resorption, among other phenomena. Integrins are composed of α and β transmembrane subunits selected from 16 alpha and 8 beta subunits that heterodimerize to produce more than 20 different receptors that bind specific ligands. Integrins link to intracellular cytoskeletal complexes and bundles of actin filaments. There are many intracellular signaling pathways activated by integrin-ligand interactions. In continuing our explanation of the role of the eosinophil in nasal polyps, the VLA family of integrins is of most importance. Very late activation antigen 4 is also the α4, β1 heterodimer on the surface of eosinophils. Beta-1 integrins are essentially involved in cell–extracellular matrix interactions and β2 subfamily in leukocyte–leukocyte and leukocyte–endothelial cell communications. The soluble α4, β1 fusion protein exhibits divalent cation-dependent binding to VCAM-1. Furthermore, anti-VLA-4 monoclonal antibodies can inhibit eosinophil infiltration in the allergic conjunctivitis model of the guinea pig, suggesting that specific monoclonal antibodies directed against specific integrins or adhesion cell molecules might be the future for the treatment of nasal polyposis. Because VLA-4 is up-regulated on the surface of eosinophils in chronic inflammation, it would be interesting to consider the role of antagonists to VLA-4, also known as integrin α4, β1. The most advanced in this endeavor is a humanized anti-α4 antibody, which is in phase 2 clinical trials for multiple sclerosis. The first reported small-molecule VLA-4 antagonist to advance to clinical trials is currently in phase 1 as an aerosol for treating asthma. Thus, the future of the treatment of many chronic inflammatory diseases, as mentioned above, may be the purification and clinical use of antagonists of various integrins that are up-regulated on the surface of inflammatory cells.

Vascular cell adhesion molecule 1 Lymphocytes migrate from the blood across endothelial cells to reach foreign substances sequestered in peripheral lymphoid organs and inflammatory sites. This migration is inhibited by anti-VCAM-1 or anti-α4 integrin, suggesting that VCAM-1 adhesion is required for migration. Studies suggest that endothelial cells not only are a scaffold for lymphocyte adhesion but also play an active role in promoting lymphocyte migration. In addition to lymphocyte migration, there is abundant evidence to suggest that TNF-α is a cytokine implicated in the pathogenesis of numerous chronic and acute inflammatory conditions. Furthermore, TNF-α is an effective inducer of eosinophil accumulation in vivo, and this response appears to be dependent on an interaction between α4 integrins and VCAM-1. Thus, it appears that the proinflammatory cytokines are capable of up-regulating integrins on the surface of eosinophils, particularly VLA-4, and that this integrin has a specific

predilection to adhere to VCAM-1 on the surface of vascular endothelial cells. The final phase of eosinophilic migration into the nasal polyp stroma is ultimately dependent on the chemokines.

Chemokines The relatively recent appreciation of a new class of cytokines, the chemokines, has done much to enhance our understanding of the extracellular signals involved in the movement of various populations of white blood cells.[84] Investigation of the molecular underpinnings of chemokine function and their involvement in inflammatory processes of all kinds is beginning to yield information about the mechanisms of pathogenesis of a number of conditions, as well as provide hope for new therapeutic insights. Chemokines share structural similarities, including four conserved cystine residues that form disulfide bonds in the tertiary structures. Traditionally, the chemokines superfamily has been divided into two subgroups: C-X-C (where X is any amino acid) and C-C according to whether an intervening residue spaces the first two cysteines in the motif. This structural distinction has been shown to delineate a general, although not absolute, distinction in the biologic properties of these molecules: most C-X-C cytokines are chemoattractants for neutrophils but not monocytes, whereas C-C chemokines appear to attract monocytes but not neutrophils. There are probably more structural distinctions that can be made that may illuminate chemokine function. However, for our purposes, there are two major chemokines of the C-C family known as RANTES and eotaxin.

Eosinophils are the predominant cell type recruited in inflammatory reactions in response not only to allergen challenge but also to the up-regulation of specific cytokines that are related to T_H2 lymphocytes. The data from many studies suggest that RANTES and eotaxin induce eosinophil transepithelial migration with great potency and efficacy. It is interesting to note that activation of the endothelial cells by either TNF-α or IL-1β promotes migration of eosinophils from the vasculature in conjunction with RANTES and eotaxin.

SUMMARY

The immune response through both T cells and B cells can release a cascade of cytokines that are responsible for the specific migration of inflammatory cells into sites of the upper respiratory tract. For the otolaryngologist, knowledge of these cytokines, integrins, and chemokines will be necessary in the future. Although the antibiotic era is still with us, eradicating chronic inflammatory disease such as nasal polyposis, chronic sinusitis, and chronic otitis media with antibiotics probably cannot be accomplished. Our knowledge of the molecular biology of the inflammatory response will require knowledge of the biology of the immune system and the inflammatory mediators that are released by immune cells such as T cells, B cells, macrophages, fibroblasts, and other cells. It may very well be that the future of clinical medicine will be an understanding of the molecular nature of these proteins and glycoproteins released by the immune cells as well as inflammatory cells and to counteract their activity with specific monoclonal antibodies directed against either the cytokine itself or the receptor for the cytokine on tissue cells. Ultimately, this will be the future of otolaryngology as well as medicine in general.

REFERENCES

1. Verdoliva A, Ruvo M, Cassani G, Fassina G. Topological mimicry of cross-reacting enantiomeric peptide antigens. J Biol Chem 1995;270:30422–7.
2. Gowans JL, McGregor DD. The immunological activities of lymphocytes. Prog Allergy 1965;9:1–78.
3. Greaves MF, Owen JJT, Raff MC. T and B lymphocytes: origins, properties and roles in immune responses. Amsterdam: Excerpta Medica; 1973.
4. Thomas Y, Rogozinski L, Chess L. Relationship between human T cell functional heterogeneity and human T cell surface molecules. Immunol Rev 1983;74:113–28.
5. Frasca D, Guidi F, Barattini P, Doria G. Haemopoietic reconstitution after sublethal irradiation: comparison of the effects of different haemopoietic cytokines on murine lymphocytes and bone marrow cells. Br J Haematol 1999;106:317–24.
6. Grossi CE, Velardi A, Cooper MD. Postnatal liver hemopoiesis in mice: generation of pre-B cells, granulocytes and erythrocytes in discreet colonies. J Immunol 1985;135:2203–11.
7. Liu CP, Auberbach R. Ontogeny of murine T cells: thymus-regulated development of T cell receptor-bearing cells derived from embryonic yoke sack. Eur J Immunol 1991;21:1849–55.

8. Yen CJ, Lin Sl, Huang KT, Lin RH. Age-associated changes in interferon-gamma and interleukin-4 secretion by purified human CD4+ and CD8+ T cells. J Biomed Sci 2000;7:317–21.
9. Ada GL, Nossal G. The clonal selection theory. Sci Am 1987;257:62–9.
10. Siiman O, Burshteyn A. Cell surface receptor-antibody association constants and enumeration of receptor sites for monoclonal antibodies. Cytometry 2000;40:316–26.
11. Claeys S, Cuvelier C, Quatacker J, Van Cauwenberg P. Ultrastructural investigation of M-cells and lymphoepithelial contacts in nasopharyngeal associated lymphoid tissue (NALT). Acta Otolaryngol Suppl (Stockh) 1996;523:40–2.
12. Bernstein JM, Gorfien J, Brandtzaeg P. The immunobiology of the tonsils and adenoids. In: Ogra PL, Lamm ME, McGhee JR, et al, editors. Mucosal immunology. 2nd ed. San Diego: Academic Press; 1999. p. 1339–62.
13. Reinherz EL, Schlossman SF. The differentiation and function of human lymphocytes. Cell 1980;19:821–7.
14. Whyte A, Licence ST, Robinson MK, VanderLienden K. Lymphocyte subsets and adhesion molecules in cutaneous inflammation induced by inflammatory antagonists: correlation between E-selectin and gamma delta CtR+ lymphocytes. Lab Invest 1996;75:439–49.
15. Butcher EC, Williams M, Youngman K, et al. Lymphocyte trafficking and regional immunity. Adv Immunol 1999;72:209–53.
16. Bernstein JM. The immunobiology of the tonsils and adenoids. In: Johnson J, Blitzer A, Ossoff R, Thomas JR, editors. Instructional courses. St. Louis: Mosby Yearbook; 1990. p. 3–20.
17. Jia GQ, Gonzalo JA, Hidlalgo A, et al. Selective eosinophil transendothelial migration triggered by eotaxin via modulation of Mac-1/ICAM-1 and VLA-4/VCAM-1 interactions. Int Immunol 1999;11:1–10.
18. Harris DE, Cairns L, Rosen FS, Borel Y. A natural model of immunological tolerance. Tolerance to murine C5 is mediated by T cells, and antigen is required to maintain unresponsiveness. J Exp Med 1982;156:567–84.
19. Lindstrom J. Immunobiology of myasthenia gravis, experimental autoimmune myasthenia gravis, and Lambert-Eaton syndrome. Annu Rev Immunol 1985;3:109–32.
20. Davies DR, Metzger R. Structural basis of antibody function. Annu Rev Immunol 1983;1:87–118.
21. Nisonoff A, Hopper JE, Spring SD. The antibody molecule. New York: Academic Press; 1975.
22. Reth M. Antigen receptors on B lymphocytes. Annu Rev Immunol 1992;10:97–122.
23. DeFranco AL. Structure and function of the B cell antigen receptor. Annu Rev Cell Biol 1993;9: 377–410 .
24. Ohno S. The origin of immunoglobulins and T-cell receptors is likely to be the cell death sensor of macrophages. Res Immunol 1996;147:247–52.
25. Edelman GM. The structure and function of antibodies. Sci Am 1970;223:34–42.
26. Porter RR. Structural studies of immunoglobulins. Science 1973;180:713–6.
27. Kabat EA. Structural concepts in immunology and immunochemistry. 2nd ed. New York: Holt, Rinehart and Winston; 1976.
28. Brandtzaeg P, Farstad IN. The human mucosal B-cell system. In: Ogra PL, et al, editors. Mucosal immunology, 2nd ed. San Diego: Academic Press; 1999. p. 439–68.
29. Koshland ME. The coming of age of the immunoglobulin J chain. Annu Rev Immunol 1985; 3:425–54.
30. Muller-Eberhard HJ. Molecular organization and function of the complement system. Annu Rev Biochem 1988;57:321–48.
31. Burton DR, Woof JM. Human antibody effector function. Adv Immunol 1992;51:1–84.
32. Kraehenbuhl JP, Neutra MR. Transepithelial transport and mucosal defense. II: secretion of IgA. Trends Cell Biol 1992;2:170–4.
33. Metzger H. The receptor with high affinity for IgE. Immunol Rev 1992;125:37–48.
34. VanderKeyl H, Gellad ZF, Owen JA. Disparity in the kinetics of onset of hypermutation in immunoglobulin heavy and light chains. Immunol Cell Biol 2000;78:224–37.
35. Reid KB. Structure-function relationships of the complement components. Immunol Today 1989;10: 177–80.
36. Tomlinson S. Complement defense mechanisms. Curr Opin Immunol 1993;5:83–9.
37. Reid KB. Activation and control of the complement system. Essays Biochem 1986;22:27–68.
38. Bernstein JM, Faden HS, Loos BG, et al. Recurrent otitis media with non-typable *Haemophilus influenzae*: the role of serum bactericidal antibody. Int J Pediatr Otorhinolaryngol 1992;23:1–13.
39. Bernstein JM, Reddy MS, Scannapieco FA, et al. The microbial ecology and immunology of the adenoid:

implications for otitis media. Ann N Y Acad Sci 1997;830:19–31.
40. Bernstein JM, Scheerer R, Shoenfeld E, Albini B. The distribution of immunocompetent cells in the compartments of the palatine tonsils in bacterial and viral infections of the upper respiratory tract. Acta Otolaryngol Suppl (Stockh) 1988;454: 153–62.
41. Capra JD, Edmundson AB. The antibody combining site. Sci Am 1977;236:50–9.
42. Tomlinson S. Complement defense mechanisms. Curr Opin Immunol 1993;5:83–9.
43. Wu TT, Kabat EA. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. J Exp Med 1970;132: 211–50.
44. Davies DR, Padlan EA, Sheriff S. Antibody-antigen complexes. Annu Rev Biochem 1991;59:439–73.
45. Sakano H, Rogers JH, Huppi JK, et al. Domains and a hinge region of an immunoglobulin heavy chain encoded in separate DNA segments. Nature 1979;277:627–33.
46. Fitzsimmons SP, Rotz BT, Shapiro MA. Asymmetric contribution to Ig repertoire diversity by V kappa exons: differences in the utilization of the V kappa 10 exons. J Immunol 1998;161:290–300.
47. Leder P. The genetics of antibody diversity. Sci Am 1982;246:72–83.
48. Hozumi N, Tonegawa S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. Proc Natl Acad Sci U S A 1976;73:3628–32.
49. Nadal DE, Soh N, Schlapfer E, et al. Distribution characteristics of immunoglobulin-secreting cells in the adenoids. Relationship to age and disease. Int J Pediatr Otorhinolaryngol 1992;24:121–30.
50. Disis ML, Schiffman K, Gooley TA, et al. Delayed-type hypersensitivity response is a predictor of peripheral blood T cell immunity after HER-2/new peptide immunization. Clin Cancer Res 2000;6: 1347–50.
51. Rothenberg EV. Stepwise specification of lymphocyte development lineages. Curr Opin Genet Dev 2000;10:370–9.
52. Kupfer A, Singer SJ. Cell biology of cytotoxic and helper T-cell function. Annu Rev Immunol 1989; 7:309–37.
53. Marrack P, Kappler J. The T cell receptor. Science 1987;238:1073–9.
54. Davis MM. T cell receptor gene diversity and selection. Annu Rev Biochem 1990;59:475–96.
55. Swain SL, Bradley LM, Croft M, et al. Helper T-cell subset: phenotype, function and the role of lymphokines in regulating their development. Immunol Rev 1991;123:115–44.
56. Rammensee H, Falk K, Rotzschke O. MHC molecules as peptide receptors. Curr Opin Immunol 1993;5:35–44.
57. Auffray C, Strominger JL. Molecular genetics of the human histocompatibility complex. Adv Hum Genet 1987;15:197–247.
58. Moller G, editor. Molecular genetics of class I and class II MHC antigens. Immunol Rev 1985;43:84, 85.
59. Matsumura M, Fremont DH, Peterson PA, Wilson IA. Emerging principles for the recognition of peptide antigens by MHC class I molecules. Science 1992;257:927–34.
60. Johnson DR, Mook-Kanamori B. Dependence of elevated human leukocyte antigen class I molecule expression on increased heavy chain-light chain (beta 2-microglobulin) transporter associated with antigen processing, tapasin, and peptide. J Biol Chem 2000;275:16643–9.
61. Bach FH, Sacks DH. Transplantation immunology. N Engl J Med 1987;317:489–92.
62. Janeway CA. The T cell receptor as a multi-component signaling machine: CD4/CD8 co-receptors and CD45 and T cell activation. Annu Rev Immunol 1992;10:645–74.
63. Townsend AR. Recognition of influenza virus proteins by cytotoxic T lymphocytes. Immunol Res 1987;6:80–100.
64. De Maeyer E, De Maeyer-Guignar J. Interferon-γ. Curr Opin Immunol 1992;4:321–6.
65. Yagita H, Nakata M, Kawasaki A, et al. Role of perforin in lymphocyte-mediated cytolysis. Adv Immunol 1992;51:215–42.
66. Kawasaki A, Shinkai Y, Yagita H, Okumura K. Expression of perforin in murine natural killer cells and cytotoxic T lymphocytes in vivo. Eur J Immunol 1992;22:1215–9.
67. Koyama AH, Fukumori T, Fujita M, et al. Physiological significance of apoptosis in animal virus infection. Microb Infect 2000;2:1111–7.
68. Tangri S, Brossay L, Burdin N, et al. Presentation of peptide antigens by mouse CD1 requires endosomal localization and protein antigen processing. Proc Natl Acad Sci U S A 1998;95:14314–9.

69. Schwartz RH. Co-stimulation of T lymphocytes: the role of CD28, CTLA-4 and B7/BB1 in interleukin-2 production and immunotherapy. Cell 1992;71: 1065–8.
70. Taniguchi T, Minami Y. The IL-2/IL-2 receptor system: a current overview. Cell 1993;73:5–8.
71. Noelle RJ, Ledbetter JA, Aruffo A. CD40 and its ligand, an essential ligand-receptor pair for thymus dependent B-cell activation. Immunol Today 1992; 13:431–3.
72. Moller G, editor. IL-4 and IL-5: biology and genetics. Immunol Rev 1988;105:102–22.
73. Bar-Or RL. Feedback mechanisms between T helpers cells and macrophages in the determination of the immune response. Math Biosci 2000;163:353–8.
74. Jinquan T, Quan S, Jacobi HH, et al. Cutting edge: expression of the NF of activated T cells in eosinophils: regulation by IL4 and IL5. J Immunol 1999;163:21–4.
75. Sgagias MK, Nieroda C, Yannelli JR, et al. Upregulation of DF3, in association with ICAM-1 and MHC class II by interferon-γ in short-term human mammary carcinoma cell cultures. Can Biother Radio Pharm 1996;11:177–85.
76. Jahnsen FL, Haraldsen G, Aanesen JP, et al. Eosinophil infiltration is related to increased expression of vascular cell adhesion molecule-1 in nasal polyps. Am J Respir Cell Mol Biol 1995;12:624–32.
77. Luttmann W, Matthiesen T, Mattys H, Virchow JC. Synergistic effects of interleukin-4 or interleukin-13 and tumor necrosis factor-α on eosinophil activation in vitro. Am J Respir Cell Mol Biol 1999;20: 474–80.
78. Bocchino V, Bertoreli G, D'Ippolito R, et al. The increased number of very late activation antigen-4-positive cells correlates with eosinophils and severity of disease in the induced sputum of asthmatic patients. J Allergy Clin Immunol 2000;105:65–70.
79. Lee CH, Lee KS, Rhee CS, et al. Distribution of RANTES and interleukin-5 in allergic nasal mucosal nasal polyps. Ann Otol Rhinol Laryngol 1999;108:594–8.
80. Jacoby DB, Ueki IF, Widdicombe JH, et al. Effect of human eosinophil major basic protein on ion transport in dog tracheal epithelium. Am Rev Respir Dis 1988;137:13–6.
81. Kronenwett R, Martin S, Haas R. The role of cytokines and adhesion molecules for mobilization of peripheral blood stem cells. Stem Cells 2000; 18:320–30.
82. Hamblin AS. Cytokines and cytokine receptors. London: IRL Press, Oxford University in association with the British Society for Immunology; 1993.
83. Lustig L, Denduchis B. Integrins: a family of cell adhesion receptors. Medicina 1993;53:357–63.
84. Schall TJ, Bacon KB. Chemokines, leukocyte trafficking and inflammation. Curr Opin Immunol 1994;6:865–73.

CHAPTER 29

Acquired Immune Deficiency Syndrome

Frank E. Lucente, MD, Samir Shah, MD, Jeffrey Vierra, MD

Acquired immune deficiency syndrome (AIDS) has emerged as one of the most extraordinary diseases in human history. We have seen the unfolding of a medical mystery in which an apparently new disease has appeared, first reported in the United States in 1981, and has since become a major worldwide epidemic. Acquired immune deficiency syndrome is caused by human immunodeficiency virus (HIV). Human immunodeficiency virus infection of macrophages, monocytes, and CD4 (helper) T lymphocytes leads to a cell-mediated immune deficiency that predisposes to a constellation of opportunistic infections and malignancies.[1,2]

More than 1,200,000 cases of AIDS have been reported in the United States since 1981, and as many as 900,000 Americans are estimated to be living with HIV infection. The epidemic is growing most rapidly among minority populations and is a leading killer of African American males. According to the US Centers for Disease Control and Prevention (CDC), the prevalence of AIDS is six times higher in African Americans and three times higher in Hispanics than among whites.[3]

Acquired immune deficiency syndrome has had a significant impact on the practice of otolaryngology. The fact that many patients originally presented with physical findings in the head and neck region, especially during the 1980s, required a high level of vigilance for detecting these findings as manifestations of HIV infection. Implementation of measures for handling instruments in the office and operating room emerged as part of the universal precautions that were promulgated as a way of controlling spread of the infection.

EPIDEMIOLOGY

To date, an estimated 36 million persons are living worldwide with HIV/AIDS. Of these, most reside in sub-Saharan Africa and Southern and Southeast Asia. Less than 3% of the HIV infected worldwide reside in North America. Since the start of the epidemic, more than 12 million persons have died of AIDS-related complications worldwide. Of the 16,000 new cases of HIV infection that are estimated to occur each day worldwide, more than 40% are among women, more than 50% are among 15 to 24 year olds, and 10% are among children less than 14 years old.[4]

Human immunodeficiency virus transmission occurs by unprotected heterosexual or homosexual sexual contact, use of HIV-infected contaminated needles, infusion of contaminated blood and blood products, or from mother to child in utero, perinatally or through breast-feeding. The patterns of transmission vary from country to country.[1,5] In sub-Saharan Africa and Southern and Southeast Asia, heterosexual transmission is the predominant mode of transmission, although intravenous drug use is a major factor in India and parts of Southeast Asia. In Northern Europe, homosexual transmission is most common, whereas intravenous drug use is the major risk factor in Southern Europe, particularly Spain and Italy. In the United States, homosexuals were the major risk group at the outset of the epidemic, but over the past several years, the rate of increase in new cases has been greatest among women, heterosexuals, and minorities.

There are significant racial and socioeconomic differences in the distribution of HIV/AIDS cases among the various risk groups. Whites predominate among homosexuals with HIV, whereas African Americans and Hispanics are disproportionately represented among HIV positives, with injection-drug use as a risk factor. Poverty has been shown to be independently associated with higher rates of HIV/AIDS, and racial minorities are disproportionately represented among those living below the poverty line.

TRANSMISSION

The known routes of transfer of the HIV virus are blood, blood products, unsafe sexual activity, and transmission from mother to child perinatally. Although researchers have detected HIV in the saliva of infected individuals, no evidence exists that the virus is spread by contact with saliva. Laboratory studies reveal that saliva has natural compounds that inhibit the infectivity of HIV.[6]

Perinatal transmission is responsible for almost all new cases of HIV in children.[7] The rate of perinatal transmission has declined more than 50% since 1993 with the introduction of voluntary testing and counseling of pregnant women and the widespread implementation of the 1994 US Public Health Service recommendations for zidovudine (AZT) therapy of pregnant women.[6] Current recommendations call for instituting antiretroviral therapy in pregnant women using complete antiretroviral regimens as one would use were the patient not pregnant.[8]

Currently, the blood supply in developed countries is safe from HIV contamination by virtue of donor deferral and routine screening for HIV antibodies. In the United States, the risk of acquiring HIV from screened blood is estimated as 1 in 225,000 units; the rare infection would be caused by blood from an infected individual who had not seroconverted at the time of donation.

Human immunodeficiency virus transmission from patient to health care worker and vice versa via accidental exposure to contaminated secretions or blood is rare. Universal precautions or standard precautions have been recommended by the CDC as a means to prevent transmission of bloodborne pathogens. The precautions applied to all patients regardless of specific diagnosis include the use of gloves when contact with blood and potentially infected secretions is possible, protective eyewear when splattering of blood or secretions is possible, and gowns when clothing is likely to be soiled. An important element of prevention of nosocomial spread of disease is needle-stick prevention. Concerns about needle-stick injury and HIV transmission has spurred the development of safer needle devices, which have been designed to retract or cover the needle after use, and "needleless" connectors for intravenous devices. The impact of these devices and other innovations on preventing nosocomial transmission of HIV and other pathogens has yet to be demonstrated.

In instances of significant occupational exposure to HIV, postexposure prophylaxis is recommended. The degree of risk is a function of many factors, including the type of exposure (deep versus superficial punctures, whether there was actual injection of blood, mucosal or cutaneous exposures on intact versus abraded or open surfaces), the volume of infected blood or secretion (exposure with a hollow, blood-filled needle versus a scalpel blade or suture needle), postinjury treatment of the wound, time between injury and evaluation for postexposure prophylaxis, and the ex vivo time of the blood or secretions prior to exposure. Factors of critical importance, but frequently unknown at the time of exposure, are the patient's prior viral therapy history, the resistance pattern of the patient's HIV strain, and the viral burden or "load." When an injury is assessed as posing a significant risk of transmission of HIV from a known or suspect source, postexposure prophylaxis should be instituted as soon as possible after the injury and certainly no longer than 24 hours after exposure. Highly active antiretroviral regimens are instituted, the selection of agents being chosen based on known prior treatment and viral resistance profile of the source when available, and are continued for 4 to 6 weeks. The exposed health care worker should be encouraged to have pretreatment HIV testing with follow-up testing at 6 weeks, 3 months, and 6 months. Postexposure counseling should include a discussion of the degree of risk from the specific exposure and a review of universal precautions and methods to prevent injuries in the future. The health care worker should practice safer sex using barrier protection to minimize potential transmission of incubating HIV infection. Counselors should be available to the health care worker and to his/her family members to provide information and support following exposure and through the follow-up period until infection has been ruled out with the 6-month HIV test.

THE HIV-VIRUS AND ITS LIFE CYCLE

Human immunodeficiency virus is a member of the *Lentivirinae* subfamily of retroviruses. The HIV virion has a lipid bilayer perforated by 72 external spikes consisting of the two major viral envelope proteins, glycoprotein 120 and 41 (Figure 29–1). The

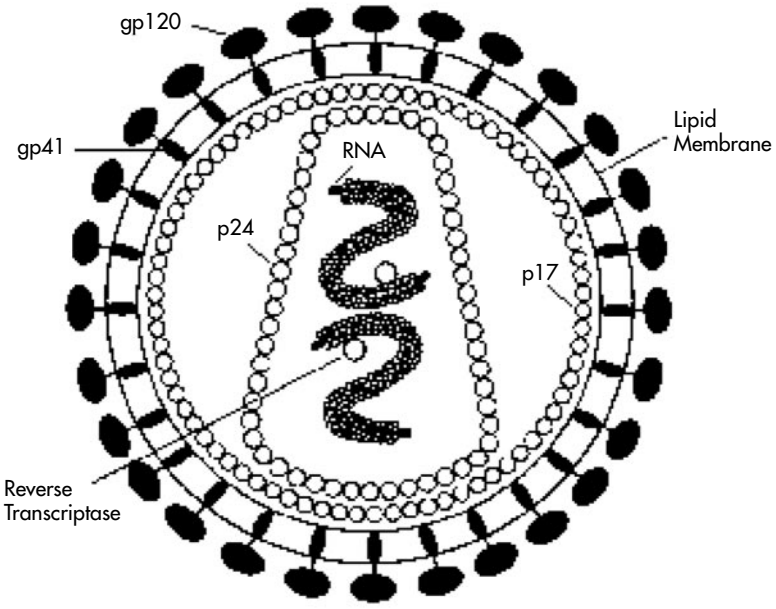


FIGURE 29–1. Human immunodeficiency virus structure. gp = glycoprotein; RNA = ribonucleic acid.

viral core contains four proteins: the p24 capsid protein, the p17 matrix protein, and the p6 and p7 nucleocapsid proteins. The core also contains two copies of single-stranded ribonucleic acid (RNA) and three viral enzymes: reverse transcriptase, integrase, and protease.

Human immunodeficiency virus infects cells by binding to a combination of cell-surface receptors. The CD4 receptor is a high-affinity target for binding by glycoprotein 120, which is necessary but not sufficient for binding of the virion to the cell. CD4 receptors are expressed on T lymphocytes, macrophages, and, to a lesser extent, a variety of somatic cells. Some strains of HIV preferentially infect CD4 (helper) T lymphocytes, whereas others preferentially infect macrophages. This preference or tropism is mediated by the differential expression of chemokine receptors on T cells and macrophages. Human immunodeficiency virus strains trophic for activated T lymphocytes bind by the CD4 receptor and CXCR4 chemokine coreceptor. Macrophage tropic strains bind by the CD4 receptor and the CCR5 chemokine receptor.

After binding to the cell receptors, the HIV envelope fuses to the cell membrane, and the virus is internalized (Figure 29–2). Viral replication begins with the viral enzyme reverse transcriptase mediating the synthesis of deoxyribonucleic acid (DNA) complementary to the viral RNA genome. This DNA-RNA heteroduplex then dissociates, and then a second DNA strand, complementary to the first, is synthesized. Viral encoded integrase enzyme then mediates the incorporation of the double-stranded DNA into the host genome. Transcription of RNA from this proviral DNA provides the coding for viral polyproteins, which are cleaved by viral proteins to produce the structural proteins and enzymes, which are eventually incorporated into progeny virus, which bud off the infected cell.

CLINICAL COURSE

The median time from infection with HIV to the development of AIDS-related symptoms among untreated patients is approximately 10 years, with a wide variation in time to progression. Approximately 10% of HIV-infected persons progress to AIDS within the first 2 to 3 years following infection. In contrast, 5 to 10% of patients are so-called nonprogressors, having stable CD4 cell counts after 15 or more years of follow-up. Factors that determine the rate of progression of untreated disease include patient age (more rapid progression over age 40 years), genetic differences, the virulence of the infecting strain, the viral load that is directly proportional to the rate of decline of the CD4 lymphocyte count, and co-infection with other microbes (eg, hepatitis C). The onset of clinical disease is related to the decline in immune competence as CD4 lymphocytes are depleted.[1–3] The most signifi-

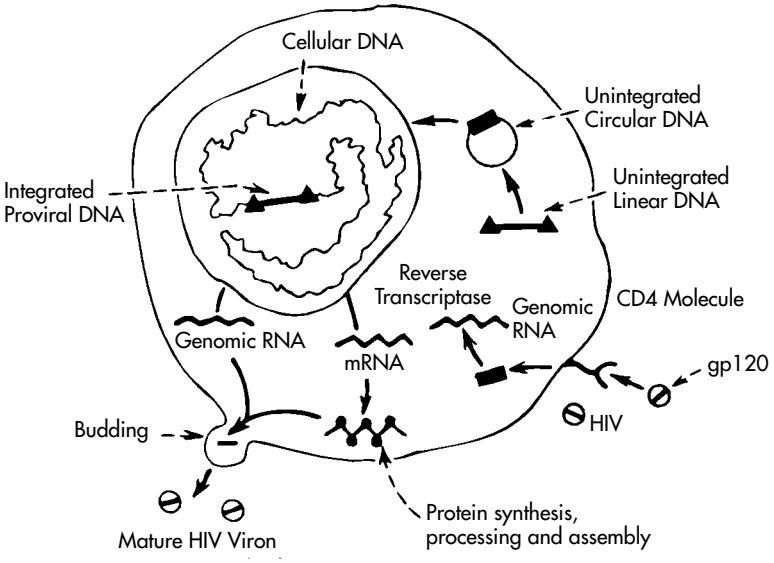


FIGURE 29–2. Human immunodeficiency virus (HIV) life cycle: on the *right side* of the diagram, the glycoprotein (gp) 120 has tropism for the CD4 molecule. The HIV genomic ribonucleic acid (RNA) is uncoated and enters the cell. The genomic RNA is transcribed by the reverse transcription enzyme to double-stranded deoxyribonucleic acid (DNA). The DNA translocates to the nucleus, where it becomes integrated into the cellular DNA by the enzyme integrase as a provirus. It can remain dormant for years or be activated to transcribe messenger and genomic RNA of the virus. The virus is assembled at the periphery and buds off as a mature HIV variant. mRNA = messenger ribonucleic acid. Adapted from Fauci AS. The human immunodeficiency virus: infectivity and pathogenesis. Science 1988;239:617.

cant complications, that is, opportunistic infections and malignancies, occur after reduction of CD4 cell below 200/mm^3.

OPPORTUNISTIC INFECTIONS

Opportunistic infections have remained the proximate cause of death for nearly all AIDS patients. Advances in prophylaxis have decreased the incidence of *Pneumocystis, Toxoplasma, Mycobacterium avium-intracellulare* complex (MAC), *Cryptococcus*, and other opportunistic infections and consequently their contribution to morbidity and mortality.[1]

Patterns of specific opportunistic infections vary geographically, among risk groups, and as a result of medical interventions. In the United States and Europe, over 90% of AIDS patients with Kaposi's sarcoma are homosexual or bisexual men, probably because they are co-infected with human herpesvirus 8, a newly identified viral cofactor (with HIV) for Kaposi's sarcoma. Toxoplasmosis and tuberculosis (TB) are more common in tropical areas, where the prevalence of latent infections with *Toxoplasma gondii* and *Mycobacterium tuberculosis* in the general population is high. Even in developed countries, where background levels of TB are low, HIV has caused increased rates and atypical presentations of TB. Widespread use of effective prophylaxis against such agents as *Pneumocystis carinii* and MAC has reduced the risk of these infections in developed countries.

ANTIRETROVIRAL THERAPY

Recommendations regarding the use of antiviral drugs in HIV are in flux. When and what to initiate, when to change regimens, and how to minimize the development of resistance and cross-resistance are continually being re-evaluated. Clearly, monotherapy results in resistance and loss of efficacy as a result of the huge viral burden, short viral half-life, and propensity to mutate.

Viral loads are critical in determining the efficacy of regimens; the goal is to make all viral loads undetectable because high loads drive CD4 loss and ultimate immune suppression. The current recommendation is to initiate a three-drug regimen in patients with a detectable viral load of over 5,000 to

10,000 HIV RNA copies per milliliter, regardless of CD4 count. This regimen offers sustained viral suppression compared with double- and single-drug regimens.

Triple combinations containing a protease inhibitor are considered the most potent of all regimens. The difficulty with multidrug therapy is that the patient may not fully comply because of the number of pills and adverse effects. Even minimal noncompliance causes drug resistance and loss of efficacy. When changing a failing regimen, two new drugs (preferably three new drugs) should be started. All regimens must be individualized, and, occasionally, when patients are unable to comply with the rigors of three drugs, double therapy is preferable to no therapy. The classes of anti-HIV drugs include reverse transcriptase inhibitors (nucleoside and non-nucleoside) and protease inhibitors (Table 29–1).[9]

Nucleoside reverse transcriptase inhibitors are phosphorylated to active metabolites that compete for incorporation into viral DNA. They inhibit the HIV reverse transcriptase enzyme competitively and act as a chain terminator of DNA synthesis. Non-nucleoside reverse transcriptase inhibitors directly bind to the reverse transcriptase enzyme at sites separate from the nucleoside-binding site. In general, viral resistance to these drugs occurs quickly, and they should not be used as monotherapy except in specific cases.

Protease inhibitors are the most potent class of antiviral drugs; they target the viral proteinase enzyme. Inhibition of viral protease prevents cleavage of an important structural polyprotein that results in noninfectious viral particles. Primary and accessory mutations of the gene encoding the viral proteinase lead to cross-resistance between this class of drugs. The only well-studied dual-protease regimen is ritonavir and saquinavir, which has demonstrated significant, sustained viral load reduction and CD4 gain. Some in vitro data suggest that saquinavir and indinavir are antagonists. Other combinations of multiple protease inhibitors are under study. The protease inhibitors are metab-

TABLE 29–1. Drugs for Human Immunodeficiency Virus Iinfection

Drug	*Usual Oral Adult Dosage*	*Cost**
Protease Inhibitors		
Amprenavir (Agenerase, Glaxo Wellcome)	1,200 mg bid	604.80
Indinavir (Crixivan, Merck)	800 mg q8h	463.50
Nelfinavir (Viracept, Agouron)	750 mg tid or 1,250 mg bid	583.20
Ritonavir (Norvir, Abbott)	600 mg bid	667.82
Saquinavir (Invirase, Roche	600 mg tid	586.36
(Fortovase, Roche)	1,200 mg tid or 1,800 mg bid	587.85
Nucleoside Reverse Transcriptase Inhibitors		
Abacavir (ABC; Ziagen, Glaxo Wellcome)	300 mg bid	349.20
Didanosine (ddI; Videx, Bristol-Myers Squibb)	200 mg bid	217.60
Lamivudine (3TC; Epivir, Glaxo Wellcome)	150 mg bid	259.86
Stavudine (d4T; Zerit, Bristol-Myers Squibb)	40 mg bid	273.70
Zalcitabine (ddC; Hivid, Roche)	0.75 mg tid	212.10
Zidovudine (ZDV; Retrovir, Glaxo Wellcome)	200 mg tid or 300 mg bid	303.60
Zidovudine plus lamivudine (Combivir, Glaxo Wellcome)	1 tablet bid	563.46
Non-nucleoside Reverse Transcriptase Inhibitors		
Delavirdine (Rescriptor, Pharmacia & Upjohn)	400 mg tid	239.77
Efavirenz (EFV; Sustiva, Dupont)	600 mg qd	394.20
Nevirapine (Viramune, Roxane)	200 mg bid	278.64

*Cost to the pharmacist for 30 days' treatment, based on wholesale price (AWP) listings, June 1999 in US dollars.
Adapted from Abramowicz M.[9]

olized through cytochrome P-450, and all concomitant drugs should be reviewed for potential interactions.

OTORHINOLARYNGOLOGIC MANIFESTATIONS

Patients who are HIV positive can acquire otorhinolaryngologic diseases similar to those experienced by patients without HIV infection. However, the patient may experience unusual presentations and recurrent or refractory disease. The evaluation and treatment of patients with HIV-related head and neck disease frequently require an interdisciplinary approach, including the input of the infectious diseases specialist, oncologist, ophthalmologist, dermatologist, dentist, and oral surgeon.

Ear

Acute bacterial or serous otitis media is frequently encountered among HIV-infected individuals.[10] Patients present with aural fullness, hearing loss, and, occasionally, otorrhea. The clinical response to appropriate therapy with antibiotics and decongestants is in part dependent on the patient's underlying immune status. Patients with well-preserved cell-mediated immune function as measured by the CD4 cell count are more likely to respond normally, whereas those with impaired immunity more frequently experience refractory or recurrent disease requiring tympanostomy tube placement.

Otitis externa is caused by the same bacterial pathogens as in the non–HIV-positive population.[11] Unusual conditions, which may involve the auricle, pinna, and external auditory canal, are Kaposi's sarcoma, molluscum contagiosum, herpes simplex virus (HSV), varicella-zoster virus, and seborrheic dermatitis. Diagnosis of these conditions is usually clinical, and treatment is specific to the etiology.

Sensorineural hearing loss can occur in up to 69% of AIDS patients; cytomegalovirus (CMV) is often implicated because CMV inclusion-bearing cells can sometimes be found in cranial nerve VIII in the internal auditory canal, in the cochlea, or in vestibular endolymphatic epithelium. Other causes of the hearing loss may include primary cochlear effects of the HIV virus, ototoxic drugs, or demyelination in the auditory tract.[12] Facial paralysis is common in HIV patients and is probably a reflection of the higher incidence of viral infections.

Skin

Dermatologic involvement in the head and neck also occurs quite commonly.[12] Seborrheic dermatitis presents with erythema and scaling of the skin of the forehead and the malar area, although it may be more generalized. Treatment is with topical corticosteroids and topical antifungal agents. Eosinophilic folliculitis or "itchy bump disease" presents as an intensely pruritic papular rash, which can occur on the head and neck but more commonly involves the upper back, upper chest, and proximal limbs. The treatment of this condition of unknown cause is difficult. Topical or systemic corticosteroids, phototherapy, and high-dose itraconazole have all been demonstrated to provide relief for some patients.

Skin diseases owing to virus include dermatomal varicella-zoster (shingles), herpes simplex, and molluscum contagiosum. Herpes zoster is frequently an early indication of underlying HIV infection. In patients with well-preserved immune function, the natural history of the infection is similar to that of non–HIV-positive persons. The painful vesicular/pustular lesions usually remain localized to a single dermatome. The time course from initial presentation as a papule to crusting and healing is 2 to 3 weeks. In patients with significant immune deficiency, the risk of dissemination is increased, and the duration of acute illness may be prolonged. Complications of herpes zoster include Ramsay Hunt syndrome (facial palsy owing to facial nerve involvement); ophthalmic involvement including keratitis, iritis, uveitis or retinitis caused by involvement of the ophthalmic division of the trigeminal nerve; and postherpetic neuralgia, which may cause prolonged or permanent dysesthesias. The diagnosis of shingles is made clinically in most cases. Definitive diagnosis by Tzanck preparation of vesicular fluid demonstrating multinucleated giant cells or culture of the fluid is rarely necessary. Treatment is indicated for all patients with immune deficiency to hasten crusting and healing, reduce infectivity, and prevent dissemination. Early treatment, that is, within 24 to 48 hours after the onset of rash, with systemic antiviral agents has been shown to decrease the risk of postherpetic neuralgia. Recur-

rences of shingles are uncommon, and thus routine use of chronic suppressive therapy, is not indicated.

Herpes simplex presents with characteristic pruritic or painful papules of the vermilion border of the lip. Over a period of 7 to 10 days, the lesions progress through vesicular, pustular, crusting, and ulcerative stages before spontaneously healing. In patients with advanced cell-mediated immune deficiency, the ulcerative lesions may persist and extend widely unless treated with systemic antiviral agents. Recurrences of orolabial herpes are common, but chronic suppression should be limited to those patients who have demonstrated a pattern of frequent recurrences.

Nose

Externally, herpex simplex can often be seen involving the nasal vestibule.[12] The infection presents as a chronic nonhealing inflammatory process that extends onto the septum, the nasal alae, and, ultimately, the upper lip and face. Again, early culture for diagnosis and treatment with appropriate antiviral agents will show substantial improvement.

Chronic rhinitis with nasal crusting and dryness is particularly common among patients with HIV infection. Although no specific treatment has been routinely successful, administration of guaifenesin (1,200 mg twice daily) and humidification are useful.

Sinusitis occurs with increasing frequency as the CD4 cell count decreases. The bacteriology of acute sinusitis in patients with a CD4 cell count above 200/mm^3 is similar to that in non-HIV patients. Treatment for these patients is directed against *Streptococcus pneumoniae* and *Haemophilus influenzae*. In patients with more severe immune deficiency, microorganisms such as *Pseudomonas aeruginosa*, *Staphylococcus aureus*, fungal pathogens, and other opportunists become more important. Empiric broad-spectrum antibiotic therapy is indicated in those patients with more advanced immune deficiency.[13] Paranasal sinus puncture may be necessary for accurate microbiologic diagnosis in a patient who fails empiric therapy.

Allergic rhinitis has been shown to occur secondary to eosinophilia and an increase in immunoglobulin E levels from HIV.[12] Treatment with topical nasal corticosteroids and second-generation antihistamines alleviates symptoms. Allergy testing may be useful to identify allergens that can be avoided. Desensitization shots are generally avoided, however, owing to the perceived risks of CD4 cell activation that may lead to increased HIV replication.

Oral Cavity and Oropharynx

Oral candidiasis is one of the most common head and neck problems encountered in HIV-infected patients. It may present with tender, erythematous mucosa of the tongue, palate, and other mucosal surfaces or with white plaques or pseudomembranes, which can be readily scraped to reveal an erythematous base. Diagnosis is usually made clinically. A microbiologic diagnosis can be made with KOH smear and/or culture of the exudate. Treatment with topical agents such as nystatin swish or clotrimazole troches is usually effective for oral disease. With more extensive disease or esophageal involvement associated with odynophagia and dysphagia, systemic therapy with an absorbable agent such as ketoconazole, fluconazole, or itraconazole is indicated.

Angular cheilitis presents as erythema and fissures at the corner of the mouth and may be caused by *Candida* species, herpes simplex, or vitamin deficiency. Angular cheilitis responds to specific therapy.

There are multiple causes for oral ulcerations including HSV, CMV, and aphthous ulcers. The small, painful, clustered ulcers caused by herpes simplex can occur on any mucosal surface and are usually self-limited. Treatment with systemic acyclovir, valacyclovir, or famciclovir is indicated for nonresolving ulcers, extensive ulceration, and chronic suppression of frequent recurrences.

Cytomegalovirus ulcers are generally fewer in number and larger in diameter than herpes simplex or aphthous ulcers. They are usually associated with systemic CMV infection in patients with profound cell-mediated immune deficiency (CD4 < 50 cells/mm^3). Diagnosis is by biopsy and culture of the ulcer. Treatment is with systemic antiviral therapy, and ganciclovir, foscarnet, or cidofovir is always indicated. Because of the high rate of recurrence, long-term suppressive therapy is required unless immune function improves with use of highly antiretroviral therapy. Antiviral therapy for CMV can be safely withdrawn when patients have stable

(> 6 months) immune reconstitution with CD4 counts above 100 cells/mm^3.

Recurrent aphthous ulcers are common in persons with HIV infection at all stages of disease. The lesions range in size from 1 to 2 mm shallow ulcerations on mucosal surfaces to giant oral ulcers as large as 2 to 3 cm. Diagnosis is usually clinical, but biopsy and cultures may be indicated in some patients to exclude HSV, CMV, or lymphoma. Treatment with topical corticosteroids or in more extensive, posterior pharyngeal, or esophageal disease with systemic corticosteroids is usually effective in decreasing pain and promoting healing. Cautery with silver nitrate can provide immediate pain relief and speed healing. Thalidomide has recently been approved for use in treating aphthous ulcers.

Good oral hygiene and regular dental care are important in patients with HIV. Gingivitis and periodontal disease are more prevalent in HIV patients, particularly with advanced disease. Linear gingival erythema occurs with or without significant plaque accumulation. Periodontal disease occurs in up to 50% of patients with advanced HIV disease. The periodontal disease may be severe with resultant halitosis, regression of the gingivae, and loosening and loss of the teeth. Therapy must be aggressive and includes débridement, topical antiseptics, and systemic antibiotics.

The Epstein-Barr virus is the etiologic agent of a unique oral finding in patients with HIV, hairy leukoplakia. This condition consists of white patches usually involving the lateral borders of the tongue. It often manifests early in the course of HIV disease and may be the initial clue to underlying HIV infection. The diagnosis is clinical. Unlike candidiasis, the white plaques cannot be scraped from the mucosa. Treatment is not indicated except for cosmetic reasons. Oral and topical acyclovir have been shown to cause regression of leukoplakia, although recurrence is invariable.

Warts on mucosal surfaces caused by human papillomavirus can present as plaque-like or verrucous lesions. The diagnosis is usually clinical with biopsy confirmation. Treatment is local excision/ablation.

Larynx

The larynx is an often difficult area to evaluate and accurately diagnose. Viral (HSV, CMV) and fungal infections (histoplasmosis, coccidiomycosis, aspergillosis, and candidiasis) have been described in the larynx of HIV-infected patients.[14] Diagnosis is often made after laryngoscopy with biopsies and cultures.

Esophagus

The esophagus is a common target in patients with HIV disease, with odynophagia and/or dysphagia being the most frequent presenting complaints. *Candida* is the most common cause of these symptoms. Esophageal candidiasis is usually but not invariably associated with oral candidiasis. Other causes include HSV, CMV, and aphthous ulcerations. The diagnosis of *Candida* esophagitis can be reliably inferred by the presence of oral thrush and odynophagia/dysphagia. An empiric trial of therapy with systemic antifungal agents is indicated. If the symptoms fail to improve within several days, endoscopic evaluation is indicated to exclude viral or aphthous esophagitis or other less common conditions, including severe reflux esophagitis, Kaposi's sarcoma, lymphoma, and tubercular and MAC esophagitis. Endoscopic biopsy and cultures provide definitive diagnosis. Therapy is directed at the specific cause.

Neck

Lymphadenopathy of the head and neck is most commonly reactive secondary to HIV itself. Generalized lymphadenopathy is a finding in acute retroviral infection and is associated with fever and other systemic symptoms. Patients in the chronic phase of HIV infection may be asymptomatic, with the exception of progressive generalized lymphadenopathy. This reactive lymphadenopathy is generally symmetric and does not require biopsy for confirmation of its benign nature. Other conditions that may present with lymphadenopathy of the head and neck include malignancies such as lymphoma, Kaposi's sarcoma, squamous cell carcinoma, and a variety of infections, including those owing to *Mycobacterium tuberculosis*, MAC, *S. aureus*, cat-scratch disease, and *Treponema pallidum*. In these instances, lymphadenopathy is commonly asymmetric and associated with other findings such as fever, weight loss, cutaneous findings (Kaposi's sarcoma, impetigo), or oropharyngeal disease (lymphoma, squamous cell carcinoma).

Salivary glands, especially the parotid glands, develop cysts that are often bilateral and are called benign lymphoepithelial cysts. These lesions typically arise from the intraparotid lymph nodes rather than the parotid salivary gland tissue. Computed tomography shows multiple cystic lesions, and the diagnosis can be confirmed with fine-needle aspiration. No treatment is required if the patient is asymptomatic. Aspiration of these cysts may be warranted for cosmetic reasons. However, the fluid is likely to recur, and it is important for the patient to understand the limitation of such treatment. Surgical excision of the cysts is not recommended owing to the possibility of facial nerve damage.

Neoplasms

Kaposi's sarcoma has a predisposition for the mucous membranes as well as for the skin of the face.[15] The primary treatment for Kaposi's sarcoma has been radiation therapy. This is done at low dosage levels ranging from 1.5 to 2.5 Gy. For localized disease, other treatment options may include laser therapy, cryotherapy, and intralesional injection with chemotherapeutic agents.

Non-Hodgkin's lymphoma is the second most common malignancy seen in HIV patients. It presents in extranodal locations, usually of B-cell origin, and can be fairly aggressive with rapid onset and progression. This malignancy is treated with a combination of chemotherapy and radiation therapy.

Squamous cell carcinoma of the head and neck in the HIV-infected patient presents at a more advanced stage, has an aggressive tumor growth rate, and tends to occur in younger patients. Treatment plans should take into account the adverse response HIV patients can have to radiation therapy. Surgical management should be considered based on the underlying health and immune status of the patient.

CONCLUSION

Over 50% of patients with HIV and AIDS present with a symptom or physical finding in the head and neck region. The most common findings include otitis media, seborrheic dermatitis, chronic rhinitis, oral candidiasis, and lymphoid hyperplasia. Treatment should be individualized and requires a team approach for management.

REFERENCES

1. HIV and AIDS. In: Berkow R, Beers M, editors. The Merck manual. 17th (centennial) ed. Philadelphia: Merck and Company; 1999. p. 1312–23.
2. Fauci AS, Lane HC. HIV disease: AIDS and related disorders. In: Braunwald E, Fauci AS, Kasper DL, et al, editors. Harrison's principles of medicine. 15th ed. Philadelphia: McGraw-Hill; 2001. p. 1852–913.
3. HIV infection and AIDS. In: AIDS 101 basic information about HIV disease and AIDS. The National Institute of Allergy and Infectious Diseases; March 1999.
4. Global AIDS surveillance—part I. Wkly Epidemiol Rec 1999;74:401–4.
5. Fauci AS. The AIDS epidemic—considerations for the 21st century. N Engl J Med 1999;341:1046–50.
6. How HIV causes AIDS. In: AIDS 101 basic information about HIV disease and AIDS. The National Institute of Allergy and Infectious Diseases; March 1999.
7. Lindegren ML, Steinberg S, Byers RH. Epidemiology of HIV/AIDS in children. Pediatr Clin North Am 2000;47:1–20.
8. Nielsen K, Bryson YJ. Diagnosis of HIV infection in children. Pediatr Clin North Am 2000;47:39–63.
9. Abramowicz M. Drugs for HIV infection. Med Lett 1999;41:65–6.
10. Reisacher WR, Finn DG, Stern J, et al. Manifestations of AIDS in the head and neck. South Med J 1999; 92:684–97.
11. Moazzez AH, Alvi A. Head and neck manifestations of AIDS in adults. Am Fam Physician 1998;58: 1813–22.
12. Weymuller EA. Universal precautions in current otolaryngology practice and head and neck manifestations of AIDS. In: Gluckman JL, editor. Renewal of certification study guide in otolaryngology head and neck surgery. Dubuque (IA): Kendall/Hunt Publishing; 1998. p. 10–3.
13. Tami TA, Gold S, Singh N. AIDS and otolaryngology. In: Johnson J, editor. Infectious diseases and antimicrobial therapy of the ears, nose and throat. Philadelphia: WB Saunders; 1997. p. 571–82.
14. Tami TA, Ferlito A, Rinaldo A, et al. Laryngeal pathology in the acquired immunodeficiency syndrome: diagnostic and therapeutic dilemmas. Ann Otol Rhinol Laryngol 1999;108:214–20.
15. Singh B, Har-el G, Lucente FE. Kaposi's sarcoma of the head and neck in patients with acquired immunodeficiency syndrome. Otolaryngol Head Neck Surg 1994;111:618–24.

CHAPTER 30

Etiology of Infectious Diseases of the Upper Respiratory Tract

John L. Boone, MD

Upper respiratory infection (URI) is one of the most common ailments for which patients seek medical care. The importance of understanding the causes of these infections is evident. This chapter is divided into five sections. The first reviews pathogens of URIs according to anatomic sites. The following sections cover those viral, bacterial, fungal, and miscellaneous infections that are of special interest or importance.

CAUSES BY ANATOMIC SITE

The pathogens principally responsible for URIs are listed by anatomic site in Table 30–1. This listing is not intended to be exhaustive but is limited to causes that are predominant for each site.

Common Cold

The common cold is an acute, self-limited viral disease characterized by rhinorrhea and nasal congestion, sometimes accompanied by throat irritation, fever, and malaise. Symptoms generally begin 1 to 2 days after the onset of viral infection and peak at 2 to 4 days. Cough is associated with approximately 30% of colds and typically does not become the most bothersome symptom until the fourth or fifth day of illness, when nasal symptoms decrease. The average incidence of the common cold in preschool children is five to seven per year. The incidence decreases with age and averages two to three per year by adulthood.[1] Infection is spread by direct contact with virus-contaminated respiratory secretions. Inoculation occurs at the nasal mucosa and, perhaps, the conjunctiva.

More than 200 viruses have been associated with the common cold. The most important come from the six viral families listed in Table 30–1. Most common colds are caused by rhinoviruses of the family *Picornaviridae*, of which there are more than 100 serotypes. Makela and coworkers, using viral culture and polymerase chain reaction (PCR) techniques, detected rhinoviruses in approximately half of 200 young adults with common colds over a 10-month period.[2] Rhinovirus infections occur year round, with peaks in the early spring and fall months in the temperate zones and during the rainy periods in the tropics. Coronaviruses account for about 10% of infections. At least two serotypes cause URIs, and re-infections are common. Parainfluenza viruses and the respiratory syncytial virus are members of the *Paramyxoviridae* family. They are primarily known as causes of lower respiratory tract infections but also account for a proportion of common colds. Re-infections occur throughout life. Influenza viruses usually produce a primary influenza-like illness. However, subsequent infections with the same or a similar serotype may result in common cold symptoms. Adenoviruses are an important cause of upper and lower respiratory tract infections, including the common cold. Specific serotypes are associated with pharyngoconjunctival fever and epidemic keratoconjunctivitis. Adenoviruses are also important causes of pharyngitis and pneumonia, especially in military recruits.

Most of the information regarding the pathogenesis of the common cold is derived from studies of experimentally induced rhinovirus infections. The rhinovirus is a nonenveloped 30 nm ribonucleic acid (RNA) virus characterized by a single positive stranded genome not only acting as a template for RNA synthesis but also encoding for a single polypeptide necessary for viral replication. After inoculation, the virus invades the host by binding to the intracellular adhesion receptor molecule 1 (ICAM-1) receptors of the basal epithelial cells,

TABLE 30–1. Anatomic Sites of Infection

Infection by Site	*Causative Agents*
Common cold (upper respiratory tract)	Viral
	Rhinoviruses
	Coronaviruses
	Parainfluenza viruses
	Respiratory syncytial virus
	Influenza viruses
	Adenoviruses
Pharyngitis	Viral
	Rhinoviruses
	Coronaviruses
	Adenoviruses
	Influenza viruses
	Parainfluenza viruses
	Respiratory syncytial virus
	Coxsackieviruses
	Bacterial
	Group A β-hemolytic streptococci
	Anaerobic bacteria
	Neisseria gonorrhoeae
Acute supraglottitis (epiglottitis)	Bacterial
	Haemophilus influenzae type b
	Group A streptococci
Acute laryngotracheobronchitis	Viral
	Parainfluenza viruses
	Influenza viruses
Acute sinusitis	Bacterial
	Streptococcus pneumoniae
	Haemophilus influenzae
	Moraxella catarrhalis
Chronic sinusitis	Bacterial (mixed infections common)
	Streptococcus pneumoniae
	Haemophilus influenzae
	Moraxella catarrhalis
	Anaerobes
	Staphylococcus aureus
	α-Hemolytic streptococci
Acute otitis media	Bacterial
	Streptococcus pneumoniae
	Haemophilus influenzae
	Moraxella catarrhalis
Chronic otitis media	Bacterial (mixed infections common)
	Streptococcus pneumoniae
	Haemophilus influenzae
	Moraxella catarrhalis
	Anaerobes
	Staphylococcus aureus
	Pseudomonas aeruginosa

mainly located in the ICAM-1-rich area of the adenoid.[3] After attachment, the entire virus is translocated across the epithelial cell membrane and uncoated to release viral RNA into the cytoplasm for replication. Translation of the entire genome leads to a large polyprotein that, when cleaved, results in newly formed viral proteins. These newly formed proteins (RNA) can aggregate and eventually be released when the host cell disintegrates. After intracellular invasion and replication, the rhinovirus infection spreads intranasally and to the pharynx. Typical for rhinovirus infections are the isolated scattered foci of infected epithelium between large areas of normal epithelium. In contrast to other common cold viruses such as influenza and adenoviruses, the epithelium does not show any striking damage or cytopathic alterations during a rhinovirus infection.[3]

The mechanism(s) by which rhinovirus infection of epithelial cells in the upper airway causes the symptoms of the common cold is incompletely understood. The original concept that symptoms resulted from destruction of nasal epithelium has not been supported by biopsies, which have shown no striking damage or cytopathic alteration other than an increased number of polymorphonuclear neutrophils in the submucosa and epithelium. It appears that the viral infection per se does not produce symptoms; rather, it is the host response to the viral infection that is responsible for the symptoms.

A rhinovirus infection of the epithelium evokes the synthesis and release of mediators and cytokines, which result in a cascade of the inflammatory reactions held responsible for common cold symptoms. Vasodilatation, increased vascular permeability, cellular infiltration, and the release of various mediators characterize this response. One cytokine, interleukin-8 (IL-8), is a strong chemoattractant for neutrophils. Interleukin-8 is also capable of activating the recruited neutrophils, resulting in the release of their cytotoxic granule content. Another cytokine, IL-6, stimulates B-lymphocyte differentiation into plasma cells, with resultant production of immunoglobulins. Interleukin-6 also induces acute-phase proteins, such as C-reactive protein, and mediates T-cell activation, growth, and differentiation. Interferon-gamma (IFN-γ), another cytokine, stimulates macrophage accumulation, activation, and cytokine production. In addition, it stimulates natural killer (NK) cell function and antigen-specific B-cell proliferation. Tumor necrosis factor alpha (TNF-α) stimulates T and B lymphocytes and activates endothelial cells to express viral adhesion molecules and to release further cytokines.[3] Bradykinin, a known potent inflammatory mediator, causes vasodilatation, increases vascular permeability, and stimulates pain and glandular secretions via neuronal reflexes.

The similarity between the clinical manifestations of allergic rhinitis and the common cold has prompted attempts to establish the role of histamine in the common cold. Several studies have reported no increase in histamine in nasal secretions during rhinovirus infection. Prostaglandin D_2, another mediator derived from mast cells, cannot be detected in the nasal secretions of subjects infected with rhinovirus. Nasal biopsy studies show no change in mast cell numbers or mast cell degranulation. These studies suggest that it is unlikely that histamine or other mast cell mediators make an important contribution to the pathogenesis of rhinovirus colds.

The current concepts of the pathogenesis of symptoms in rhinovirus infection may one day allow the development of methods to block or dampen the host response. It may be that such an approach would allow the host to become immune to the infecting virus without the accompanying symptomatic illness.[4]

Pharyngitis

In most cases of acute pharyngitis, particularly mild cases, it is not possible to determine a specific microbiologic cause on clinical grounds alone. In a study of pharyngitis in adults, clinical findings such as fever, lymphadenopathy, and exudate correlated imperfectly with microbiologic findings.[5]

About 70% of acute sore throats are caused by viruses. Most of these occur as part of common cold or influenza syndromes. The most common viral agents are rhinovirus, coronavirus, adenovirus, and influenza and parainfluenza viruses.[6] Other associated viruses include respiratory syncytial virus, herpes simplex virus (types 1 and 2), coxsackievirus, Epstein-Barr virus, cytomegalovirus (CMV), and human immunodeficiency virus (HIV).

Adenoviruses are common agents of acute pharyngitis and cause a major share of pharyngitis in military recruits. Sore throat is severe and may be accompanied by cough, hoarseness, and substernal

pain. Fever, chills, malaise, myalgias, and headache are generally present. Adenoviruses are also the cause of pharyngoconjunctival fever, a syndrome characterized by fever, pharyngitis, and follicular conjunctivitis. Community epidemics of this infection most often center around contaminated public swimming pools.

Herpangina is a vesicular and ulcerative pharyngitis caused by coxsackieviruses or, less frequently, echoviruses. It is seen primarily in children 3 to 10 years old or, less commonly, in adolescents and young adults. Outbreaks are most common in the summer and fall. Onset is heralded by fever, which is followed by the development of small (1 to 2 mm) mucosal vesicles and ulcers. Over 2 to 3 days, these enlarge to 3 to 4 mm and are surrounded by erythematous rings. Lesions are most commonly noted on the anterior tonsillar pillars, as well as the soft palate, tonsils, pharyngeal walls, and posterior aspects of the buccal mucosa. Uninvolved areas of the pharynx appear normal or are only slightly erythematous. Fever and sore throat are sometimes accompanied by constitutional symptoms, including anorexia and abdominal pain, which can mimic appendicitis. Most cases of herpangina are mild and resolve without complications in 3 to 6 days.

Herpes simplex viruses (types 1 and 2) can cause acute pharyngitis indistinguishable from that caused by other respiratory viruses. More typically, vesicles and shallow ulcers accompany an infection. Herpes simplex virus is to be suspected when palatal or pharyngeal ulcers are accompanied by gingivostomatitis. Chronic herpetic lesions can occur in immunocompromised patients and are characterized by progressively large, shallow, painful ulcers.

Infectious mononucleosis most often presents with fever, pharyngotonsillitis, fatigue, and cervical lymphadenopathy. The illness is caused predominantly by the Epstein-Barr virus (EBV). Cytomegalovirus is an occasional cause of infectious mononucleosis-like disease. In such cases, pharyngitis and cervical lymphadenopathy may be mild or absent. These two viruses are addressed in more detail later in this chapter.

Primary infection with HIV may cause an acute febrile pharyngitis.[7] Fever, sore throat, and varying constitutional symptoms precede onset of lymphadenopathy by about a week.

Group A β-hemolytic streptococcus (GABHS), also known as *Streptococcus pyogenes*, is the most common bacterial cause of acute pharyngitis. Incidence is age related. In one study of school-aged children, GABHS accounted for 40% of cases of pharyngitist.[6] A prospective study of symptomatic adults patients found 10% with culture-proven GABHS pharyngitis.[8] Although GABHS pharyngitis was previously thought rare in infants, a more recent study found a 25% incidence of GABHS in children less than 3 years old.[9] As mentioned, attempts to correlate clinical findings with the presence of GABHS in acute pharyngitis have shown imperfect results. Features suggestive of GABHS, however, include pharyngeal exudate, dysphagia, headache, painful cervical lymphadenitis, fever, and the absence of cough, coryza, and hoarseness. Diagnosis may be confirmed by rapid antigen tests and/or throat culture. Timely diagnosis is important to prevent suppurative or nonsuppurative complications of GABHS infection. These complications are discussed later in this chapter.

The role of groups B, C, and G streptococci in the pathogenesis of symptomatic pharyngitis is uncertain. These non–group A streptococci may be responsible for a small portion of pharyngitis but are not associated with the nonsuppurative complications seen in GABHS.

Vincent's angina is an acute oropharyngeal ulcerative condition caused by a combination of anaerobic gram-negative and fusobacterial microorganisms. *Fusobacterium necrophorum, Treponema vincentii, Peptococcus, Bacteroides,* and other anaerobic microorganisms have been identified. Poor oral hygiene, fatigue, stress, malnutrition, and endocrine or metabolic disturbances are predisposing factors. Infection typically manifests as unilateral pseudomembranous tonsillar ulceration with pain, fetid breath, and cervical lymphadenopathy. Acute necrotizing ulcerative gingivitis (ANUG), or "trench mouth," is a similar ulcerative process of the gingiva. Penicillin and metronidazole are effective in the treatment of Vincent's angina and ANUG.

Neisseria gonorrhoeae should be considered as a cause of pharyngitis in sexually active patients, especially those with known orogenital contact. In a study of adults with gonorrhea, gonococcal pharyngitis was found in 20% of homosexual men, 10% of women, and 3% of heterosexual men. About 50% of infected individuals were asymptomatic, although odynophagia, low-grade fever, and erythema may occur. Gonococcal pharyngitis may occur without

associated urethritis and may serve as a source of disseminated disease.[10]

Acute Supraglottitis (Epiglottitis)

Acute supraglottitis is cellulitis of supraglottic structures including the epiglottis and aryepiglottic folds. Edema progresses rapidly and may result in airway obstruction.

Formerly, *Haemophilus influenzae* type b (Hib) was the etiologic agent in almost all cases of supraglottitis in children. However, with the advent of the Hib vaccine, the rates of disease caused by this microorganism have dramatically decreased.[11] Although *H. influenzae* is still identified in areas where the vaccine is available, group A streptococcus has been more frequently recovered in more recent studies.[11,12] When caused by group A streptococcus, supraglottitis was seen in an older age at onset (mean 6.2 years) and associated with slower recovery, greater involvement of the aryepiglottic folds, and negative blood cultures. This differs from the presentation of Hib supraglottitis, which includes younger age, involvement of the epiglottis, faster recovery, and positive blood cultures.[11] Other causes of acute supraglottitis in children are uncommon and include *Staphylococcus aureus*, other streptococci, *H. influenzae* non-type b, and *H. parainfluenzae*. Immunocompromised patients are more likely to have these and other, more unusual pathogens.

Acute supraglottitis is increasingly recognized in adults. In most cases, no causative microorganism is found, either by blood or throat culture. Using serology, PCR, and blood culture to investigate their series of adult epiglottitis patients, Trollfors and coworkers found Hib in about 25% and *Streptococcus pneumoniae* in 20% of patients.[13] No pathogen was identified in 43% of patients. Other pathogenic bacteria have been cultured from adult patients with supraglottitis, including *H. parainfluenzae, S. aureus,* and group A streptococci.

Acute Laryngotracheobronchitis (Croup)

Acute laryngotracheobronchitis, or croup, produces inflammation of the subglottic areas that result in stridor and respiratory distress. Most patients are children between the ages of 3 months and 3 years. The viral infection begins in the nasopharynx and, after a prodromal period of coryza and low-grade fever lasting for 12 to 72 hours, spreads to the respiratory epithelium of the larynx, trachea, and bronchi.

Parainfluenza viruses (types 1, 2, and 3) are the most frequent etiologic agents, accounting for almost 75% of isolated viral agents.[14] Influenza A virus, respiratory syncytial virus, adenoviruses, enteroviruses, and rhinoviruses are less commonly isolated. Rarely, *Mycoplasma pneumoniae* has been isolated from children with croup. Bacterial agents such as *S. pneumoniae, S. aureus*, and *Moraxella catarrhalis* have been associated with croup. They likely represent superinfection of a preceding viral disease.

Sinusitis

The microbiology of sinusitis varies according to the chronicity of the infection, as well as the age and underlying condition of the patient. Accurate culture requires aspiration or irrigation because of the presence of resident microflora in the nasal cavity and nasopharynx.

Acute community-acquired sinusitis in adults is most frequently associated with *S. pneumoniae* and nontypable *H. influenzae*. In children, to these pathogens may be added *M. catarrhalis.*[15] Anaerobic bacteria are recovered in about 5 to 10% of cases of acute sinusitis in adults, usually in the presence of dental disease. Anaerobes are rarely recovered from cases of acute sinusitis in children, presumably because of the relative absence of odontogenic infections in the young. *Staphylococcus aureus* is an uncommon cause of acute sinusitis, found in only about 3% of cases. In contrast, *S. aureus*, α-hemolytic streptococci, and *S. pneumoniae* were reported as the predominant pathogens of isolated sphenoid sinusitis. Although this may reflect increased contamination by resident nasal flora, antimicrobial coverage for *S. aureus* is prudent in patients with acute sphenoid sinusitis. Viral URIs undoubtedly play an important role in predisposing a patient to acute sinusitis; however, viruses, including rhinoviruses, adenoviruses, and influenzae and parainfluenzae viruses, are isolated in only 15 to 20% of antral aspirates taken from patients with acute sinusitis.

In studies of chronic sinusitis, the most common microorganisms identified were those seen in

acute sinusitis plus *S. aureus*, coagulase-negative *Staphylococcus*, and anaerobic bacteria. Infection is often polymicrobial, and anaerobic bacteria play a far greater role in both adults and children than in acute sinusitis.[16] Such anaerobes include *Bacteroides, Peptostreptococcus*, and *Fusobacterium*. Staphylococcal species, including *S. aureus*, are also far more frequently encountered in chronic sinusitis, as are α-hemolytic streptococci.

Nosocomial sinusitis, such as that seen following prolonged nasotracheal intubation, is commonly polymicrobial and has a high incidence of *S. aureus*, anaerobes, and gram-negative bacteria such as *Pseudomonas aeruginosa, Escherichia coli, Klebsiella pneumoniae, Proteus mirabilis,* and *Enterobacter* species.[17]

In cystic fibrosis patients with acute sinusitis, *P. aeruginosa* and *H. influenzae* were the most frequently isolated pathogens.[18]

Otitis Media

Otitis media is the most common illness diagnosed by pediatricians. Acute otitis media (AOM) is a suppurative infection of the middle ear with a relatively rapid onset. Recurrent otitis media refers to episodes of repeated AOM separated by a month or more. Chronic otitis media includes the diagnoses of otitis media with effusion, as well as chronic suppurative otitis media with and without cholesteatoma.

In children with AOM, *S. pneumoniae, H. influenzae*, and *M. catarrhalis* are the microorganisms most frequently isolated after tympanocentesis and culture of middle ear fluid (MEF). Several studies have demonstrated that *S. pneumoniae* causes about 35% of AOM episodes, nontypable *H. influenzae* 25%, and *M. catarrhalis* 15%.[19,20] Most AOM associated with *H. influenzae* is caused by nontypable strains. In about 10% of *H. influenzae* infections, the otitis is caused by type b. Such type b infections may have a higher incidence of concomitant bacteremia or meningitis and should be preventable by immunization. *Staphylococcus aureus*, other streptococci, and *P. aeruginosa* were found in less than 3% of AOM. *Chlamydia pneumoniae*, which is susceptible only to macrolides and tetracyclines, was isolated in 8% in one study of young children.[21] The most notable recent trends in the bacteriology of AOM have been the rise in the proportion of patients infected with drug-resistant *S. pneumoniae* and an overall increase in β-lactamase-producing *H. influenzae* and *M. catarrhalis*. Bluestone and colleagues also studied the bacteriology of 34 adults with AOM. *Streptococcus pneumoniae* (21%) and *H. influenzae* (26%) were again the most frequently isolated microorganisms. This study confirmed the importance of *H. influenzae* beyond the age of preschool children.[19]

Streptococcus pneumoniae is more likely than *H. influenzae* or *M. catarrhalis* to cause a more severe and persistent AOM. In the absence of therapy, *S. pneumoniae* tends to persist longer than these other microorganisms. Also, there is an increased likelihood of finding *S. pneumoniae* in AOM less responsive to antibiotic therapy. Acute otitis media caused by *S. pneumoniae* is less likely to require a preceding viral infection and more likely to induce a greater inflammatory response.[22]

Viruses play a significant role in the pathogenesis of AOM. Clinical experience shows that AOM usually occurs concurrently with, or immediately after, a viral URI. Also, the seasonal rates of AOM parallel that of viral URI. Several studies have demonstrated the effect of viral URIs in inducing eustachian tube dysfunction. The pathogenesis by which viral URIs turn into AOM or facilitate bacterial AOM is not yet clear. Heikkinen noted that in vitro studies have shown an increase in levels of inflammatory mediators (bradykinin, IL-1, IL-6, IL-8, leukotriene C4, and TNF) in nasopharyngeal secretions in response to a viral infection. Intranasal challenge of these mediators has been shown to provoke eustachian tube dysfunction in adult volunteers. Respiratory viruses have also been shown to have a suppressive effect on polymorphonuclear and cell-mediated immune function, presumably increasing the host's susceptibilty to bacterial infections.[23] Fainstein et al have shown that experimentally induced influenza virus infection increases the adherence of *S. pneumoniae* and *H. influenzae* to pharyngeal epithelial cell.[24]

Viral culture and antigen detection techniques have allowed demonstration of viruses or viral antigens in the MEF of about 20% of children with AOM. Respiratory syncytial virus, influenza viruses, parainfluenza viruses, rhinovirus, and adenovirus are most commonly identified. In about two-thirds of the cases, bacteria were also identified, indicating a mixed infection. Virus as the only pathogen in MEF has been detected in 6% of AOM cases.[24]

Recently, studies using the more sensitive PCR technique have detected RNA genomic sequences in about 50% of MEF in AOM. Heikkinen et al noted that respiratory syncytial virus is especially capable of causing inflammation in the middle ear, as well as facilitating subsequent bacterial infection.[25] Vaccination against influenza A has been shown to reduce the incidence of bacterial AOM in children during an expected influenza epidemic. It is to be expected that effective vaccines against the major viral pathogens predisposing to AOM, especially RSV, will reduce the incidence of AOM in children.[23]

Neonatal otitis media is most frequently associated with microorganisms similar to those seen in infants; however, gram-negative enteric microorganisms and *S. aureus* are more frequently identified than in older children.[26]

The bacteriology of recurrent otitis media in children with episodes separated by a month or more is essentially that of AOM.

In their 10-year review, Bluestone and associates studied the bacteriology of otitis media with effusion in children.[19] *Haemophilus influenzae* (15%), *M. catarrhalis* (10%), and *S. pneumoniae* (7%) were most frequently isolated. Other streptococci, *S. aureus*, and *P. aeruginosa* were isolated in less than 3% of cases. No growth was found in 30%. This and other reports refute previous assumptions that MEF in otitis media with effusion is sterile.

Chronic suppurative otitis media is defined by the presence of chronic infection in the middle ear and discharge through a tympanic membrane perforation or patent pressure equalization tube. Mastoiditis is invariably present. Bacteria isolated are seldom those seen in an initial AOM. Polymicrobial infections are common. *Pseudomonas aeruginosa* is most commonly isolated, and *S. aureus* is seen more frequently than in AOM.[27] Anaerobes, including anaerobic gram-positive cocci, *Bacteroides* species, and *Fusobacterium*, are present in 50% of cases.[28] In chronic suppurative otitis media with cholesteatoma, *P. aeruginosa, S. aureus*, and anaerobes again predominate.[29]

SPECIFIC VIRUSES

Epstein-Barr Virus

Epstein-Barr virus, a member of the herpesvirus family, is a double-stranded deoxyribonucleic acid (DNA) virus. Two strains have been identified in humans. Epstein-Barr is ubiquitous, infecting more than 90% of the world's population. In less industrialized societies, nearly all inhabitants are infected during childhood, when symptoms of infection are mild. In the United States, western Europe, and other industrialized countries, infection may be delayed until adolescence or young adulthood and may manifest as infectious mononucleosis because of the reaction of a more mature immune system.

Infection by EBV is largely limited to epithelial cells of the pharynx and to B lymphocytes. The surface receptor for the virus on these cells is the CD21 molecule, which is also the receptor for the CD3 component of complement. Primary infection is initiated in the oropharyngeal epithelium, where replication, host cell lysis, and release of virions take place. Adjacent B lymphocytes are infected and disseminate the virus to local and distant lymphoid tissues and to other epithelial cells, including the nasopharynx and salivary glands. Within B lymphocytes, the virus establishes a latent infection. The DNA genome circularizes to form an episome, or extrachromosomal element, in the nucleus of the cell. Transformation of the infected B lymphocyte causes them to become "immortalized," which is to say that they acquire the ability to proliferate indefinitely. Such resting memory B cells are thought to be the site of persistence, or latent infection, of EBV within the body.[30] The host's immune system responds to infection first with NK cells, then with human leukocyte antigen (HLA) class I restricted cytotoxic and suppressor T lymphocytes and by antibody-producing anti-EBV B lymphocytes. The NK and T lymphocytes appear in the peripheral blood as "atypical lymphocytes." Cytokines released from these cells are largely responsible for the clinical manifestations of primary EBV infection. In infants and small children, these immune responses are weak, and symptoms of infection are mild. The immune systems of adolescents and young adults respond more vigorously and cause the symptoms of infectious mononucleosis.

Epstein-Barr virus is the predominant cause of infectious mononucleosis, which is characterized by fever, fatigue, pharyngitis, and lymphadenopathy. Splenomegaly, mild hepatitis, and cerebritis may also occur. Uncommon manifestations include jaundice, pneumonitis, blood dyscrasias, and central nervous

system syndromes. The diagnosis is suggested by the clinical features. Abnormal laboratory findings include the presence of atypical lymphocytes in the peripheral blood, elevation of hepatic enzymes, and the presence of positive heterophil antibodies (the "Monospot" or Paul-Bunnell test). Heterophil antibodies are immunoglobulin (Ig)M antibodies that bind to red blood cells from nonhuman species such as sheep or horses. Many rapid test kits for their detection are commercially available. The percentage of patients with acute EBV infection who have heterophil antibodies in their serum increases relative to the time since onset of symptoms. The percentage is also higher in patients over 4 years of age. In one study of young adults, 69% tested positive for heterophil antibodies by 1 week and 80% by 3 weeks after the onset of symptoms. A pediatric study found fewer than 50% of children under age 4 to have any detectable heterophil antibodies at any time.[31]

Epstein-Barr virus is closely linked to nonkeratinizing nasopharyngeal carcinoma and to Burkitt's lymphoma, a tumor of B lymphocytes limited largely to children in tropical Africa and New Guinea. Epstein-Barr virus genomes are also found in about one-third of the non-Hodgkin's B-cell lymphoma found in patients with acquired immune deficiency syndrome (AIDS).

Oral hairy leukoplakia is common in patients with HIV and represents infection by EBV of epithelial cells. It consists of white lesions with a hairy or corrugated appearance located on the lateral surface of the tongue. Epstein-Barr virus replication is evident only in the upper layers of the epithelium.

Cytomegalovirus

Cytomegalovirus, the largest member of the herpesvirus family, consists of a core of double-stranded DNA genome surrounded by a protein coat. The virus is ubiquitous, and serologic evidence suggests that up to 80% of adults have been infected. Most primary infections in immunocompetent persons are asymptomatic. An upsurge in interest in CMV has resulted from an appreciation of its role as an opportunistic pathogen in immunocompromised patients, including those with HIV infection.

The details of pathogenesis in CMV infection are less well understood than for EBV. Infection may be acquired by contact with the saliva, sputum, urine, or genital fluid of infected persons; through blood transfusion or solid organ transplantation; or by transplacental transfer to a fetus. Ductal epithelial cells of the major salivary glands and proximal renal tubules are especially involved by primary infection. Once acquired, the virus infects nucleated blood cells and spreads systemically. After primary infection, CMV generally remains in an asymptomatic latent state, but immunosuppression may result in reactivation and disease.

In infants and children, CMV infection generally causes few symptoms. In young adults, however, CMV can produce a mononucleosis syndrome with fever, fatigue, pharyngitis, lymphadenopathy, and relative lymphocytosis. It is estimated that 20% of infectious mononucleosis is caused by acute CMV infection. Pharyngitis, lymphadenopathy, and splenomegaly are less common and severe than in cases caused by EBV. Patients with CMV mononucleosis will show a relative lymphocytosis but will have a negative heterophil antibody (Paul-Bunnell) test.

Primary or reactivated CMV infections commonly cause disease in immunosuppressed hosts, such as persons with HIV infection or those who have undergone bone marrow or solid organ transplantation. Cytomegalovirus itself may adversely affect the immune system, causing suppression of NK-cell and T-lymphocyte function. Clinical features in immunosuppressed patients include encephalitis, chorioretinitis, hepatitis, pneumonitis, colitis, and esophagitis. Cytomegalovirus has also been identified as a cause of oral ulcers in immunosuppressed patients. Diagnosis in these cases in based on tissue biopsy, which demonstrates the characteristic histologic findings of intranuclear CMV inclusions surrounded by clear halos ("owl eyes"), as well as clusters of intracytoplasmic inclusions seen with periodic acid–Schiff or Gomori's methenamine silver techniques.

Cytomegalovirus is among the most common viral pathogens causing congenital abnormalities. Fetal involvement follows a primary CMV infection of a woman lacking antibody to the virus during pregnancy. Clinical features include mental retardation, hearing loss, jaundice, chorioretinitis, microcephaly, and pneumonitis, among others.

Human Immunodeficiency Virus

Human immunodeficiency virus is a retrovirus. It is able to convert its own single-stranded RNA to dou-

ble-stranded DNA for incorporation into a host cell genome. The principal targets are T helper (CD4) cells, which are central to the function of cell-mediated immunity. Impairment of these cells renders a patient susceptible to a variety of opportunistic infections and unusual malignant diseases.

Human immunodeficiency virus attaches to CD4 receptors and enters T-helper cell lymphocytes. The single-stranded RNA is converted by the enzyme reverse transcriptase into double-stranded DNA. The viral DNA is incorporated into the host cell genome. A latent phase follows that corresponds to clinically quiescent infection. Activation and viral replication lead to cell death and the destruction of large numbers of CD4 cells. It is possible that activation is speeded by the presence of DNA viruses such as herpes simplex virus and CMV. The presence or absence of opportunistic infections corresponds with the patient's CD4 count. A normal count is 700 to 1,200 cells/mm^3. CD4 counts below 200 cells/mm3 are associated with increased susceptibility to opportunistic infections. Such patients benefit from antibiotic prophylaxis.

Humoral immunity is adversely affected by HIV as well. Immunoglobulin G2 and IgG4 subclass deficiencies predispose patients to recurrent sinopulmonary infections. Patients with HIV may also develop elevated IgE levels, which may be responsible for an increased incidence of allergic rhinitis.

The reported prevalence of sinusitis in HIV patients varies widely, although clinicians generally agree that the incidence is higher than in immunocompetent patients. Retrospective studies have estimated the prevalence of disease at 10 to 20%. Two prospective studies placed the incidence at 30 to 68%.[32] Several reports suggest that the incidence of sinus disease increases with progressive immunodefiency, as measured by declining CD4 counts.[33,34] In addition, IgE levels increase in many HIV-infected patients with disease progression. Such immune dysfunction, with the resultant increase in allergic disease, may predispose HIV-infected patients to recurrent rhinosinusitis. The prevalence of specific pathogens in sinusitis of those HIV-infected patients with a CD4 count of greater than 200 cells/mm^3 is similar to that of non–HIV-infected patients. In acute sinusitis, *S. pneumoniae* and *H. influenzae* are most frequently recovered. In chronic sinusitis, *S. aureus*, anaerobic bacteria, and *P. aeruginosa* are also described.[35] Patients with HIV with CD4 counts less than 200 cells/mm^3 are more susceptible to unusual pathogens. Infections with *P. aeruginosa* are more frequent. Other pathogens reported include CMV, *Pneumocystis carinii, Acanthamoeba castellanii*, and *Legionella pneumophila.* Patients with AIDS are at increased risk for invasive fungal sinusitis. *Aspergillus fumigatus* has been most frequently reported. Other fungal pathogens have included *Pseudallescheria boydii, Candida albicans, Cryptococcus neoformans, Rhizopus arrhizus*, and mucormycosis, among others.[36]

Studies have suggested that children with HIV infection have an increased incidence of otitis media. Moreover, those HIV-infected children with a deteriorating immune system, as indicated by a decline in the CD4 count, are at greater risk for recurrent or chronic disease than those who retain an intact immune system.[37] Causative microorganisms do not seem to differ from those found in immunocompetent hosts. *Streptococcus pneumoniae, H. influenzae*, and *M. catarrhalis* predominate. Unusual pathogens, however, should be suspected when a patient has a severely depressed immune system or responds poorly to standard antibiotic therapy.

Adult HIV-infected patients appear to have an increased incidence of serous otitis media. This may be partly attributable to poor eustachian tube function. The eustachian tube may be affected by adenoidal hypertrophy, nasopharyngeal tumors such as Kaposi's sarcoma and non-Hodgkin's lymphoma, or allergic rhinitis prompted by increased IgE levels. *Pneumocystis carinii* is a common cause of pneumonia in AIDS patients. Although extrapulmonary infections are rare, several reports have described otitic involvement. Patients present with aural discharge and the presence of a polyp in the middle ear or external auditory canal. Several patients have had no history of prior *P. carinii* pneumonia. Diagnosis is made by biopsy of the polyp and demonstration of pneumocystis on Gomori's methenamine silver stain. These aural infections have responded well to trimethoprim-sulfamethoxazole. Other unusual pathogens have caused otitis media or mastoiditis, including *A. fumigatus, Candida albicans*, and *Mycobacterium tuberculosis*. Chronic otitis media, cholesteatoma, and intracranial complications of otitis media do not appear to occur more often in the HIV-infected population than in other patients.[38]

Oral candidiasis is a common feature of HIV infection and is often the first finding of the disease. Candidal infections are considered in more detail later in this chapter. As noted earlier, oral hairy leukoplakia is thought to be caused by EBV. The lesion is most frequently located on the lateral surface of the tongue and has a white corrugated or filiform appearance. Epstein-Barr virus replication is evident only in the upper layer of the epithelium. Herpes simplex viral infections are more common in HIV-infected patients. Lesions are generally more severe, numerous, and persistent than in immunocompetent patients. Herpes labialis is frequent and may extend onto the face to form a giant herpetic lesion.

Opportunistic infections of the larynx must be considered in HIV-infected patients with hoarseness or stridor. Candidiasis is the most common cause. Laryngeal candidiasis is most often part of a more widespread infection involving the oral cavity or esophagus. Absence of candidiasis in these sites, however, does not exclude laryngeal infection. The vocal folds are most frequently involved. Candidal infection of the epiglottis is much less common but is capable of causing stridor and airway compromise, especially in children.[39,40] Laryngeal infections caused by CMV and *Cryptococcus neoformans*, among others, have been described in HIV-infected patients. A series of five HIV-infected patients with supraglottitis was reported by Rothstein and associates.[41] A pale, floppy epiglottis with supraglottic edema, cervical lymphadenopathy, normal to low white blood cell count without a shift to the left, and rapidly progressive airway obstruction comprised the typical presentation. No unusual pathogens were recovered from epiglottic cultures in these patients. The most common pathogens were *S. pneumoniae* and *S. aureus*. Conservative medical management was not successful. Patients required airway intervention with appropriate antibiotic therapy.

SPECIFIC BACTERIA

Group A β-Hemolytic Streptococcus (Streptococcus pyogenes)

Group A β-hemolytic streptococcus, or *S. pyogenes*, is a chain-forming, gram-positive coccus. More than 80 serotypes have been recognized based on a series of distinct surface proteins, the M proteins. The M protein strains of GABHS that cause pharyngitis are generally distinct from those causing impetigo or pyoderma. M proteins inhibit phagocytosis of GABHS by host leukocytes in the absence of type-specific antibodies and therefore play an important role in the virulence of these microorganisms.[42]

Group A β-hemolytic streptococcus secretes certain toxins and enzymes that have clinical effects or cause antibodies to be produced that may be used for serologic identification. Streptolysins O and S induce lysis of host cell membanes, including red blood cells. Streptolysin O is antigenic and is the basis of the antistreptolysin O (ASO) antibody assay, the most widely used of the many streptococcal antibody tests. Pyogenic (also called erythrogenic) exotoxins are produced by some serotypes of GABHS. Three exotoxins have been identified: streptococcal pyogenic exotoxin A (SPE A), SPE B, and SPE C. These exotoxins are thought to be in part responsible for the clinical manifestations of scarlet fever and of streptococcal toxic shock-like syndrome (TSLS).

The leading role of GABHS as a cause of pharyngitis is discussed earlier in this chapter. The greater significance of these microorganisms lies in their capacity to cause complications. These complications may be broadly separated into three categories: suppurative, toxin mediated, and nonsuppurative.

Suppurative complications of GABHS include peritonsillar abscess, retropharyngeal abscess, suppurative cervical lymphadenitis, sinusitis, otitis media, and mastoiditis. Group A β-hemolytic streptococcus is a prominent cause of peritonsillar and retropharyngeal abscesses. Similarly, GABHS and *S. aureus* are the most important causes of suppurative cervical lymphadenitis. In contrast, GABHS plays a minor role in the pathogenesis of sinusitis, otitis media, and mastoiditis.

Toxin-mediated complications of GABHS pharyngitis include scarlet fever and streptococcal TSLS. Both are thought to be associated with the release of streptococcal pyogenic exotoxins by certain M protein serotypes. In the 1980s, an increase in the incidence and severity of scarlet fever, as well as the emergence of TSLS, suggested an increase in the number of virulent GABHS strains capable of producing pyogenic exotoxins. Serotypes M-1 and M-3, in particular, were found to be associated with TSLS.[43] Investigators have postulated that manifestations of scarlet fever, including the characteristic

rash, are attributable to host hypersensitivity to these pyogenic exotoxins.

Initial manifestations of scarlet fever include sore throat, fever, and constitutional symptoms such as vomiting, headache, and malaise. The characteristic rash follows 12 to 48 hours later, beginning on the trunk before quickly becoming generalized. The rash appears diffusely erythematous with fine red papules. It blanches on pressure, may become petechial, and fades in 2 to 5 days, leaving a fine desquamation. Skin folds, such as the groin and axilla, are most intensely involved. The face is flushed, with circumoral pallor. Enlargement and erythema of the papillae of the dorsal tongue produce a characteristic "strawberry tongue."

In the late 1980s, reports emerged of GABHS infections associated with hypotension, progressive multisystem organ failure, and other features suggestive of a toxin-mediated disorder. Because of a partial resemblance to staphylococcal toxic shock syndrome, the new disorder was named streptococcal TSLS. The syndrome occurs most often in healthy adults. Most patients have a severe soft tissue focus of infection, such as necrotizing fasciitis and/or myositis; however, 10 to 20% of cases have been associated with pharyngitis.[43] The syndrome most typically consists of hypotension, fever, rash, desquamation, and multisystem organ failure, which is out of proportion to the extent of local signs and symptoms. Bacteremia is present in about half of the cases. In a reported series of 50 patients, the mortality rate was 24%. Group A β-hemolytic streptococcus isolated from patients with TSLS appears to come from a limited group of M protein serotypes. M-1 and M-3 strains, in particular, have been associated with TSLS. Streptococcal pyogenic exotoxin A stimulates production of TNF by peripheral monocytes in vitro. A cytokine, IL-1β, is thought to contribute to the pathophysiology of shock and tissue injury. Treatment of streptococcal TSLS involves support of failing organ systems, appropriate antibiotics, and, if present, aggressive surgical débridement of necrotizing fasciitis or myositis.

Nonsuppurative complications of GABHS pharyngitis include poststreptococcal glomerulonephritis and acute rheumatic fever (ARF). The pathogenesis of these conditions is unclear, but each is suspected to be immune mediated.

Poststreptococcal acute glomerulonephritis (AGN) is a delayed, nonsuppurative complication resulting in inflammation of the renal glomeruli. Acute glomerulonephritis may follow either pharyngitis or cutaneous infections caused by GABHS. A limited number of strains are "nephritogenic." M-12 is the M protein serotype most frequently related to pharyngitis-associated nephritis.[42] The pathogenesis of AGN is unclear. Antibodies to antigens of nephritogenic streptococci may cross-react with structurally similar renal antigens, producing injury. Alternatively, inflammation may result from deposition of streptococcal antigen-antibody complexes within the kidney. Acute glomerulonephritis associated with GABHS pharyngitis occurs most frequently in school-aged children. The latent period following pharyngitis averages 10 days. Clinical features include fever, moderate edema, hypertension, and azotemia. Mild cases may have few symptoms. The urine is smoky or brown in color with a high specific gravity, proteinuria, and hemoglobinuria. Microscopic examination reveals red blood cells, white blood cells, and hyaline, granular, or red blood cell casts. Other laboratory findings include a mild normocytic normochromic anemia, prolonged erythrocyte sedimentation rate (ESR), decreases in C3 complement, and elevations of the blood urea nitrogen and serum creatinine. Unlike in ARF, recurrences are rare. The long-term prognosis in children is excellent, with few cases leading to residual renal disease. Prognosis in adults is more guarded.

Acute rheumatic fever is a nonsuppurative complication of GABHS characterized by inflammatory lesions of the heart, joints, subcutaneous tissues, and central nervous system. Pyodermic infections are not associated with ARF. The disease has become uncommon in more prosperous areas of the world, although local outbreaks thought to be caused by highly rheumatogenic strains of GABHS are periodically reported. In poorer regions, ARF remains prevalent and accounts for a large proportion of cardiac disease in children and young adults.[44] The pathogenesis of ARF is poorly understood. Current evidence supports an immune mechanism. Humor- and cell-mediated immune re- sponses to rheumatogenic antigens of GABHS are thought to cross-react mistakenly with structurally similar antigens in host tissues. Pathologic findings of ARF include inflammatory lesions of connective tissues in the heart, joints, blood vessels, and subcutaneous tissues. A unique feature, seen mainly in the heart, is the Aschoff's nodule, a granuloma that results from

injury to collagen. Cardiac findings also include pericarditis, myocarditis, and/or endocarditis. Valvular involvement begins as edema and inflammation of the leaflets and chorda tendinae. With healing, the valves may become thickened and deformed, the chordae shortened, and the valve commissures fused. These changes result in valvular stenosis or insufficiency. The mitral valve is affected in 75 to 80% of cases, the aortic valve in 30%, and the tricuspid and pulmonary valves in under 5%. Acute rheumatic fever occurs most often in children 5 to 15 years of age; it is rare before age 4 or after age 40. Symptoms and signs usually commence 1 to 3 weeks after the streptococcal pharyngitis. No single symptom, sign, or laboratory test is diagnostic of the disease. The Jones criteria, most recently updated in 1992, remain useful in making the diagnosis.[45] If supported by evidence of preceding group A streptococcal infection, the presence of two major manifestations, or one major and two minor manifestations, a high probability of ARF is indicated. Supporting evidence includes a positive throat culture, positive rapid streptococcal antigen test, or an elevated or rising streptococcal antibody titer. Major manifestations include carditis, polyarthritis, chorea, erythema marginatum, and subcutaneous nodules. Minor manifestations include clinical findings such as arthralgia or fever, laboratory findings such as an elevated ESR or C-reactive protein, and a prolonged PR interval on an electrocardiogram. Treatment of ARF includes the use of salicylates and corticosteroids to quiet inflammation. Neither agent prevents the development of rheumatic heart disease, the only long-term sequelae of ARF. Rheumatic patients are at extremely high risk of developing recurrent ARF after streptococcal infections and require antibiotic prophylaxis.

Corynebacterium diphtheriae

The corynebacteria are nonmotile, pleomorphic, unencapsulated, gram-positive bacilli. Diphtheria is caused by toxin-producing strains of *Corynebacterium diphtheriae* or, occasionally, the related microorganism *Corynbacterium ulcerans*. The infection is now rare in countries with immunization programs but remains widespread elsewhere. In 1990, an epidemic centered in the former Soviet Union demonstrated the potential for resurgence of this disease should efforts to maintain high levels of immunity in a population lapse.

Humans are the only known reservoir for *C. diphtheriae*. Transmission occurs via contact with respiratory secretions or exudate from infected skin lesions. The virulence of *C. diphtheriae* results from the action of a potent exotoxin capable of inhibiting intracellular protein synthesis. After host cell uptake, the exotoxin catalyzes inactivation of the transfer RNA translocase elongation factor 2. Loss of this enzyme prevents the interaction of messenger RNA and transfer RNA, stopping further addition of amino acids to developing polypeptide chains. Local action of the exotoxin in the upper respiratory tract induces a characteristic pseudomembrane comprised of fibrin, leukocytes, erythrocytes, dead respiratory cells, and microorganisms. Systemically, the exotoxin has a predilection for the heart (myocarditis), nerves (demyelination), and kidneys (tubular necrosis). Some strains of *C. diphtheriae* do not produce exotoxin. Such strains will not produce severe systemic symptoms or the characteristic pseudomembrane.

The posterior structures of the mouth and the proximal parts of the pharynx are the most common sites for respiratory tract diphtheria. Onset is abrupt and involves sore throat, malaise, and low-grade fever. Initial erythema usually progresses to development of an adherent pseudomembrane typically on one or both tonsils with extension to the soft palate, oropharynx, and nasopharynx. Its color may vary from creamy to grayish green. The pseudomembrane is surrounded by a deep red border, and attempts at removal result in bleeding. Cervical lymph nodes are usually enlarged, and severely ill patients may have a bull-neck appearance. Non–exotoxin-producing strains of *C. diphtheriae* will not produce a pseudomembrane. Infections produced by such strains may closely resemble streptococcal pharyngitis.

Laryngeal and tracheobronchial involvement usually results from direct spread from the pharynx, although the disease may occasionally begin there. Symptoms include hoarseness, respiratory distress, and a brassy cough. If the presence of a laryngeal or tracheal membrane is found on laryngoscopy, consideration must be given to early endotracheal intubation or tracheostomy to forestall airway obstruction, especially in children.

Nasal diphtheria initially causes purulent or serosanguinous nasal discharge. Patients with chronic nasal diphtheria may act as carriers of the disease.

Diphtheria exotoxin has its most striking effects on the heart and nervous system. Clinically significant myocarditis occurs in 7 to 20% of respiratory tract diphtheria cases. Neuropathy can involve cranial nerves IX and X, causing palatal or pharyngeal paralysis with resultant regurgitation of swallowed fluids through the nose.

Diphtheria must be suspected in any unimmunized patient with a severe, rapidly spreading pharyngeal exudate. Diagnosis is confirmed by culture on Loeffler's or Tindale's medium. Because routine methods of throat culture do not promote the isolation and identification of *C. diphtheriae*, the laboratory must be alerted to use selective media when the disease is suspected. Toxin production is demonstrated by immunodiffusion on solid culture media (Elek test), by monoclonal enzyme immunoassay for toxin (EIA), or by PCR testing of the isolate's DNA, using primers for the *tox* gene.

Treatment of diphtheria involves the use of antitoxin and antibiotics and protection of a compromised airway. Diphtheria antitoxin only neutralizes the exotoxin before its entry into cells, so it is critical that it be administered as soon as a presumptive diagnosis has been made, without awaiting diagnostic confirmation. Penicillin and erythromycin are the antibiotics of choice. By eradicating the pathogen, they end local disease, further production of exotoxin, and transmission of disease.

Klebsiella rhinoscleromatis (Rhinoscleroma, Scleroma)

Klebsiella rhinoscleromatis (*K. pneumoniae*, subspecies *rhinoscleromatis*) is a capsulated gram-negative coccobacillus that causes a chronic granulomatous disease of the upper respiratory tract. It is endemic to tropical and subtropical areas of Africa, Central and South America, and Asia. Most European cases are reported from Eastern Europe. In temperate zones, cases are generally imported from endemic areas. Poor hygiene, crowded living, and malnutrition are predisposing factors. Females are more frequently affected than males. The disease tends to present in the second and third decades of life. The nose is almost always involved. Less commonly, the paranasal sinuses, lacrimal sacs, pharynx, larynx, and trachea are affected. Nasal lesions often begin at the mucocutaneous junction of the vestibule. The disease usually progresses through three stages. The catarrhal stage is manifested by prolonged purulent rhinorrhea, crusting, and nasal obstruction. Diagnosis in this stage is difficult because of the nonspecific findings. In the granulomatous stage, multiple nodules form and coalesce to form bluish-red or pale, rubbery granulomas.[46] Most cases are diagnosed in this stage. Severe cases can progress to local destruction and cosmetic deformity. Broadening of the nasal dorsum produces the characteristic Hebra nose. The final fibrotic stage causes cicatrix formation. Nasal or nasopharyngeal stenosis is common. Stenosis of the larynx or trachea can cause life-threatening airway obstruction. Diagnosis of rhinoscleroma is made by clinical findings and biopsy. Histologic findings include granulomatous inflammation and the presence of Mikulicz's cells and Russell's bodies. These findings are most characteristic of the granulomatous stage, in which *K. rhinoscleromatis* is most easily isolated. Mikulicz's cells are "foamy histiocytes" caused by gram-negative bacteria in the cytoplasm of macrophages. Russell's bodies are eosinophilic structures within the cytoplasm of plasma cells. In advanced cases, fibrosis may obscure characteristic findings. Treatment of rhinoscleroma is with prolonged antibiotics. The most commonly used antibiotics are tetracycline (2 g/day), ciprofloxacin (1 g/day), and streptomycin (1 g/day). The criteria for cessation of therapy are clinical improvement and repeated negative tissue culture. Surgical débridement may be indicated to alleviate airway obstruction. Laryngeal dilatation, endoscopic resection, or tracheostomy may be required for laryngeal stenosis.

Mycobacterium tuberculosis

Mycobacterium tuberculosis, an acid-fast bacillus, is a cause of infection throughout the upper respiratory tract. Its incidence has again increased with the advent of AIDS. Two sites of special interest are the larynx and middle ear.

Tuberculosis is the most common cause of granulomatous disease of the larynx. It is usually, but not invariably, associated with active pulmonary disease. Spread from the lung to larynx is largely ascribed to direct inoculation by bacilli-laden sputum, but hematogenous and lymphatic spread from pulmonary foci is recognized as well. Symptoms of laryngeal tuberculosis may include hoarseness, odynophagia, dysphagia, and referred otalgia. Exam-

ination may show edema, submucosal nodules, and shallow ulcerations. Lesions may progress to deeper ulceration, perichondritis, and arytenoid fixation. The posterior part of the larynx is most commonly involved, although any portion of the larynx may be involved.[47] Laryngeal tuberculosis may mimic carcinoma or chronic laryngitis. Syphilis, mycotic infections, sarcoidosis, and Wegener's granulomatosis must also be considered. A chest radiograph is helpful in suggesting the diagnosis if active pulmonary disease is present. A sputum smear may show acid-fast bacilli. The diagnosis may be confirmed by biopsy. Laryngeal tuberculosis responds well to antitubercular chemotherapy. Fibrotic changes from advance disease may require surgical reconstruction.

Tuberculous otitis media may result from either hematogenous spread from pulmonary foci or from direct inoculation by bacilli-laden sputum refluxed through the eustachian tube. Otorrhea may be scant or profuse. Although the disorder is often painless, many patients experience otalgia. Granulation tissue and occasionally polyps are seen in the middle ear. Conductive hearing loss may result from effusion, mucosal thickening, tympanic membrane perforation, or ossicular destruction. Classically, tympanic membrane perforations have been described as initial small multiple perforations that coalesce to a near-total perforation. More recent studies show that simple perforations are more common, however. Denudation and destruction of the ossicular chain are common. Facial nerve paralysis is more common in aural tuberculosis than in otitis media caused by other bacteria. Tuberculous otitis media can also cause mastoiditis, which, in turn, may progress to subperiosteal abscess and a draining fistula. Other complications may include labyrinthitis, meningitis, and osteomyelitis of the petrous pyramid. The diagnosis of aural tuberculosis is best made by biopsy of granulation tissue or polyps. Histologic presence of caseating granulomata and acid-fast bacilli confirms the diagnosis. Smear and culture of aural discharge are unreliable because the bacilli can be demonstrated in only a minority of patients. Chest radiograph abnormalities suggestive of pulmonary disease are present in only 50% of patients with aural tuberculosis.

Treatment of aural tuberculosis consists of long-standing antitubercular chemotherapy. Surgical intervention may be indicated in the presence of facial nerve paralysis, subperiosteal abscess, persistent postauricular fistula, labyrinthitis, or extension of the infection into the central nervous system. Cortical mastoidectomy may disclose thick granulation tissue and possible bony necrosis. Tympanoplasty or ossicular reconstruction should be delayed until the infection has been controlled.

Mycobacterium leprae

Mycobacterium leprae is a gram-positive, acid-fast bacillus. It prefers an ambient temperature of less than 37°C and, as a result, has a predilection for skin, peripheral nerves, and the upper aerodigestive tract. Leprosy presents in three general forms. In the lepromatous form, the patient is unable to generate an effective immune response to *M. leprae* because of a selective unresponsiveness of T lymphocytes. This form of leprosy tends to be the most widespread and destructive. Most upper aerodigestive tract leprosy is of the lepromatous form. In the tuberculoid form, the host has a normal T-cell response. Tuberculoid form lesions are generally less widespread. They consist of plaques and macules with sharp, raised borders. An intermediate third form of leprosy is the borderline form, which reflects a partial reduction in immune response.

The nose is by far the most frequently involved site in the upper aerodigestive tract. Approximately 95% of patients with lepromatous leprosy have nasal involvement. Lesions begin as a pale, nodular, plaque-like thickening of the nasal mucosa. Untreated lesions can progress to ulceration, septal perforation, and saddle-nose deformity. Symptoms include rhinitis, obstruction, crusting, and bleeding. The anterior end of the inferior turbinate is the site most likely to give a positive biopsy.

The larynx is less commonly involved. Involvement nearly always begins at the tip of the epiglottis. Mucosal thickening is followed by granuloma formation and ulceration. Symptoms may include hoarseness and pain. Laryngeal leprosy can closely mimic carcinoma. Untreated lesions may progress to fibrosis and airway obstruction, requiring tracheostomy.

Mycobacterium leprae has never been cultured in vitro. Diagnosis is most reliably made by tissue smear or biopsy, which may show granulomatous changes and the presence of bacilli. Rifampin, dapsone, and clofazimine are the most widely used

antileprosy drugs. Treatment is generally with the standard World Health Organization regimen.[48]

Actinomyces

Actinomycosis is caused by anaerobic gram-positive bacterium of the family *Actinomycetaceae*. The filamentous, branching appearance of these oral saprophytes on microscopy caused their initial misidentification as fungi. *Actinomyces israelii* is the microorganism most frequently incriminated in human disease. These inhabitants of the normal oral flora take advantage of infection, trauma, or surgical injury to penetrate mucosa and invade adjacent tissue. Cervicofacial infections commonly follow dental infections or manipulations. Such infections most commonly affect otherwise healthy individuals.

The clinical manifestations of cervicofacial actinomycosis may vary from an indolent infection to a more rapidly progressive, tender, fluctuant swelling resembling a typical pyogenic infection.[49] Most cases, however, present as a chronic infection with few systemic symptoms. A typical presentation shows a slowly progressive, mildly tender induration of the neck, buccal, and mandibular regions. Infection spreads without regard to tissue planes. Progression leads to formation of a hard, board-like lesion ("lumpy jaw"), suppuration, draining fistulae, and the presence of "sulfur granules" (white-yellow granules of bacterial filaments, 1 to 5 mm in diameter) in exudates or tissues. Overlying skin can be dark red to purple, and associated cutaneous sinus tracts remain chronic, with little tendency to heal. Osseous involvement of the mandible occurs in 15% of cases.

A diagnosis of actinomycosis is based on the clinical presentation, positive culturing of the microorganism, and observation of sulfur granules or acid-fast–negative pleomorphic filaments in exudate or tissue sections.[50] Anaerobic culturing must be meticulous to obtain growth. Even with proper precautions, actinomycosis is identified from culture in fewer than 50% of cases. Fine-needle aspiration has been shown to be useful in obtaining specimens for histologic examination and culture.

Treatment of cervicofacial actinomycosis consists of prolonged antibiotics and surgical débridement. Although sensitive to penicillin, nearly all cases require surgical drainage, curettage, or lesion resection.[49] The duration of penicillin therapy will vary from 2 to 18 months. Concomitant hyperbaric oxygen treatment in severe cases should be considered as well.

SPECIFIC FUNGAL INFECTIONS

Candidiasis

Candida albicans is the most frequent of several species that cause candidiasis. It is a dimorphic fungus with both yeast and hyphal forms. *Candida albicans* is a commensal of the oral cavity and pharynx. Infection usually only follows changes in the local bacterial flora, mucosal integrity, or host immunity. Local factors that may predispose patients to candidiasis include xerostomia and dentures. Systemic factors are numerous and include prolonged antibiotic use, diabetes mellitus, immunocompromised states (eg, HIV, cytotoxic therapy), pernicious anemia, and hematologic malignant diseases.

Oral candidiasis ("thrush") can vary greatly in appearance. Zegarelli has described several variants.[51] The pseudomembranous variant resembles white milk curds overlying an erythematous base. The erythematous variant appears flat, irregular, and dusky red. When candidiasis occurs beneath dentures, it is flat and red and extends up to but not beyond the denture border. A hyperplastic variant resembling lichen planus manifests as a true leukoplakia without pseudomembrane or erythematous base. Oral candidiasis is often asymptomatic but may also cause a "burning" pain. Candidiasis may also be responsible for median rhomboid glossitis, as well as some cases of angular cheilitis.

Candidiasis less commonly infects the larynx and pharynx. When involved, the larynx is covered by a thick, white exudate. Hoarseness is the most common symptom, followed by pharyngeal pain. Severe cases may proceed to airway obstruction and laryngeal scarring. Diagnosis is best made by direct laryngoscopy and biopsy.

Aspergillosis

The *Aspergillus* species are ubiquitous saprophytic fungi found in soil and decaying organic material. *Aspergillus fumigatus* and *Aspergillus flavus* are the most common of many species to cause disease. Hyphae show regular septa with frequent acute

branching. *Aspergillus* infections are distributed worldwide, although chronic disease is especially prevalent in the Sudan and Saudi Arabia. Infection is acquired by the inhalation of spores. The lungs, nose, and paranasal sinuses are the most commonly affected sites.

Aspergillus is the most common cause of fungal sinusitis. The maxillary and ethmoid sinuses are most frequently involved. Conditions that obstruct sinus drainage such as allergic rhinitis, nasal polyps, and chronic bacterial sinusitis may predispose patients to fungal sinusitis. Infections take one of four forms. Acute invasive fungal sinusitis occurs primarily in immunocompromised patients. Vascular invasion by hyphae with resultant thrombosis and necrosis may cause a fulminant course resembling craniofacial mucormycosis. Erosion into the orbits, cranium, palate, and skin is common. Chronic invasive fungal sinusitis occurs more frequently in immunocompetent patients. Progression of the infection is generally slow, with mild early signs and symptoms. Extension into the orbit or cranium may produce proptosis, ophthalmoplegia, and headache. The third form of fungal sinusitis is the "fungus ball." This is a tangle of fungal hyphae within a sinus cavity but without invasion of sinus mucosa. The maxillary sinus is most commonly involved. A fourth form of fungal sinusitis is allergic fungal sinusitis (AFS). Most patients with AFS are immunocompetent and present with features of refractory sinusitis and nasal polyposis. Many also have a history of atopy or asthma. The pathogenesis of AFS is incompletely understood, but evidence supports a Gell and Coombs type I (IgE-mediated) immune response to fungal antigens. Type III and IV immune responses may also be involved. The pathologic hallmark of AFS is "allergic mucin," consisting of eosinophils, Charcot-Leyden crystals, and scattered fungal hyphae. Although AFS is not thought to be invasive, bony erosion is seen on computed tomographic (CT) scans in 20% of reported cases. Laboratory abnormalities include eosinophilia, elevated total serum IgE, and elevated IgE radioallergosorbent test (RAST) titers to specific fungi. The diagnosis of AFS is confirmed by demonstration of allergic mucin and culture of the fungus.

The presence of *Aspergillus* or other fungi must be suspected in cases of sinusitis that do not respond to the usual medical treatment. Computed tomographic scans may show a characteristic pattern of alternating high- and low-density signals and may also be useful in determining the extent of the disease. Often the proper diagnosis is suggested only at surgery, with discovery of mucus-covered, inspissated, cheesy material typical of fungal sinusitis. Diagnosis is confirmed by histology and culture. Treatment of acute invasive and chronic invasive forms requires surgical drainage, débridement, and the use of systemic antifungal agents. Treatment for a fungus ball requires only surgical removal and aeration of the involved sinus. Systemic antifungal agents are not required. Allergic fungal sinusitis is treated by surgical débridement and aeration of the involved sinus followed by the use of systemic and topical intranasal corticosteroids. Recurrence is common and may be heralded by rising serum total IgE levels. Specific immunotherapy with multiple relevant fungal and nonfungal antigens has reduced the necessity for corticosteroid therapy, while preventing recurrence of the disease and the need for repeat surgery.[52]

Phaeohyphomycoses (Dematiaceous Fungi)

Dematiaceous fungi are defined by the presence of melanin in the walls of their hyphae or spores. These fungi are increasingly recognized as causes of fungal sinusitis previously ascribed to *Aspergillus* species. The list of dematiaceous fungi causing sinusitis is rapidly expanding and currently includes *Drechslera, Bipolaris, Exserohilum, Curvularia, Alternaria,* and *Cladosporium* species. Signs and symptoms are essentially indistinguishable from those of chronic indolent *Aspergillus* sinusitis. These fungi are also common causes of AFS. Diagnosis and treatment are similar to that of aspergillosis.

Mucormycosis

Mucormycosis (zygomycosis) is an infection caused by fungi of the *Mucoraceae* family, which includes *Rhizopus, Mucor*, and *Absidia* species. Nearly all cases occur in patients with diabetes mellitus or who are otherwise immunocompromised. These fungi are found in soil and decaying organic matter. When examined microscopically, their hyphae are broad (7 to 20 μm), with few septa and irregular 90-degree branching.

Craniofacial (rhino-orbital-cerebral) mucormycosis usually originates in the nose or sinuses.

Vascular invasion by hyphae leads to thrombosis with resultant ischemic infarction and necrosis. This accounts for the potentially rapid erosion and spread of this infection into the orbit, cranium, palate, or face. Patients with nasal involvement present with purulent or serosanguinous drainage, ulceration, or necrosis. The lateral nasal wall and turbinates are the most frequent site of initial involvement. Edema and erythema are followed by cyanosis and black necrosis. Orbital extension may cause periorbital swelling, proptosis, and ophthalmoplegia. Cranial involvement may cause cranial neuropathies, focal neurologic deficits, and coma. Anesthesia of an involved area may provide an early clue to the presence of mucormycosis.[53]

Prompt diagnosis requires a high index of suspicion and is confirmed by biopsy. Frozen sections may speed diagnosis. A CT scan delineates the extent of the infection. Treatment requires a prolonged course of amphotericin B, as well as aggressive surgical débridement of all nonviable tissue. Antifungal azoles are ineffective. The underlying condition must be corrected if possible.

Histoplasmosis

Histoplasma capsulatum is a dimorphic fungus that exists in a mycelial form in the soil and a budding yeast form at body temperature. It is found throughout the world and in the United States is especially prevalent in the Mississippi and Ohio River valleys. The disease is acquired by inhalation of airborne conidia, often during contact with soil contaminated by avian or bat feces. An increased incidence has been seen in patients with AIDS.

Disseminated histoplasmosis occurs most frequently in patients with impaired cell-mediated immunity. In such patients, lesions of the upper aerodigestive tract are most often seen. The larynx is most commonly involved, with preference for the anterior part of the larynx and epiglottis. Symptoms include sore throat, hoarseness, and dysphagia. Other sites of involvement include the tongue, palate, buccal mucosa, and pharynx. Lesions may appear nodular or ulcerative and may mimic carcinoma, syphilis, tuberculosis, or any other granulomatous lesion. Biopsy is a reliable method of diagnosis. It shows a granulomatous response with budding yeast cells. Amphotericin B is the drug of choice for severe, refractory, or relapsing infections. An oral azole, particularly itraconazole, may be used for milder infections or suppressive therapy.

Blastomycosis

Blastomyces dermatitidis is a dimorphic fungus that grows in a mycelial form at room temperature and as a budding yeast at body temperature. Blastomycosis is most prevalent in the eastern portion of North America and over a wide area of Africa. Infection is thought to be acquired by the inhalation of spores. Dissemination occurs by hematogenous seeding from primary pulmonary foci and, when present, is most frequent in the skin, bone, and genitourinary tract. Skin lesions are most commonly located on the face, extremities, scalp, and neck. The appearance of these lesions may vary. Most begin as papules that, over a course of weeks or months, progress to relatively painless ulcerative or verrucous lesions. Long-standing lesions may show central healing with extensive fibrosis. Subcutaneous abscesses and draining sinuses may also be present.

Involvement of the nose, sinuses, larynx, or mouth may cause well-circumscribed, indurated lesions of the mucosa. As in skin involvement, these lesions are relatively painless, may be ulcerative or verrucous, and may closely mimic squamous cell carcinoma in appearance. Biopsy shows pseudoepitheliomatous hyperplasia of the mucosa with microabscesses, granulomas, and the presence of fungi. Treatment involves the use of systemic antifungal agents. Many patients with indolent head and neck manifestations of blastomycosis can be treated with a prolonged course of oral itraconazole.[54] Intravenous amphotericin B is used in more severe infections.

Cryptococcosis

Cryptococcus neoformans, a yeast with a thick polysaccharide capsule, is saprophytic and distributed worldwide. A common natural source is avian droppings. Many patients with cryptococcosis are immunosuppressed. Acquired immune deficiency syndrome is now the most important predisposing factor. Cryptococcosis occurs in about 8% of AIDS patients and is a commonly recognized cause of life-threatening infection. The fungus is acquired by inhalation, with primary pulmonary infection followed by hematogenous spread. The next most com-

mon site of involvement is the central nervous system. Other frequently involved sites include the bones and skin. The upper respiratory tract is rarely involved. Laryngeal lesions have been described as edematous, exudative, or verrucous. Diagnosis is made by biopsy, which shows pseudoepitheliomatous hyperplasia, granulomatous changes, and the presence of encapsulated fungi. The antifungal azoles and amphotericin B are used for treatment.

Coccidioidomycosis

Coccidioides immitis is a dimorphic fungus, growing in mycelial form in the soil and as thick-walled endospore-forming spherules at body temperature. Coccidioidomycosis is endemic to desert and semiarid portions of the southwestern United States and northern Mexico, as well as scattered areas of Central and South America.

Infection is acquired by the inhalation of spores. Most symptomatic patients develop only a mild to moderate flu-like illness that resolves spontaneously. A few patients develop a severe pulmonary illness that may include hematogenous dissemination. Immunocompromised hosts are especially at risk. The most frequent sites of spread are to the skin, bones, and meninges. In the upper respiratory tract, the larynx is most frequently involved.[55] Lesions may involve any area of the larynx and may appear nodular or ulcerative. Biopsy shows granulomatous changes. Fungal staining shows the characteristic spherules filled with endospores. Antifungal azoles may be effective in indolent disseminated disease, but amphotericin B is required for severe infections.

Paracoccidioidomycosis

Paracoccidioides brasiliensis is a dimorphic fungus found in mycelial form at cooler temperatures and in yeast form at body temperature. Paracoccidiodomycosis is endemic to those areas of Latin America with hot, humid summers and dry, temperate winters. The greatest number of cases are from Brazil, although the disease has been reported from Mexico to Argentina. The fungus is a saprophyte found in soil and decaying organic matter. Infection is acquired through the respiratory tract. Hematogenous spread may then occur to other areas of the body. The mucous membranes of the oral cavity, nose, pharynx, and larynx are involved in approximately half of all cases. In a smaller number of cases, the fungus may be directly inoculated into the mucosa by chewing or cleaning of the teeth with plant fragments. Pulmonary infection is found in the majority of symptomatic patients. Dyspnea, hemoptysis, and chest pain may occur. Lymphadenitis and hepatosplenomegaly are also common.

Lesions of the mucosa of the oral cavity, nose, pharynx, and larynx are painful and manifest as well-defined areas of hyperemic granulation tissue and ulceration. Enlargement of lymph nodes draining these lesions may be present. Clinical manifestations include hoarseness, odynophagia, sore throat, and dyspnea. Destruction of underlying cartilages may lead to nasal collapse. Cutaneous involvement may be present, with areas of the face around the mouth and nose most frequently involved.

Diagnosis is made by the results of smear, culture, and biopsy. Serologic tests may be helpful. Biopsy demonstrates pseudoepitheliomatous hyperplasia with intraepithelial abscesses, granulomatous changes, and the presence of fungi. Treatment is with antifungal azoles, amphotericin B, and sulfonamides.[56]

Conidiobolomycosis

Conidiobolus coronatus is present in soil and decaying organic matter and is endemic to tropical regions of Africa and America. This chronic granulomatous infection begins in the nasal submucosa before extending to the contiguous facial skin. Nasal symptoms may include obstruction, rhinorrhea, and epistaxis. Infection may slowly spread bilaterally to the skin of the nose, upper lip, and cheeks or to the paranasal sinuses, orbits, and pharynx. The skin remains intact but turns shiny and erythematous. Diagnosis is suggested by the clinical picture and is confirmed by potassium hydroxide smear, culture, and biopsy. Treatment is with antifungal agents, although long-term results have been disappointing.

MISCELLANEOUS INFECTIONS

Leishmaniases

The leishmaniases are caused by species of the protozoan genus *Leishmania*. These parasites are transmitted from animal hosts to humans by sandflies.

The disease is endemic to Central and South America, South Asia, the Middle East, Africa, and the shores of the Mediterranean. *Leishmania* are obligate intracellular parasites that primarily inhabit macrophages. Expression of the disease can vary and depends on the species of parasite and on the immune status of the human host. Three major clinical forms of the disease are recognized: visceral, cutaneous, and mucosal.

Visceral leishmaniasis, also known as kala-azar, reflects parasitic infection of the spleen, liver, bone marrow, and lymph nodes. Fever and weight loss are followed by hepatomegaly and massive splenomegaly. Uncommonly, the mucosa of the upper respiratory tract is involved. Untreated patients with kala-azar may die of hepatic failure.

Cutaneous leishmaniasis typically progresses from small papules and nodules to ulcerative lesions, with a central depression and raised, indurated border. Healing may be accompanied by considerable scarring. The external nose is frequently involved. Cutaneous lesions of the nose may extend to the skin of the vestibule but do not ordinarily involve nasal mucosa.

Mucosal (mucocutaneous) leishmaniasis is the least common form of the disease. It occurs as a sequela of the cutaneous form of the disease and appears when the initial skin lesions have healed. The nasal mucosa is most frequently involved. Septal perforation is common, and nasal deformity may result. Advanced lesions may extend onto the external nose, lip, and palate. Mucosal lesions of the oral cavity and larynx have also been described.

The diagnosis of leishmaniasis can be made directly by identification of the intracellular amastigote in stained histologic sections of tissue or by culture of the promastigote from biopsy. Indirect diagnosis can be made by serologic tests (enzyme-linked immunosorbent assay, direct agglutination, others), PCR testing, or monoclonal antibody staining of tissue smears.[57]

Pentavalent antimony in the form of sodium stibogluconate is the most widespread treatment for leishmaniasis. Amphotericin B, pentamidine, and paromomycin are alternate therapies.

Rhinosporidiosis

The etiologic agent of rhinosporidiosis, *Rhinosporidium seeberi*, has been difficult to classify. Because its mature stage consisted of large, thick-walled, spherical structures containing smaller daughter cells (endospores), it was thought to be a fungus. Recent studies, however, by amplifying and sequencing a portion of the *R. seeberi* 18S ribosomal DNA subunit, have demonstrated it to be phylogenetically related to a group of fish parasites referred to as the *Dermocystidium,* rosette agent, *Ichthyophonus*, and *Psorospermium* (DRIPs) clade. These novel aquatic Protistan parasites are near the animal-fungal divergence.[58]

Rhinosporidiosis is endemic to South Asia and is occasionally seen in tropical Africa and America. It is most frequently contracted by immersion in contaminated water. Rhinosporidiosis is a chronic inflammatory disease that most frequently involves the nasal mucosa. The conjunctiva, oral cavity, and larynx are less commonly involved. In the nose, the disease typically manifests as a friable, painless, polypoid growth. Nasal obstruction and bleeding are the most common symptoms. Histologic findings include chronic inflammatory changes and the presence of sporangia. Surgical excision is the preferred treatment. No effective medical treatment as yet exists.

REFERENCES

1. Turner RB. The common cold. Pediatr Ann 1998; 27:790–5.
2. Makela MJ, Puhakka T, Ruuskanen O, et al. Viruses and bacteria in the etiology of the common cold. J Clin Microbiol 1998;36:539–42.
3. Van Kempen M, Bachert C, Van Cauwenberge P. An update on the pathophysiology of rhinovirus upper respiratory tract infections. Rhinology 1999;37: 97–103.
4. Hendley JO. Editorial comment: the host response, not the virus, causes the symptoms of the common cold. Clin Infect Dis 1998;26:847–8.
5. Huovinen P, Lahtonen R, Zeigler T, et al. Pharyngitis in adults: the presence and coexistence of viruses and bacterial organisms. Ann Intern Med 1989; 110:612–6.
6. McMillan JA, Sandstrom C, Weiner LB, et al. Viral and bacterial organism asssociated with acute pharyngitis in a school-aged population. J Pediatr 1986;109:747–52.
7. Kessler HA, Blaauw B, Spear J, et al. Diagnosis of human immunodeficiency virus infection in sero-

negative homosexuals presenting with an acute viral syndrome. JAMA 1987;258:1196–9.
8. Komaroff AL, Pass TM, Aronson MD, et al. The prediction of streptococcal pharyngitis in adults. J Gen Intern Med 1986;1:1–7.
9. Schwartz RH, Hayden GF, Wientzen R. Children less than three years old with pharyngitis. Are group A really that common? Clin Pediatr 1986;25:185–8.
10. Hutt DM, Judson FN. Epidemiology and treatment of oropharyngeal gonorrhea. Ann Intern Med 1986; 104:655–8.
11. Wenger JK. Supraglottitis and group A streptococcus. Pediatr Infect Dis J 1997;16:1005–7.
12. Gorelick MH, Baker MD. Epiglottitis in children, 1979 through 1992. Effects of *Haemophilus influenzae* type b immunization. Arch Pediatr Adolesc Med 1994;148:47–50.
13. Trollfors B, Nylen O, Carenfelt C, et al. Aetiology of acute epiglottitis in adults. Scand J Infect Dis 1998; 30:49–51.
14. Kaditis AG, Wald ER. Viral croup: current diagnosis and treatment. Pediatr Infect Dis J 1998;17:827–34.
15. Wald ER. Sinusitis. Pediatr Ann 1998;27:811–8.
16. Hamilos DL. Chronic sinusitis. J Allergy Clin Immunol 2000;106:213–27.
17. Brook I. Microbiology of nosocomial sinusitis in mechanically ventilated children. Arch Otolaryngol Head Neck Surg 1998;124:35–8.
18. Shapiro ED, Milmoe GJ, Wald ER, et al. Bacteriology of the maxillary sinuses in patients with cystic fibrosis. J Infect Dis 1982;146:589–93.
19. Bluestone CD, Stephenson JS, Martin LM. Ten-year review of otitis media pathogens. Pediatr Infect Dis J 1992;11:S7–11.
20. Pelton SI. Otoscopy for the diagnosis of otitis media. Pediatr Infect Dis J 1998;17:540–3.
21. Block SL. Causative pathogens, antibotic resistance and therapeutic considerations in acute otitis media. Pediatr Infect Dis J 1997;16:449–56.
22. Musher D, Dagan R. Is the pneumococcus the one and only in acute otitis media? Pediatr Infect Dis J 2000;19:399–400.
23. Heikkinen T. Role of viruses in the pathogenesis of acute otitis media. Pediatr Infect Dis J 2000;19:S17–23.
24. Fainstein V, Musher DM, Cate TR. Bacterial adherence to pharyngeal cells during viral infection. J Infect Dis 1980;141:172–6.
25. Heikkinen T, Thint M Chonmaitree T. Prevalence of various respiratory viruses in the middle ear during acute otitis media. N Engl J Med 1999;340:260–4.
26. Berman SA, Balkany TJ, Simmons MA. Otitis media in infants less than 12 weeks of age: differing bacteriology among in-patients and out-patients. J Pediatr 1978;93:453–4.
27. Kenna MA, Bluestone CD, Reilly JS, Lusk RP. Medical management of chronic otitis media without cholesteatoma. Laryngoscope 1986;96:146–51.
28. Wald ER. Anaerobes in otitis media and sinusitis. Ann Otol Rhinol Laryngol 1991;154:14–6.
29. Brook I. Aerobic and anaerobic bacteriology of cholesteotoma. Laryngoscope 1981;91:250–3.
30. Cohen JI. Epstein-Barr virus infection. N Engl J Med 2000;343:481–92.
31. Godshall SE, Kirchner JT. Infectious mononucleosis: complexities of a common syndrome. Postgrad Med 2000;107:175–9.
32. Rubin JS, Honigberg R. Sinusitis in patients with the acquired immunodeficiency syndrome. Ear Nose Throat J 1990;69:460–3.
33. Armstrong M, McArthur JC, Zinreich SJ. Radiographic imaging of sinusitis in HIV infection. Otolaryngol Head Neck Surg 1993;108:36–43.
34. Thompson C, et al. Etiology of acute sinusitis in HIV infection [abstract]. Int Conf AIDS 1993;9:51.
35. Tami TA, Wawrose SF. Diseases of the nose and paranasal sinuses in the human immunodefiency virus-infected population. Otolaryngol Clin North Am 1992;25:1199.
36. Hunt SM, Miyamoto RC, Cornelius RS, Tami TA. Invasive fungal sinusitis in the acquired immunodeficiency syndrome. Otolaryngol Clin North Am 2000;33:335–47.
37. Barnett ED, Klein JO, Pelton SI, Luginbuhl LM. Otitis media in children born to immunodeficiency virus-infected mothers. Pediatr Infect Dis J 1992;11:360–4.
38. Lalwani AK, Sooy CD. Otologic and neurotologic manifestations of acquired immunodeficiency syndrome. Otolaryngol Clin North Am 1992:25:1183.
39. Marelli RA, Biddinger PW, Gluckman JL. Cytomegalovirus infection of the larynx in the acquired immunodeficiency syndrome. Otolaryngol Head Neck Surg 1992;106:296–301.
40. Browning DG, Schwartz DA, Jurado RL. Cryptococcosis of the larynx in a patient with AIDS. South Med J 1992;85:762–4.
41. Rothstein SG, Persky MS, Edelman BA, et al. Epiglottitis in AIDS patients. An unusual cause of fungal laryngitis. Laryngoscope 1989;99:389–92.
42. Cunningham MA. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev 2000;13:470.

43. Shulman ST. Complications of streptococcal pharyngitis. Pediatr Infect Dis J 1994;13:S70–4.
44. Wood TF, Potter MA, Jonasson O. Streptococcal toxic shock-like syndrome. The importance of surgical intervention. Ann Surg 1993;217:109–14.
45. Special Writing Group of the American Heart Association. Guidelines for the diagnosis of rheumatic fever: Jones criteria, 1992 update. JAMA 1992;268: 2069.
46. Hart CA, Rao SK. Rhinoscleroma. J Med Microbiol 2000;49:395–6.
47. Kandiloris DC, et al. Laryngeal tuberculosis at the end of the 20th century. J Laryngol Otol 1997;111: 619.
48. Jacobson RR, Krahenbuhl JL. Leprosy. Lancet 1999; 353:655–60.
49. Miller M, Haddad AJ. Cervicofacial actinomycosis. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1998;85:496–508.
50. Nagler R, Peled M, Laufer D. Cervicofacial actinomycosis: a diagnostic challenge. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1997;83:652–6.
51. Zegarelli DJ. Fungal infections of the oral cavity. Otolaryngol Clin North Am 1993;26:1069–89.
52. Mabry RL, Mabry CS. Allergic fungal sinusitis: the role of immunotherapy. Otolaryngol Clin North Am 2000;33:433–40.
53. Ferguson BJ. Mucormycosis of the nose and paranasal sinuses. Otolaryngol Clin North Am 2000;33:349–65.
54. Hanson JM, Spector G, El-Mofty SK. Laryngeal blastomycosis: a commonly missed diagnosis. Ann Otol Rhinol Laryngol 2000;109:281–6.
55. Boyle JO, Coulthard SW, Mandel RM. Laryngeal involvement in disseminated coccidioidomycosis. Arch Otolaryngol Head Neck Surg 1991;117:433–8.
56. Sant'Anna GD, Mauri M, Arrarte JL, Camargo H. Laryngeal manifestations of paracoccidioidomycosis (South American blatomycosis). Arch Otolaryngol Head Neck Surg 1999;125:1375–8.
57. Lohuis PJ, Lipovsky MM, Hoepelman AI, et al. *Leishmania braziliensis* presenting as a granulomatous lesion of the nasal septum mucosa. J Laryngol Otol 1997;111:973–5.
58. Herr RA, Ajello L, Taylor JW, et al. Phylogenetic analysis of *Rhinosporidium seeberi's* 18S small-subunit ribosomal DNA groups this pathogen among members of the protoctistan *Mesomycetozoa clade*. J Clin Microbiol 1999;37:2750–4.

CHAPTER 31

Imaging of the Nasal Cavities, Paranasal Sinuses, Nasopharynx, Orbits, Infratemporal Fossa, Pterygomaxillary Fissure, Parapharyngeal Space, and Base of the Skull

Mahmood F. Mafee, MD

The development of computed tomography (CT) and magnetic resonance imaging (MRI) has been regarded as the most important contribution to diagnostic medical imaging techniques since Roentgen discovered the x-ray in 1895. Computers were first used in nuclear medicine in the 1960s, but it was the introduction of CT in 1971, an invention of Godfrey N. Hounsfield, who won the Nobel Prize for his pioneering research, that accelerated interest in the application of computers to the broader aspects of radiology. Like CT, MRI has proven to be a major breakthrough in diagnostic medical imaging. Based on the pioneering work of Paul Lauterbur in the early 1970s,[1] MRI, with its continued refinement, is considered to be the most important medical imaging advance of the century.

COMPUTED TOMOGRAPHY

The technique of CT employs a narrow highly collimated rotating (scanning) x-ray beam that is used to irradiate the patient from numerous projections. Detectors with high detection efficiency, located on the other side, opposite to the source of the beam, register the amount of radiation that has passed through the part of the body.[2] The data collected are processed by a computer, which, by solving thousands of equations instantaneously, produces a true tomographic section.[2,3] The reconstruction of the subject cross-section resides in the computer as a series of digital numbers, usually referred to as "CT numbers."[2] These numbers are related to the linear x-ray attenuation of the material in the volume element (voxel) at the location of interest.[2]

Zero is arbitrarily chosen to represent water. The numbering system –500 to +500 used on the original scanners is referred to as old Hounsfield units.[2] The numbering system, –1,000 or less to +1000 or more, is referred to now as Hounsfield units (HU). In this system, air is black (dense air, –1,000 HU) and bone is white (dense bone, +1,000 HU). Fat absorbs the x-ray less than water and more than air; therefore, its CT numbers will be negative, usually between –50 and –140 HU. The CT number of circulating blood is +40 to +50; congealed blood, +50 to +80; calcium, +80 to +500; muscle, +30 to +40; cerebrospinal fluid (CSF), +5 to +12; bone, +100 to +1000; and air, –100 to –1000.

The viewing devices (console) provided with CT scanners usually have a set number of gray scales (eg, 50, 100, 150, 300, 1,000, or more). The range of CT numbers covered by these gray scales (the contrast) is adjustable (eg, window-width control) (Figure 31–1).[2] The location of the center of the gray scale also is adjustable (window-level control) (see Figure 31–1). This manipulation of window width (image contrast) and window level (mean level; see Figure 31–1) by the CT radiologist at the console gives the most effective diagnostic information, ensuring that the diagnostic sensitivity is limited by the recorded image and not the observer's eye. A window width of 80 units with the level of 35 examines values ranging from +75 to –5 (see Figure 31–1). With the window level = L = +35 and window width = W = +80 setting, values higher than +75 will be white, those lower than –5 will be black, and those between +75 and –5 will be shades of gray (see Figure 31–1).

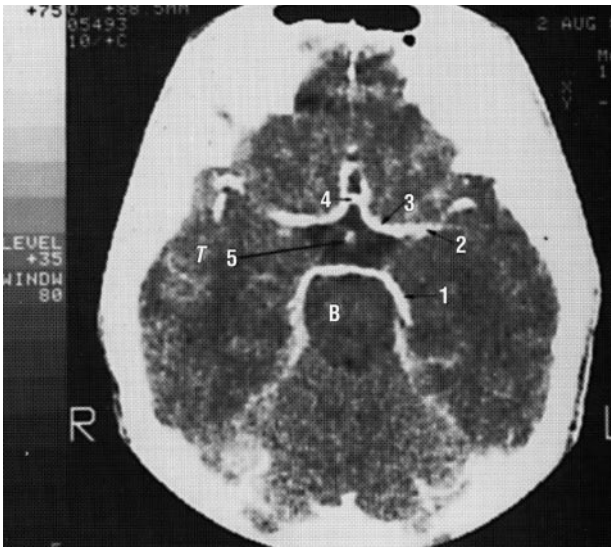


FIGURE 31–1. Normal circle of Willis. Postcontrast axial computed tomographic (CT) scan of the lower part of the head at the level of the suprasellar cistern demonstrates the enlarged pituitary stalk (5) in the center of the low-density star-shaped suprasellar cistern. The posterior cerebral (1), middle cerebral (2), anterior cerebral (3), and anterior communicating (4) arteries are clearly seen as high-density images because of their higher iodine-contrast content. This image was obtained with a window width of +80 (+75, –5) and a window level of +35. With this setting, values higher than +75 will be white (bones), and those lower than –5 will be black (air in the frontal sinus). Note the differential enhancement of the white and gray matter in the brainstem (B) and temporal lobe (T). The gray matter, by virtue of its richer vascularity, shows more enhancement than the white matter. The cerebrospinal fluid (suprasellar cistern, cerebral aqueduct) appears as a low-density image in CT.

Spiral or Helical Computed Tomography

Since the introduction of CT, significant improvements in CT technology have been made. The introduction of spiral or helical CT has a great impact in cross-sectional imaging. Spiral CT has introduced the concept of volumetric (three-dimensional) imaging with x-ray. In spiral CT technology, continuous patient motion through the gantry combined with uninterrupted beam rotation leads to the spiral pattern of volumetric data acquisition. Spiral CT affords us with new opportunities. The speed with which studies can be carried out allows much faster patient throughput. The new-generation multislice volumetric spiral CT provides unprecedented power and speed, allowing scanning of the whole body in a few seconds. This has opened the door to new applications, such as CT angiography, cardiac imaging, and rapid dynamic CT. With spiral CT, the quality of reformatted or postprocessed images has significantly improved. For children and agitated patients in whom speed is of the utmost importance, spiral CT is extremely useful to provide acceptable diagnostic information in a matter of a few seconds.

MAGNETIC RESONANCE IMAGING

Nuclear Magnetic Resonance Phenomenon and Magnetic Resonance Imaging

Magnetic resonance imaging is a fascinating field and has introduced an entirely new and important parameter in the evaluation of abnormalities of the head and neck and neurologic disorders.[4–6] The technique of MRI scanning is considerably more complex than CT, and its full diagnostic potential is often not realized because clinicians have an inadequate understanding of the basic principles of MRI and the various technological modifications available for study of specific lesions. Advances in MRI technology are occurring on an ongoing basis; therefore, MRI protocols can vary from institution to institution.

Historical Background

In 1936, C. J. Gorter, using calorimetric techniques, described his unsuccessful attempt to observe the lithium (Li) nuclear magnetic resonance (NMR) in lithium chloride (LiCl).[7] In 1938, I. Rabi and coworkers described the first successful molecular beam NMR.[8] Subsequent to the successful molecular beam nuclear magnetic experiments, Gorter and Broer, in 1942, again attempted to observe NMR in powders of LiCl and potassium fluoride, but no resonance was observed.[9] Despite Gorter's failure, two independent groups, Felix Bloch and coworkers and Edward Purcell and coworkers in 1946, were successful in demonstrating NMR in condensed matter, work for which they shared the Nobel Prize in physics in 1952.[10,11]

Historically, the impetus of the development of MRI is attributable to Lauterbur, who, in 1973, by

using the interaction of magnetic fields, generated the first two-dimensional MRI.[1] This depicted the proton density and the spin-lattice relaxation time (T_1) distribution (see section of MR phenomenon) in two 1 mm tubes of water.[1] Subsequently, numerous NMR techniques have been developed generating data from a point, line, plane, or multiple planes within a volume and from the entire volume.[12]

MAGNETIC RESONANCE PHENOMENON

Elementary Physics

In discussing x-rays and their effects on atoms, we are concerned with their extranuclear structure, that is, the arrangement of planetary (orbital) electrons, outside the nucleus. In MR phenomenon, we are concerned with nuclear structures of atoms, that is, the nucleons ("nucleon" is the generic term for a proton or neutron). The existence of nuclear spin and nuclear angular momentum, the entity essential to the MR phenomenon, was first suggested by the Austrian physicist Wolfgang Pauli in 1924 when he observed the behavior of light in a magnetic field.[4] Since then, it has been verified that atomic nuclei have an angular momentum arising from their inherent property of rotation or spin. Pairs of protons or neutrons, however, align in such a way that their spins cancel out, and a net spin (net rotation) will therefore exist for a nucleus only when it contains an odd (unpaired) proton, an odd neutron, or both.[13] Since protons carry a positive charge, the nuclei have an associated electric charge distribution; therefore, the spin generates a current flowing about the spin axis, which, in turn, generates a small magnetic field. Each nucleon of nonzero spin has an inherent magnetic field (magnetic dipoles) and a magnetic moment (a magnetic moment is a vector quality represented by an arrow, which describes the magnetic field's strength and direction), or dipole, associated with it that can be thought of as behaving like a tiny spinning bar magnet with north and south poles (Figure 31–2). Therefore, these tiny spinning magnets can interact with the electromagnetic field.[12–14] In a sample of materials, the nuclear magnets (magnetic moments or dipoles) will be pointing in random directions,[12–14] and the result of many nuclei in a random arrangement is that there is no net magnetization (one cancels the other) in any direction (see Figure 31–2). However, if the nuclei are placed in an external static magnetic field, the randomly oriented magnetic dipoles respond to the force of the field by trying to line up in the direction of the field (see Figure 31–2). In the magnetic field, therefore, there is a net (sum) magnetization produced parallel to and in the same direction as the applied field (see Figure 31–2). The magnetic behavior of the entire population of nuclei can be defined as a macroscopic or bulk magnetization vector or moment (M) (see Figure 31–2), which represents the net effect of all of the magnetic moments of the nuclei of a given nuclear species in the sample.

In the absence of an external magnetic field, the net magnetization is, of course, zero owing to random orientation of nuclei. When the external magnetic field B_o (see Figure 31–2) is imposed on the sample, however, the nuclear dipoles become orientated parallel to the applied field, and the net magnetization (spin vector, M) (see Figure 31–2) yields a finite equilibrium that will point in a direction parallel to the external magnetic field B_o (see Figure 31–2). This direction conventionally defines the longitudinal (Z) axis (see Figure 31–2). By convention, the spin system is presented in a cartesian coordinate system (X, Y, Z), with the Z-axis parallel with the direction of the external magnetic field B_o. (see Figure 31–2).

Precession Motion

Let us focus particular attention on the nuclei of the hydrogen atoms, protons, which, from the standpoint of the MR phenomenon, are particularly favorable nuclei. The hydrogen atom is now referred to as a proton. Each proton can be regarded as a small, freely suspended bar magnet spinning rapidly about its magnetic axis. If a group of protons are placed in a static magnetic field of particular strength (B_o), the magnetic moment of each proton experiences a couple (from the north and south pole), a force or torque tending to turn its magnetic moment (M) parallel to the static magnetic field (see Figure 31–2). Because it spins, the proton responds to this couple by rotation about the line of force (like a gyroscope precessing about the gravitational force); therefore, its axis is tilted and begins to precess (rotate) about the direction of the magnetic field in a movement known as "precessing" (see Figure 31–2). The precessional motion traces a cone in a manner analogous to the wobbling of a toy top as it spins and tries to orient itself in the earth's gravitational field. The frequency of the precessional motion given by the famous theorem of Larmor and the Larmor equation, ω (omega)

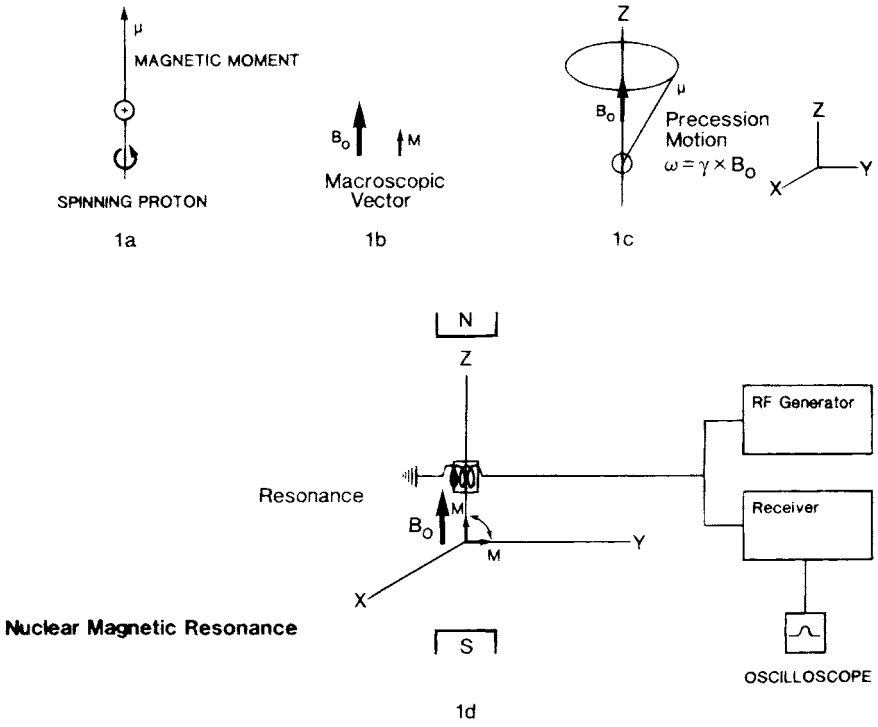


FIGURE 31–2. Nuclei of odd number of neutrons, protons, or both, such as hydrogen (proton), have a net spin or "nuclear angular momentum." Each spinning nucleus generates a magnetic field and hence a magnetic moment that is a vectoral quantity with both direction and magnitude (1a). Nuclear magnets of the nuclei with spin will be pointing in random directions. Therefore, there is no net magnetization in any direction. If the nuclei are placed in an external static magnetic field B_o (1b), the randomly oriented magnetic moments line up in the direction of field at a state of equilibrium. The magnetic behavior of the entire population of nuclei is defined as a macroscopic or bulk magnetization vector M (1b). When a proton is placed in a magnetic field B_o (1c), it experiences a couple (from the north and south poles), tending to turn its magnetic moment parallel to the field. Because it spins, the proton responds to this force like a gyroscope, and its axis of the magnetic moment vector M wobbles or precesses about the direction of the magnetic field B_o (1c). The frequency (ω) of this precessional motion is unique for each nuclear magnetic sensitive element and is given by the famous theorem of Larmor ($\omega = \gamma \times B_o$). By convention, the spin system is presented in a cartesian coordinate system (X, Y, Z) with the Z-axis parallel with the external magnet field B_o. When the spin system (M) is placed in an external magnetic field (B_o), besides the creation of precession, another phenomenon, a resonance condition, is produced in which the nuclei can absorb energy at sharply defined frequency (Larmor frequency) from an electromagnetic pulse (wave) and re-emit the energy as electromagnetic radiation of the same frequency after the incident beam has been turned off. Let us wind a coil around a nuclear magnetic sensitive specimen and place it inside a magnet (with north, N, and south, S, poles), as shown in 1d, and supply it with radiofrequency (RF) current at Larmor frequency. A short pulse of each resonant RF current can be arranged to tip the net magnetization vector (M) through 90 degrees in the X-Y plane (1d) that is perpendicular to B_o. This is called 90-degree pulse. The magnetization M now continues to rotate in the X-Y plane at the same frequency and induces a voltage (current) in a receiver coil (antenna) in the X-Y plane. If the pulse is then turned off, the excited spin system (the nuclei) returns to its state of equilibrium and loses the excess energy. The vector (M) continues to rotate freely in the X-Y plane, causing the induced current (signal) in the antenna to fade. This decaying signal is called the "free induction decay" (FID), which can be amplified, detected, and displayed on an oscilloscope or fed into a computer to provide pictorial information. The return of the excited spin system to the state of equilibrium, the so-called relaxation, is characterized by two sample-related time constants. T_1 (spin-lattice relaxation time) is the time constant required for the net magnetization (M) to return from the X-Y plane to its equilibrium along the Z-axis. T_2 (spin-spin relaxation time or transverse relaxation time) is the time constant of the decay of FID.

$= \gamma$ (gamma) $\times B_0$, states that the frequency of the precessional motion (ω) is proportional to the couple and therefore to the strength of the applied static external magnetic field (B_0) and the constant of proportionality gamma, the so-called gyromagnetic property (or ratio) of the specific nucleus.[4] The equation shows that as the magnetic field changes, so does the nuclear precessional rate of motion (angular frequency ω or Larmor frequency). This is an important point in MRI because it enables us to locate the origin of the MR signal and to pinpoint spatially the source of each frequency (ω).

Nuclear Magnetic Resonance

Resonance is a physical phenomenon that permits the transfer of energy from an object, particle, or system vibrating mechanically or electronically to another similar object, particle, or system, causing the latter to vibrate at the same frequency. When a substance containing nuclei of nonzero spin is placed in a magnetic field, a resonance condition is produced, which means that there will be a strong interaction or resonant effect.[4,15] The nuclei can then absorb energy at a sharply defined frequency from electromagnetic radiation and re-emit the energy as electromagnetic radiation of the same frequency after the incident beam has been turned off.[4,13–15] Either the absorption or the re-emission can be measured. To secure this resonant effect, the applied electromagnetic radiation (pulse) from the radiofrequency (RF) generator (see Figure 31–2) must have the same frequency as the Larmor precession frequency (natural resonant frequency) of the sample (ie, it has to be resonant). Hence, the name "nuclear magnetic resonance" has been designated for the phenomenon. In short, the precession is a resonance effect, and if a system has a natural resonance (frequency resonance, frequency of oscillation), energy needs can most efficiently be imparted at the resonance frequency. For example, at the field strength of 1.5 tesla (1 tesla = 10,000 gauss), a pulse at 63.87 megahertz (MHz) will excite hydrogen (protons), whereas a pulse at 25.86 MHz will excite phosphorus nuclei.[4]

Relaxation Phenomenon and Times

As mentioned earlier, in the presence of a static magnetic field, B_o, nuclei with nonzero spin orient themselves with the magnetic field lines that point in the "Z" direction at a state of equilibrium "rest" (see Figure 31–2). At equilibrium, the nuclei possess their lower energy level. When an electromagnetic radiation (RF) or RF pulse (wave) is used at a sharply designated frequency (Larmor frequency) because of resonant condition, some of the nuclei can absorb energy from the RF pulse, become excited (attain higher energy level), and change the direction of their magnetic moments. When the RF pulse is turned off, the excited nuclei relax and return to their original orientation (equilibrium, lower energy level), losing the excess energy by emitting electromagnetic radiation (the source of the MR signal) of the same frequency as the applied RF pulse by transferring energy to surrounding (lattice) molecules. This process is called "relaxation" and is characterized by two sample-related relaxation time constants: the longitudinal (spin-lattice) relaxation time (T_1) and the transverse (spin-spin) relaxation time (T_2). When the RF pulse of specific frequency (Larmor frequency) is applied, it affects the NMR-sensitive nuclei in two ways: (1) the RF pulse excites the nuclei and makes them move into a higher energy level, and (2) the net magnetization vector (M) (see Figure 31–2) will precess along the direction of the main field (Z axis) with an ever-increasing angle depending on the intensity and duration of the RF pulses (Figure 31–3). The magnetization vector (M) will consequently spiral down toward the X-Y plane (90-degree pulse) and then continue to rotate in the X-Y plane (see Figure 31–3). The magnetization vector M can be considered as having a component, M_z (longitudinal magnetization), along the z-axis (see Figures 31–2 and 31–3) and a component, M_{xy} (transverse magnetization), in the x-y plane perpendicular to the field axis (Z) (see Figure 31–3). At equilibrium, the "longitudinal magnetization" M_z is equal to the net magnetization M, and the "transverse magnetization" M_{xy} is zero (see Figure 31–3). After RF excitation, the precessional angle (see Figures 31–2 and 31–3) continues to increase as long as the RF pulse is applied (see Figure 31–3). The precessional or Larmor frequency (cycles per second), however, remains constant, being fixed by the inherent properties of the nuclei and the strength of the static magnetic field, B_o. After the excitation of the RF pulse, M_z decreases and M_{xy} increases (see Figures 31–2 and 31–3). The RF needed to tip the vector (M) through an angle of 15 degrees is described as a 15-degree pulse. Pulses of 90 and 180 degrees cause corresponding increases in the

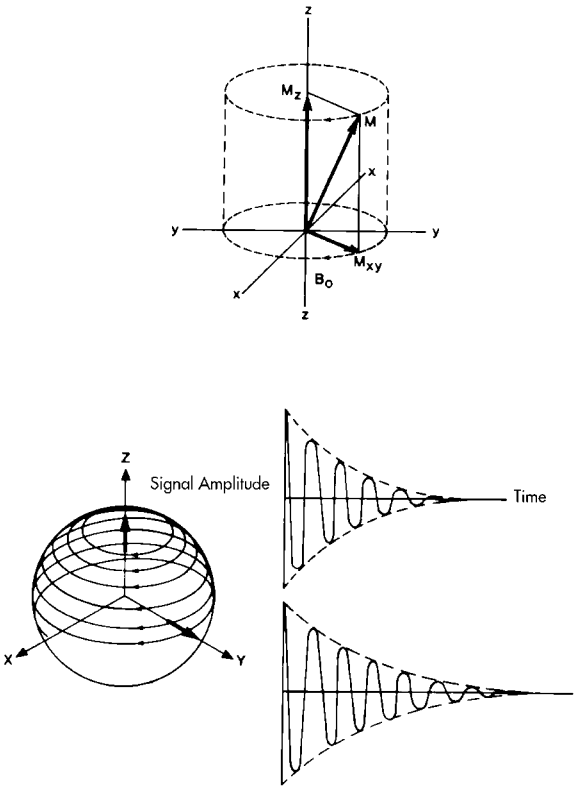


FIGURE 31–3. The precessing net magnetization vector (M) can be considered as having a component, M_z, along the axis of the main field (B_o) and a component, M_{xy}, in a plane perpendicular to the field axis such as the x-y plane (*top image*). At equilibrium, the net magnetization is aligned with the main field so that M_z has its maximum (M_z = M) and M_{xy} has its minimum (zero) magnitude. After a 90-degree pulse, the M_{xy} has its maximum and M_z its minimum (zero) magnitude. Following application of a radiofrequency (RF) pulse, the net magnetization vector M, will precess at an ever-increasing angle to the main field direction (bottom image), with the precession occurring at the inherent or natural Larmor frequency. If the RF pulse is long enough or strong enough, the net magnetization vector (M) spirals down toward the X-Y plane (*bottom image*). When the RF pulse is turned off, the spin system will relax, and the excess energy is released as decaying signal (free induction decay) with a time constant T_2. The T_2 relaxation time for most tumors is longer than that for normal tissue. In other words, the nuclear magnetic resonance signal of the tumor (lower decaying pulse) decays longer than the normal tissue (upper decaying pulse).

precessional angle.[13] After a 90-degree pulse, the vector (M) will be in the x-y plane, and when the nuclei are flipped through 180 degrees, the net magnetization vector (M) will be in the opposite direction of the main field, B_o. The "transverse magnetization" M_{xy} is time dependent and has a rotational or "precessional property" in the x-y plane. Such rotating magnetic current acts as a radiator of RF energy and therefore, according to Faraday's law of induction, can induce a current (voltage) in a receiver coil (inductance antenna) in the x-y plane (see Figure 31–2). Immediately after the application of a specific pulse through the transmitting coil (see Figure 31–2), the transverse magnetization M_{xy} continues to rotate freely in the x-y plane, generating an induced current (signal) that can be detected either by the same coil (acting as transmitting and receiving coil) that transmitted the pulse or by a specific receiver coil (antenna). The emitted signal is called the free induction signal or the free induction decay (FID). The magnitude and length of the FID (NMR signal) are determined by the relaxation times.[15]

T_1 Relaxation Time

The T_1 relaxation time represents the time constant required for the net magnetization vector (M) to return to its equilibrium after it has been perturbed by the RF pulse. In other words, T_1 is the time that characterizes the recovery of the net magnetization (M), after it is perturbed, to its equilibrium state in the field direction (see Figures 31–2 and 31–3). The alignment of nuclei is not instantaneous but rather grows exponentially, increasing to approximately 63% of its final value within a period equal to the time constant (T_1). It will reach 95% of the final value at a time equal to three times the constant. Therefore, in reality, T_1 is the time that characterizes the recovery (return of excited spin to equilibrium) of 63% of nuclear magnetization. T_1 is called "spin-lattice relaxation" because the thermal equilibrium state, hot state (excited spin) to cold state (spin at rest or equilibrium), is reached through an exchange of energy with its molecular environment or lattice. The alternative term, "longitudinal relaxation time," is used to reflect the fact that T_1 characterizes the behavior of the M_z component of the magnetic vector (M) (see Figure 31–3) when magnetization reverts to its equilibrium orientation in the direction of the applied field (Z) (see Figure 31–3). Basically, T_1 is a measure of how fast energy can be transferred from the spinning nuclei to their surrounding environment or lattice by random collision between

molecules or the rate at which thermal equilibrium is restored after being disturbed.

T_2 Relaxation Time

As mentioned, the NMR signal picked up after an RF pulse is known as FID. The signal is proportional to the magnetization in the sample and would be produced continuously if the individual nuclear spins were able to maintain their coherence (synchronization) with no shifting of their relative phases.[14] The FID signal, however, decays because of spin-spin ("dipole-dipole") interaction (magnetic interaction of neighboring nuclei) and the inhomogeneity of the static magnetic field B_o, both causing fluctuation in local magnetization with resultant fluctuation (any error) in precessional frequency (around Larmor frequency) of the individual proton. This causes the resonance of protons within the sample to lose synchronization and their vectors to develop a phase difference. The signal decays exponentially in time, with a time constant characterized by the T_2 (see Figure 31–3). Basically, T_2 is the time taken for the transverse magnetization, M_{xy}, to decrease by 63% of its original value. In a period three times as long as this, the transverse magnetization will almost have disappeared. In short, T_2 relaxation time is the time taken for the decay of FID or MR signal to decrease by 63% of its original intensity value.

For water, T_1 and T_2 are several seconds, the precise value depending on the purity of the water. For cellular water, T_1 and T_2 are shorter. T_1 is shorter in liquids (seconds) than in solids (minutes). In solids and/or in a sample of materials at low temperatures, where the atoms and molecules move about very little, transfer of thermal energy is slow; therefore, T_1 can last for hours.[13] T_1 is short in liquids because of the mobility of the atoms and molecules and the consequent rapid transfer of thermal energy.

For protons in pure water, T_1 and T_2 are approximately equal.[13] T_2 is short in solids (microseconds) and long in fluids. In pure fluids, because of rapid motion of the atoms and molecules (random process), the magnetic field contribution from neighboring nuclei averages zero, and only the main magnetic field determines the resonant frequency. In this case, T_2 is very long. If the nuclei find themselves in impure liquids where molecules attach to proteins and have low translational and rotational speeds, this averaging random process of the magnetic field contribution from neighboring nuclei is more effective, and the resonant frequency is therefore different (rapid phase difference and loss of coherence) for each nucleus; this makes T_2 short. Solids, on the other hand, have fixed atoms and molecules that maintain local field variation (spin–spin or dipole–dipole interaction), which result in a rapid loss of coherence and a very short T_2. Relaxation times are affected by viscosity, temperature, and the presence of dissolved ions and molecules.[14]

Nuclear Magnetic Resonance Spectroscopy and Chemical Shift

Nuclear magnetic resonance spectroscopy is a plot of signal intensity versus resonance frequency. The appearance of NMR spectra depends on the density and molecular environment surrounding a nucleus. The presence of neighboring atoms and the "shielding currents" that are associated with the distribution of electrons around adjacent atoms slightly modifies the strength of local magnetic fields at particular nuclei and hence modifies the resonance frequency[12–14,16] so that the details of molecular structure change the NMR frequency of a nucleus in that molecule. The extent of this field change depends on the electronic environment of the observed nucleus; consequently, identical nuclides have slightly different resonance frequencies (lines) if they are residing in different chemical structures (environments). For example, the resonant frequency of a hydrogen proton within a water molecule differs slightly from that of a hydrogen proton within a fat molecule, and the resonant frequency of a 31P nucleus bound in an adenosine triphosphate (ATP) molecule varies slightly from that of a 31P nucleus within a phosphocreatine molecule. These small but specific displacements or shifts in frequency produce separate absorption peaks in a spectrum (chemical shift), which are "fingerprints" of the molecular conformations and directly aid in the determination of chemical structures.[13–18] The chemical shift is measured and listed in parts per million with respect to the signal from a reference such as a signal of tetramethylsilane. The chemical shift is of paramount importance in analytic chemistry since it allows the assignment of different signals (lines) in an NMR spectrum to corresponding substances or molecular groups. For example, the proton spectrum of acetic acid CH_3-COOH provides two lines. One line represents the three protons (hydrogen) of the

methyl group and another line the carboxyl proton. The three protons of the methyl group have an identical electronic environment, so their signal appears as a single resonance line. In 31P spectra of muscle tissue, the signal of phosphocreatine is used as a reference signal. The 31P spectrum of a muscle typically shows five signals that are related to inorganic phosphate, phosphocreatine, and the three phosphorus atoms of ATP (alpha, beta, and gamma).

Currently, NMR spectroscopy is one of the most powerful tools used in chemical analysis.[13] The combined MRI, in vivo spectroscopy, has become a routine procedure to provide morphologic as well as noninvasive biochemical "noninvasive biopsy" information of internal structures of human tissues.[16–18]

Clinical Use of T_1-Weighted, Proton-Weighted, and T_2-Weighted Magnetic Resonance Imaging

Magnetic resonance scanners produce images that represent one or more of NMR parameters. These are (1) hydrogen (proton) density (also called spin density), (2) the state of motion of hydrogen, (3) the tissue relaxation times T_1 and T_2, and (4) chemical shift. All forms of unprocessed MR signals depend on nuclear spin density.

The ability to rotate net magnetization (M) (see Figures 31–2 and 31–3) from the direction of the static magnetic field (B_o) into some other orientation by applying an RF pulse is fundamental to MR spectroscopy and imaging techniques because it provides a flexible method for investigating relaxation processes and other phenomena.[14] By choosing an appropriate pulse sequence, the intensity of an MR image can be made to reflect one or more of several NMR parameters inherent to the tissue being examined.[13,14] Although the proton density, chemical makeup, T_1 and T_2, and motion represent intrinsic tissue properties, the transition of these parameters onto the MR image is also dependent on the particulars of the imaging technique used to form that image. This subject has been described by many authors in several articles.[19,20]

The 90- and 180-degree pulses are used for NMR techniques.[14] In forming an MR image, the RF pulse must be applied repeatedly. After each 90-degree pulse, the magnetization of the sample is reduced to zero as this rotates the longitudinal magnetization vector (M) into the transverse (x-y) plane to become transverse magnetization M_{xy} (see Figures 31–2 and 31–3). Therefore, the longitudinal component of magnetization M_z is zero. The longitudinal component of magnetization will recover exponentially with a time constant of T_1. Rapid repetition of pulses (irradiation of sample) would not allow much recovery of magnetization, and little signal would be detected after each successive 90-degree exciting pulse. Thus, a certain time (repetition time = TR) is introduced between successive RF pulses. As the TR increases, the signal and imaging time increase. After RF pulse, as discussed earlier, the FID signal decays exponentially in time with a time constant T_2. The interval between application of the RF pulse and reception of a signal or spin echo is called echo time or time of echo (TE). As TE increases, signal decreases (see Figure 31–3), and contrast between tissues of different values of T_2 changes. If the signal is obtained immediately after RF excitation, it will not undergo any T_2 decay; therefore, although it is a strong signal, it has no T_2 dependence. If, instead, the signal is obtained sometime after RF excitation, the available signal, although weaker, has an exponential dependence on T_2. Strong signals (proton-weighted [PW] and T_1-weighted [T_1W] images) yield better signal-to-noise levels or ratios and, therefore, better spatial resolution. With better optical resolution, anatomy is better appreciated. However, in the image, the differences owing to T_2 relaxation time (T_2-weighted [T_2W] image) contribute to contrast resolution with better pathologic details to the extent that T_2 often contributes to tissue characterization.

Spin-Echo Pulse Sequence

The spin-echo (SE) pulse sequence is the most commonly used technique to acquire MR images. T_1-weighted MR images can be obtained using the SE pulse sequence with a short TR of 400 to 1,000 ms and a short TE of 20 to 25 ms. The SE pulse sequence of long TR (1,000 to 3,000 or more ms) and short TE (20 to 40 ms) provides proton density (PW) MR images. Long TR (1,000 to 3,000 or more ms) and long TE (40 to 120 or more ms) provide the T_2W MR images.

Inversion Recovery Pulse Sequence

Various pulse patterns that emphasize different aspects of the evoked NMR signal can be exploited to

alter the contrast of MR images.[13] In an "inversion recovery" (IR) pulse sequence, a 180-degree (inverting) pulse, followed by a 90-degree (read) pulse, must be applied repetitively for signal averaging purposes.[12,17] Because of application of a 180-degree pulse and the fact that magnetization must be allowed to recover after the 180-degree inverting pulse, the IR sequence is slower than SE in data collection.

Fluid-Attenuated Inversion Recovery

Fluid-attenuated inversion recovery (FLAIR) is used to suppress the high CSF signal in T_2W MR images so that pathology adjacent to CSF is seen more clearly. This is a powerful pulse sequence to detect demyelinating disease, subarachnoid hemorrhage, leptomeningeal pathology, and most of intracranial pathology.

Use of Gadolinium Contrast Material

Gadolinium, a rare earth element, is used in conjunction with T_1W pulse sequences. With a standard dose of 1 mmol/kg, it has little effect on T_2W MR images. The use of gadolinium-diethylenetriamine pentaacetic acid contrast material improves lesion delineation. It is often used with fat suppression T_1W MR technique in the head and neck or where the lesion may be obscured within high-intensity fat.

The gadolinium T_1W pulse sequence is the best to evaluate intracranial extension of head and neck lesions. Gadolinium also enhances normal structures, including mucosa, lymphoid tissue, extraocular muscles, and slow-flowing blood in veins.

The viewing devices (console) provided with MR scanners, like CT scanners, have a set number of gray scales that can be used to adjust the contrast. The MR parameters (TR, TE) and pulse sequences (SE, IR), the strength of the magnet (eg, 1.5 tesla), and the time and number of images in each series of pulse sequences are recorded on the screen and reflected on each frame of hard-copy MR films. Therefore, a simple method for beginners is to look for TR and TE to recognize the T_1W, PW, and T_2W images. A short TR (less than 1,000 ms) and a short TE (20 to 25 ms) is a T_1W image (Figure 31–4). A long TR (more than 1,000 ms) and a short TE (20 to 40 ms) is a PW image, and a long TR (1,000 ms or more) and long TE (more than 40 ms) is a T_2W image (Figure 31–5). T_1-weighted and PW images provide better anatomic details, and T_2W images are best for pathologic details.

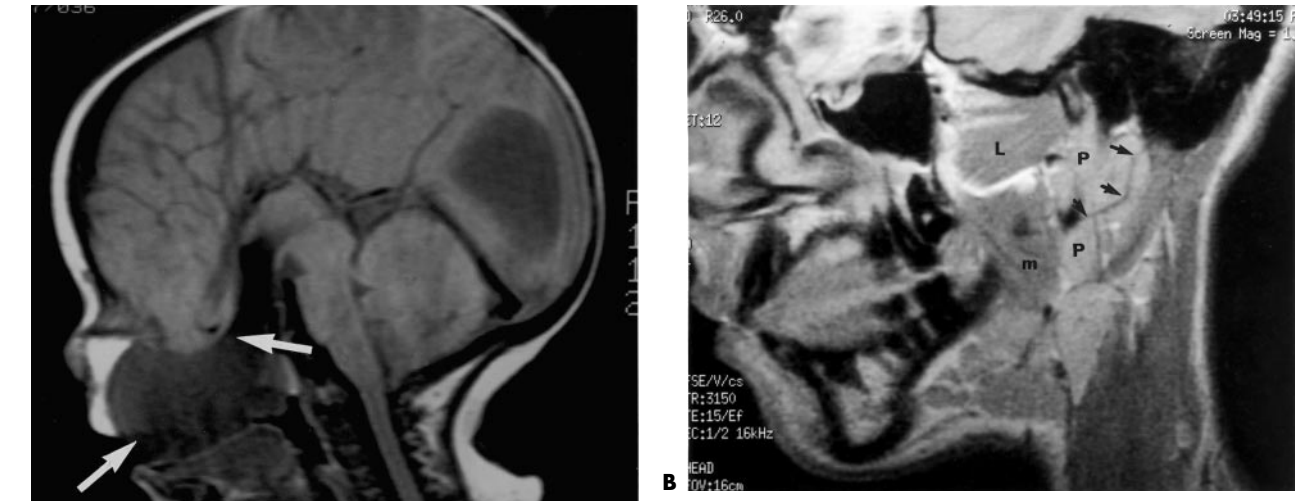


FIGURE 31–4. *A*, Sagittal T_1-weighted magnetic resonance image showing a large nasal meningoencephalocele (*arrows*) in this child. *B*, Oblique sagittal proton-weighted (TR = 3150, TE = 15 ms) magnetic resonance image showing parotid gland (P), lateral (L) and medial (m) pterygoid muscles, and the course of facial nerve (*arrows*).

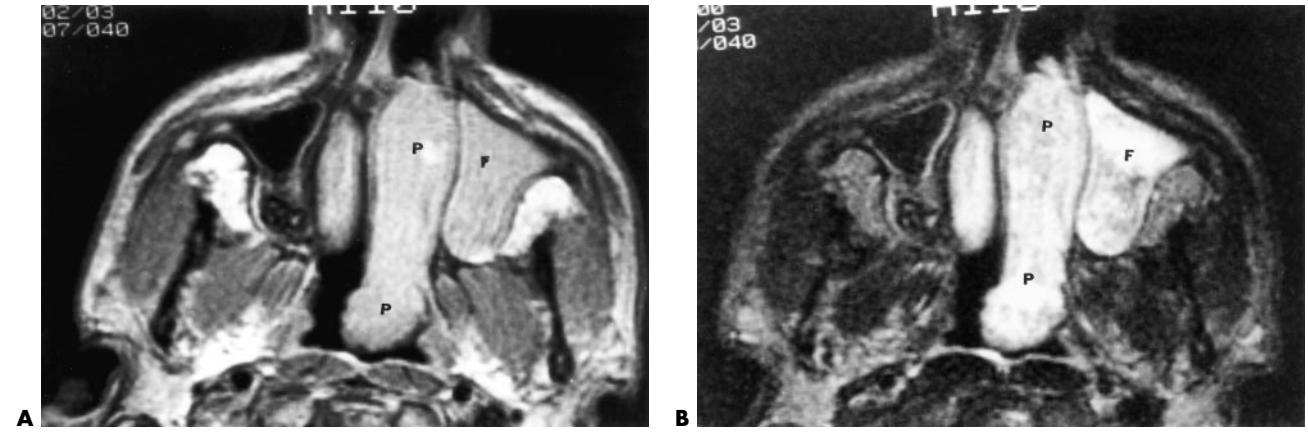


FIGURE 31–5. Inverted papilloma. Proton-weighted (PW) (2000/02, TR/TE) (*A*) and T_2-weighted (T_2W) (2000/80, TR/TE) (*B*) magnetic resonance (MR) images showing a large inverted papilloma (P). Notice the hyperintensity of the tumor on the T_2W MR image. Extension of the tumor in the left maxillary sinus is better demonstrated on T_2W scan. Fluid (F) in the sinus has an intermediate signal on PW and a hyperintense signal on T_2W MR scans.

Magnetic Resonance Angiography and Venography

Magnetic resonance angiography (MRA) and venography (MRV) are noninvasive techniques used to depict blood vessels in a projective format similar to conventional invasive angiography, without the use of ionizing radiation or iodinated contrast materials (Figure 31–6, A and B). On standard SE pulse sequences, rapidly moving blood will usually yield a signal void (dark) regardless of the orientation of the plane of the image. Therefore, MRA must employ

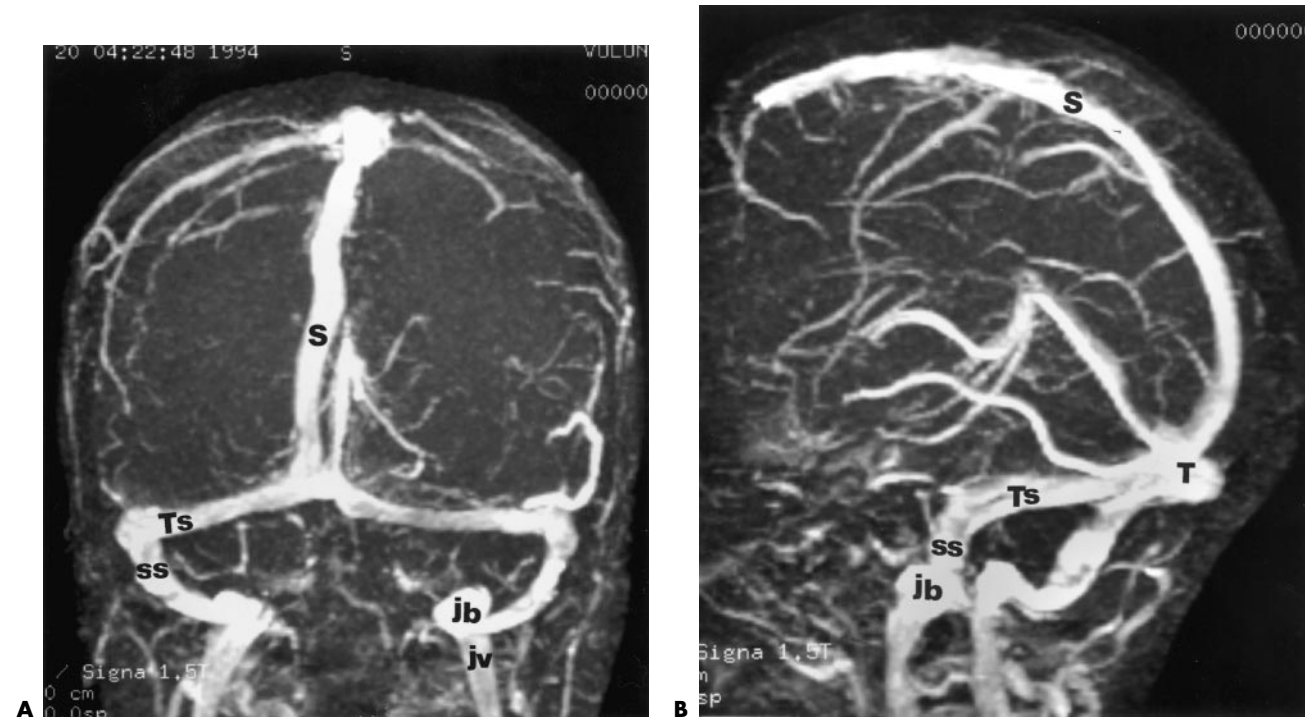


FIGURE 31–6. A, Magentic resonance venogram: two-dimensional time-of-flight coronal venography shows normal venous structures. S = superior sagittal sinus; Ts = transverse sinus; ss = sigmoid sinus; jb = jugular bulb; jv = internal jugular vein. *B*, Sagittal magnetic resonance venogram shows normal venous structures. S = superior sagittal sinus; T = torcula herophili; Ts = transverse sinus; ss = sigmoid sinus; jb = jugular bulb.

certain RF pulse sequences to produce signal from flowing blood. In MRA, the appearance of flowing blood depends on two main flow-related effects: time-of-flight (TOF) and phase contrast (PC).

Technical advances, such as the development of the two-dimensional and three-dimensional gradient-echo (GRE) pulse sequences (Figure 31–7), led to more widespread use of MRA and MRV. These noninvasive techniques of studying blood flow not only permit the evaluation of the vascular anatomy of the head and neck but also provide information on flow direction, velocity, and blood vessel patency. Both TOF and PC MRA can be obtained as a series of two-dimensional images or as a three-dimensional data set. This leads to the logical terminology of two-dimensional TOF, three-dimensional TOF, two-dimensional PC, and three-dimensional PC for the two fundamental approaches to vascular MRI techniques.

Clinical application best suited for two-dimensional TOF MRA includes mapping dural venous sinuses, evaluating the vascular structures of the neck for evidence of thrombosis or invasion by tumor, evaluating for carotid or vertebral dissection, and assessing the draining veins of arterial venous malformations. Two-dimensional TOF MRV can be used to evaluate suspected intracranial venous thrombosis; however, thrombus may appear bright (as a result of the short T_1 of methemoglobin) and can be confused for a patent venous sinus, resulting in a false-negative examination.[21]

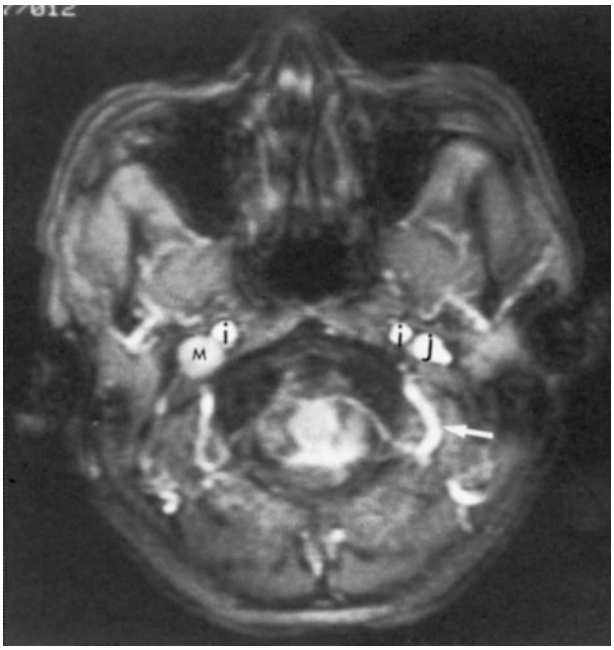


FIGURE 31–7. Schwannoma of the jugular fossa extending into the neck. A GRASS (gradient-recalled acquisition in the steady state) image shows the hyperintense internal carotid artery (i), internal jugular vein (j), vertebral artery (*arrow*), and hyperintense mass (M) within the right jugular vein.

Standard Film Radiography versus Computed Tomography and Magnetic Resonance Imaging

The modern diagnostic radiology department is equipped with a variety of medical imaging systems. Each has advantages and limitations. Plain film radiography is simple, fast, and inexpensive and produces sharp images of high-contrast structures (air, fluid, bone) (Figures 31–8 to 31–15); however, high-contrast objects can obscure low-contrast objects, and the presence of overlying objects can sometimes be misleading.

Tomography

The conventional radiographic films and screens are inherently nonuniform (less sensitive than the computer) and would limit subject-contrast discrimination (resolution) to possibly 3 to 5%, which is greater than the range for some soft tissue. Because of these limitations, standard tomography such as polytomography is no longer used if a CT unit is available.

Computed tomography and MRI allow the reconstruction of tissue contrast in true cross-sections (ie, without being influenced by nearby or distant tissues in the section), with excellent tissue-contrast discrimination (resolution).[2] Both modalities, and in particular MRI, provide the best anatomic cross-sectional imaging technology available to date. The new-generation high-resolution scanner, with a smaller pixel (picture element), the combination of thin section (1 to 1.5 mm) and extended bone range possibility (expansion of the HU scale up to 4,000 units, ie, +3,000 to –1,000), provides clear definition and differentiation of soft tissue structures while also providing superior bone detail to complex motion tomography (Figures 31–16 to 31–18). Dynamic CT provides excellent definition of vascular structures and can differentiate vascular tumors such as chemodectomas from

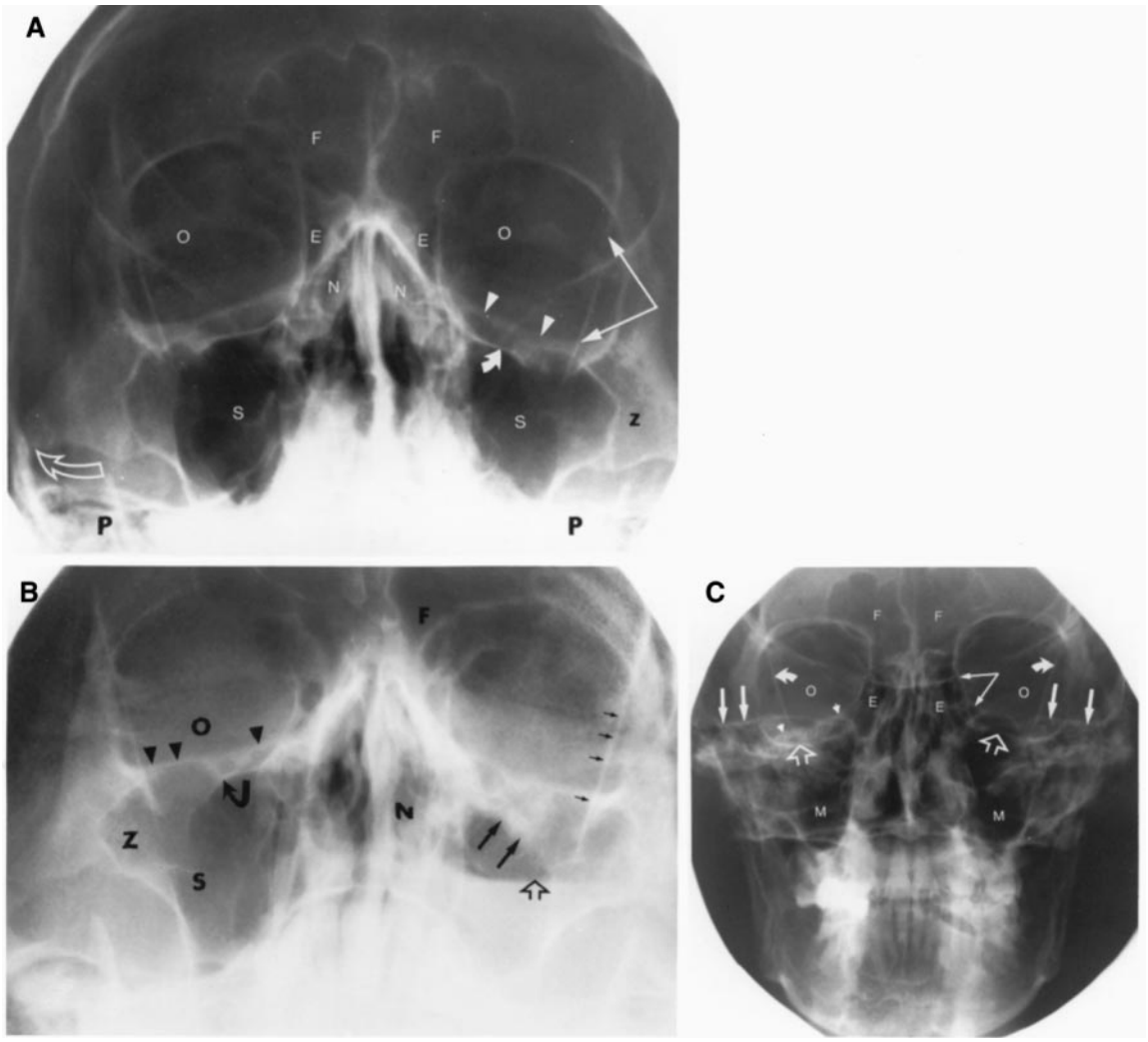


FIGURE 31–8. *A*, Waters' view showing maxillary sinuses (S), frontal sinuses (F), nasal cavities (N), body of the zygoma (Z), orbits (O), inferior orbital rim (*arrowheads*), floor of the orbit (*curved arrow*), temporal line (*crossed arrows*), zygomatic arch (*curved hollow arrow*), and petrous pyramid (P). *B*, Waters' view shows blow-out fracture (depressed floor) of the left orbit (*arrows*) with associated air-fluid level (*hollow arrow*). Note soft tissue swelling of the eyelid and cheek. F = frontal sinus; N = nasal cavity; O = orbit; S = maxillary sinus; Z = zygoma. *Arrowheads* point to the inferior orbital rim, the *curved arrow* points to the normal floor of the right orbit, and *short arrows* refer to the innominate line. *C*, Caldwell's view showing the maxillary (M), ethmoid (E), and frontal (F) sinuses. Notice orbits (O), inferior orbital rim (*short arrows*), orbital floor (*hollow arrows*), lamina papyracea (*crossed arrows*), and innominate lines (*curved arrows*).

less vascular lesions such as neurogenic tumors (Figure 31–19).[22–24]

Magnetic Resonance Imaging

One of the advantages of MRI is the ability to obtain transverse (axial) and direct coronal, sagittal, and oblique images without needing to change the position of the patient. The contrast relationships of different tissues on MRI are much different than on CT scanning (see Figure 31–17). The technical factors of various pulse sequences to obtain an MR image, in addition to physical and biochemical composition of the sample, play a significant role in tissue contrast in MR images. Water, such as in vitreous cavity (see Figure 31–17, B), has long T_1 and T_2 relaxation times and would appear as hypointense (dark) in T_1W and PW (proton-density) (see Figure 31–17, B) images and as hyperintense (bright) in T_2W MR scans. The lens has an

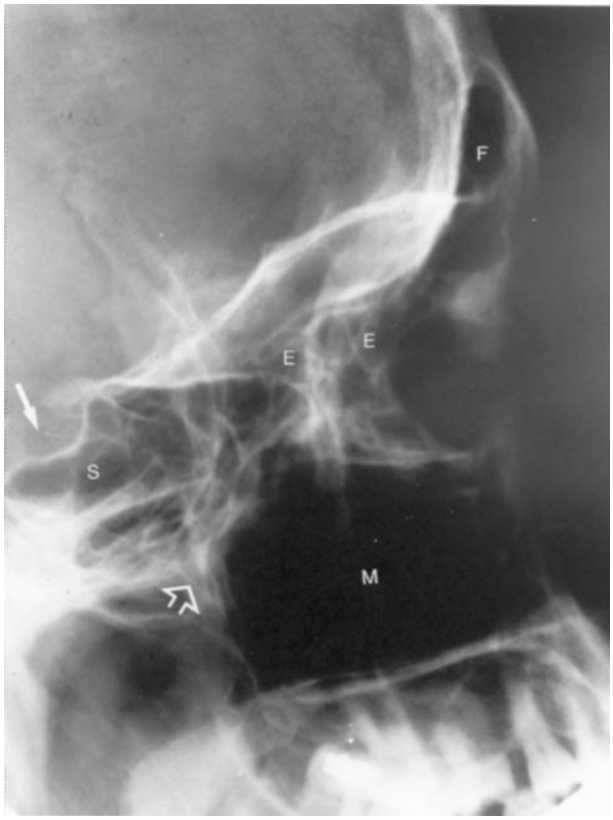


FIGURE 31–9. Lateral view showing the sphenoid sinus (S), sella turcica (*arrow*), pterygoid plates (*hollow arrow*), maxillary sinus (M), ethmoid air cell (E), and frontal sinus (F).

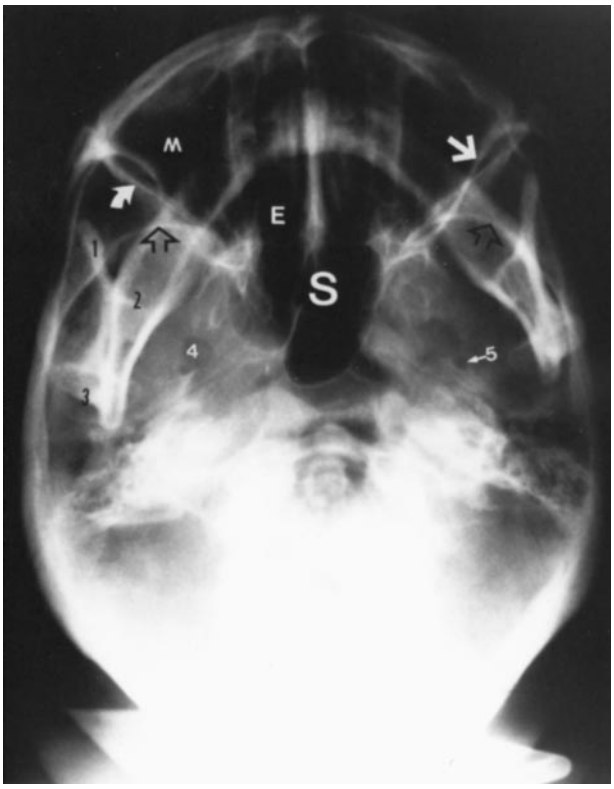
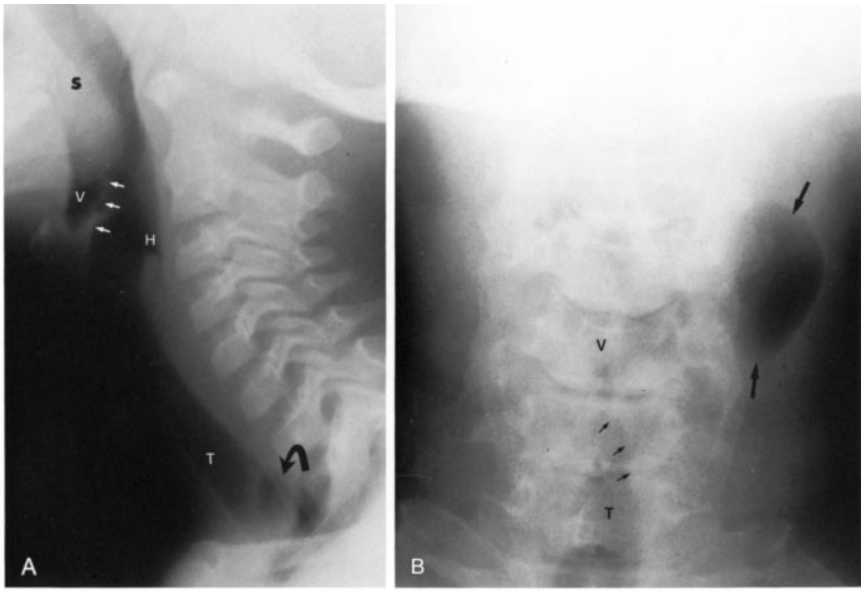


FIGURE 31–10. Basal view showing the sphenoid (S) and ethmoid (E) sinuses. Notice the "three lines": sinus line (*arrow*), orbital line (curved arrow), and middle cranial fossa line (hollow arrow). Note the maxillary sinus (M), coronoid process (1), mandible (2), mandibular condyle (3), foramen ovale (4), and foramen spinosum (5).

FIGURE 31–11. *A*, Normal lateral view of the neck, showing the soft palate (S), valleculae (V), epiglottis (*arrows*), hypopharynx (H), trachea (T), and apex of the lung (*curved arrow*). *B*, Frontal view of the neck, showing the trachea (T), subglottic space (*small arrows*), air in the laryngeal vestibule (V), and a large laryngocele (*large arrows*).



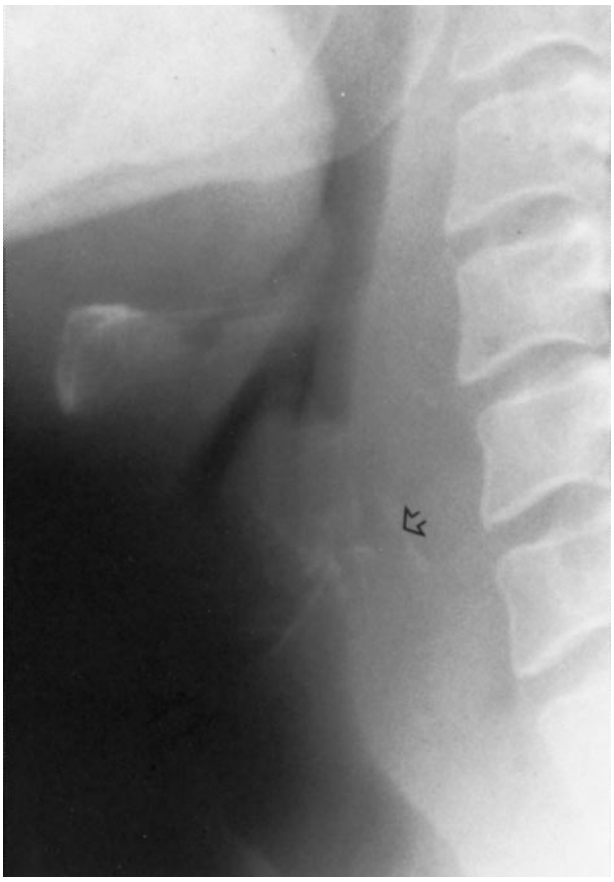


FIGURE 31–12. Fish bone and edema of the retropharyngeal space. Lateral neck, showing marked increased thickness of the prevertebral space (compare with normal Figure 31–11, A). Note the fish bone (*arrow*). Anterior to the fish bone are partially calcified cricoid and thyroid cartilages.

intermediate signal intensity (gray) in T_1W and PW scans (see Figure 31–17, B) and appears as a hypointense image in T_2W scans.

Fat has short T_1 and intermediate T_2 relaxation times and would therefore appear as a hyperintense image in T_1W and PW (see Figure 31–17, B) scans and as a hypointense image in T_2W scans. In the SE fat suppression technique, T_1W MR images can be obtained so that fat appears hypointense. This pulse sequence is commonly used in head and neck regions to suppress the signal of fat in T_1W MR images, allowing better delineation of pathologic tissues adjacent to normal fat. Tissues of intermediate structure (muscles) have an intermediate signal on T_1W and PW scans and a low signal on T_2W MR scans (see Figure 31–17, B). Solid structures such as cortical bone, enamel, dentin, and fibrous tissues (sclera, tendons, fasciae) are hypointense (dark) on T_1W, PW, and T_2W scans because of their sparsity of mobile hydrogen atoms. The cortical bone will appear black, whereas the medullary cavity will be bright because of signals from fat content. Air in paranasal sinuses, because of the lack of hydrogen atoms, appears dark in all MR scans. In SE pulse sequences, vascular structures such as carotid arteries and internal jugular veins are identified readily as low–signal intensity (dark) structures. With normal rapid blood flow, excited spins are carried away from the selected imaging section before a signal (echo) is formed (signal void phenomenon), creating a dark image with signal intensity similar to that of air, cortical bone, enamel, and dentin. In fast spin-echo (FSE) T_2W MR images, the fat, unlike conventional SE T_2W MR images, will be hypertense. However, FSE pulse sequences can be obtained with a fat suppression technique in which the fat appears hypointense, similar to conventional SE T_2W MR images.

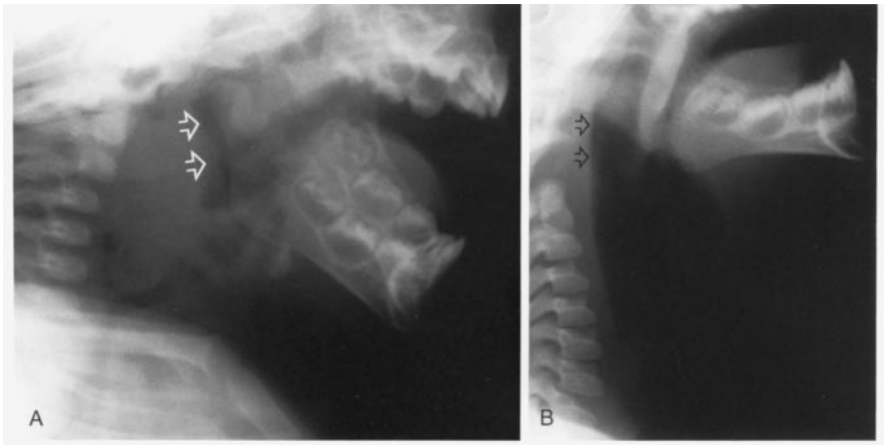


FIGURE 31–13. Normal neck of an infant. Lateral neck taken with flexion during expiration (*A*) and with head in extension (same patient) during inspiration (*B*). In *A*, there is bulging of the retropharyngeal soft tissue (*arrows*), mimicking retropharyngeal abscess or cellulitis. In *B*, the thickness of the retropharyngeal soft tissue is normal (*arrows*).

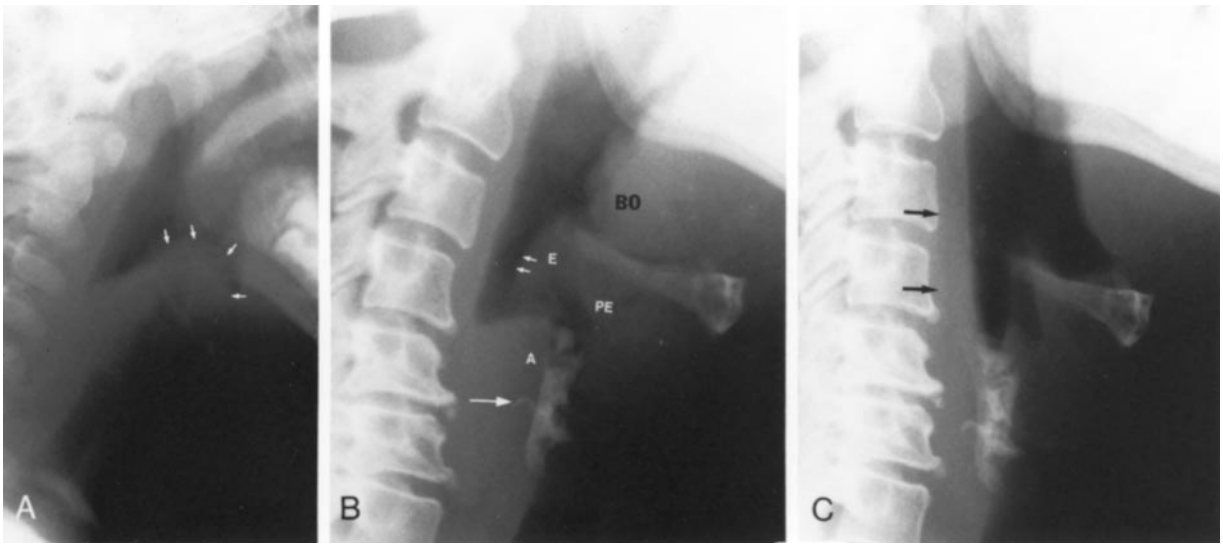


FIGURE 31–14. Epiglottitis and angioedema. *A*, Lateral neck film, showing marked enlargement of the epiglottis (*arrows*). Notice marked thickening of the aryepiglottic folds. *B*, Lateral neck in another patient with angioedema showing thickening of the epiglottis (E), aryepiglottic folds (*small arrows*), arytenoid region (A), pre-epiglottic space (PE), and base of the tongue (BO). The *large arrow* points to the cricoid cartilage. *C*, Lateral neck of the same patient as in *B*, showing marked resolution of the edema. There is now slight retropharyngeal soft tissue thickening (edema) (*arrows*).

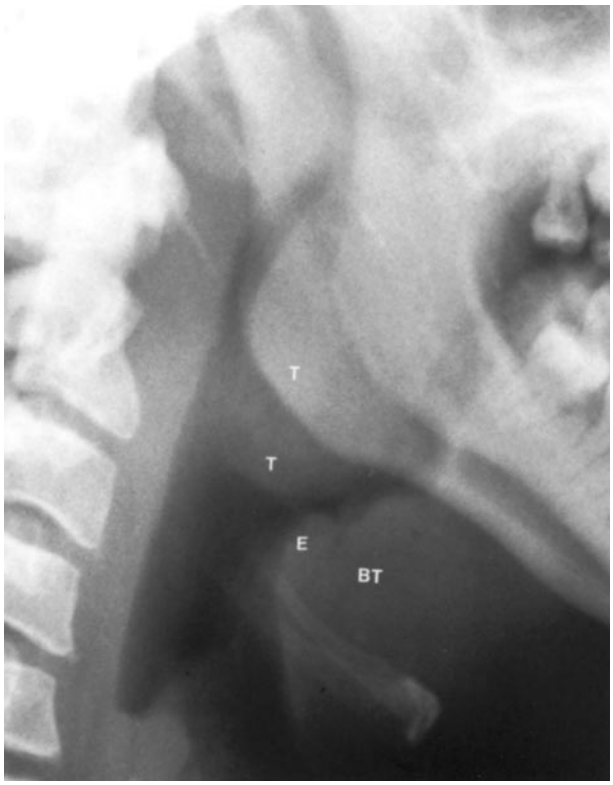


FIGURE 31–15. Tonsillitis and retropharyngeal abscess. Lateral neck shows large tonsils (T) with edema of the epiglottis (E) and base of the tongue (BT). The prevertebral soft tissue is thickened as a result of early cellulitis.

Dynamic Magnetic Resonance Imaging

Fast-scan or narrow flip angle MRI (dynamic MRI) is performed using the GRE pulse sequence such as "GRASS" technique. The GRASS, which stands for gradient-recalled acquisition in the steady state, is a new imaging technique. Its prime characteristic is the ability to provide T_1, T_2, and PW images in a short time. A striking feature of GRASS images is the high-intensity appearance of vascular structures, "MRA," especially when vessels transect the image slice (see Figure 31–7). The GRASS images are obtained with very short TR (25 to 100 ms), short TE (8 to 15 ms), and a variable flip angle (10 to 60 degrees).

Tissue Characterization of T_1- and T_2-Weighted Images

Damadian and coworkers showed that malignant tissues, except melanotic tumors, have a long T_1 and T_2 relaxation time.[25] Nonmelanotic tumors, therefore, are hypointense on T_1W scans and would appear as hyperintense images on T_2W scans. Most melanotic lesions on T_1W and PW MR scans would appear relatively hyperintense and become hypointense on T_2W MR scans.[26,27] Acute hemorrhage appears isointense to slightly hypointense on T_1W scans and would appear very hypointense on

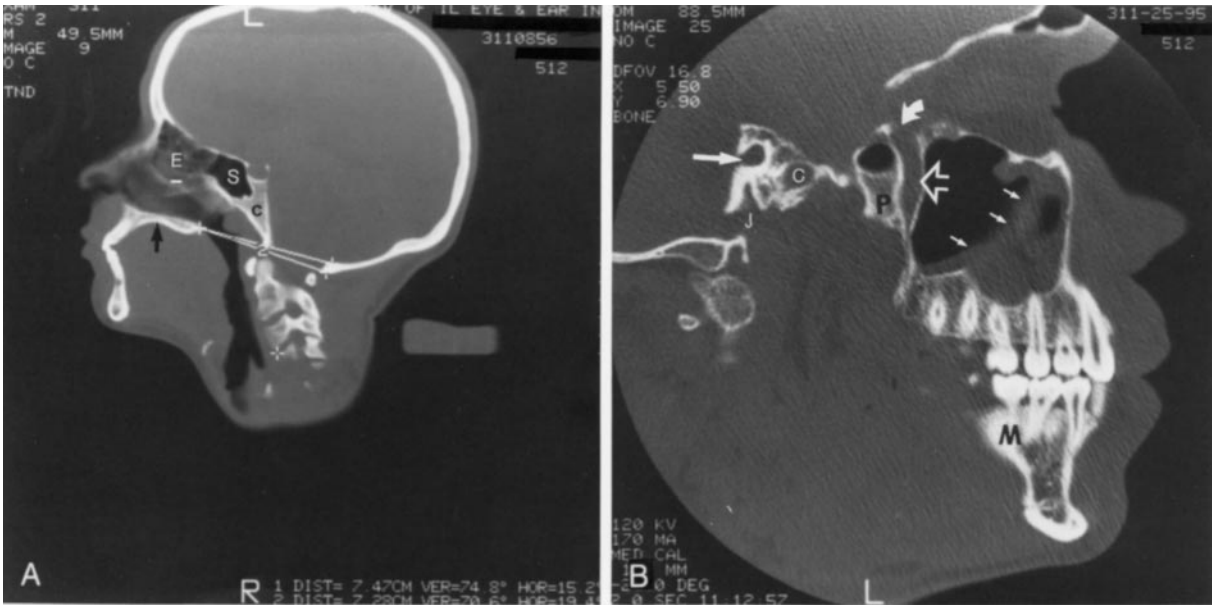


FIGURE 31–16. *A*, Direct sagittal computed tomographic (CT) scan of the midline head, showing the hard palate (*arrow*), sphenoid sinus (S), clivus (C), ethmoid air cells (E), and craniocervical junction. *B*, Sagittal CT scan showing mucosal thickening of the maxillary sinus (*short arrows*), pterygomaxillary fissure (*hollow arrow*), pterygoid plate (P), pterygopalatine fossa and inferior orbital fissure (*curved arrow*), carotid canal (C), jugular fossa (J), internal auditory canal (*long arrow*), and mandible (M).

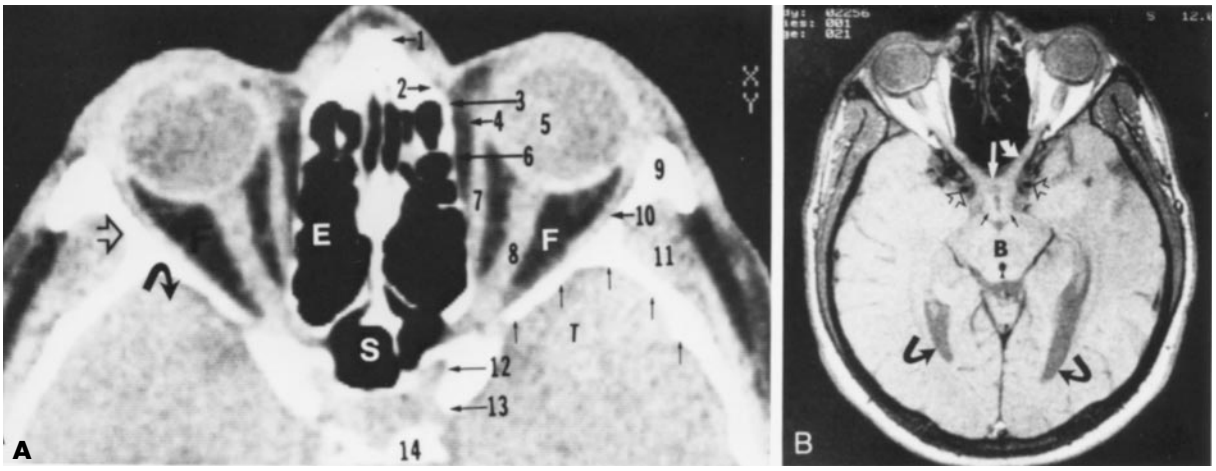


FIGURE 31–17. *A*, Normal ethmoid: axial computed tomographic (CT) scan through the plane of optic nerves, showing the nasal bone (1), nasal process of the maxilla (2), lacrimal bone (3), periorbital fat (4), globe (5), lamina papyracea (6), medial rectus muscle (7), optic nerve (8), zygomatic process of the frontal bone (9), lateral rectus muscle (10), temporal fossa (11), temporal line (*curved arrow*), cranial opening of the optic canal (12), anterior clinoid process (13), and posterior clinoid process (14). Note ethmoid (E) and sphenoid (S) sinuses and retroglobar fat (F). Two of the "triple lines" seen in the submentovertical projection (see Figure 31–10), the orbital line (posterior wall of the orbit) (*open arrow*) and middle cranial fossa line (*arrows*), are seen. *B*, Proton-weighted axial magnetic resonance scan showing optic nerves, optic chiasm (*long white arrow*), hypothalamus (*short black arrows*), brainstem (B), and cerebral aqueduct (*arrowhead*). Note visualization of the intracanalicular segment of the optic nerve (*curved white arrow*) that is hardly seen in the CT scan. Notice that in this proton-weighted image, the water in the vitreous cavity and cerebrospinal fluid in the lateral ventricles (*curved black arrows*) are hypointense and fat in the retrobulbar space and is hyperintense in the subcutaneous space. The blood vessels are hypointense (*hollow arrows*).

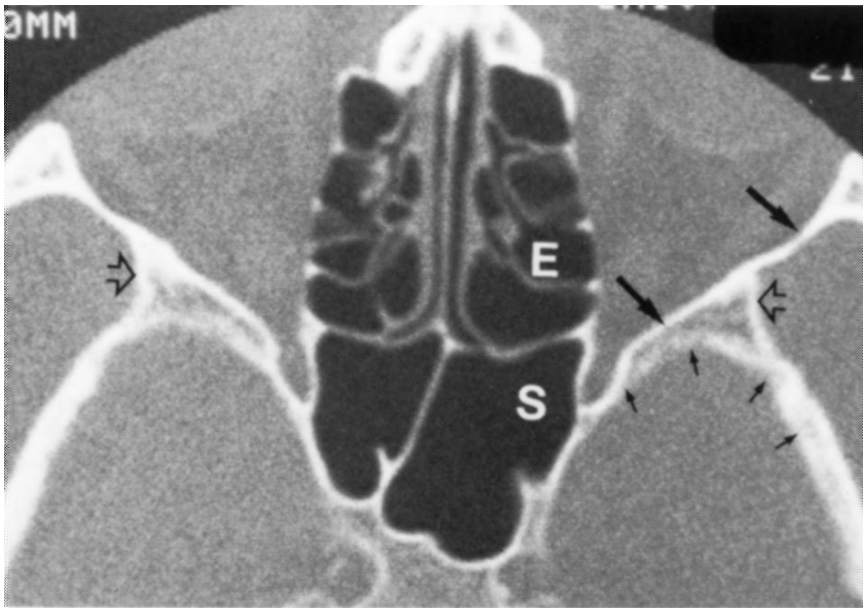


FIGURE 31–18. Normal ethmoid and sphenoid sinuses. High-resolution axial computed tomographic (CT) scan at the level of the middle of the orbit, using the extended bone CT range technique showing the ethmoid labyrinth (E) with their intact and sharply delineated septa and sphenoid sinuses (S). Note the sphenoethmoidal plate, intersphenoid septum, and sharply delineated orbital (*long arrows*) and anterior middle cranial fossa (*short arrows*) lines. Part of the innominate lines are also well visualized (*hollow arrows*). The innominate line (see Figure 31–10) is related to the outer cortex of the greater wing of the sphenoid bone.

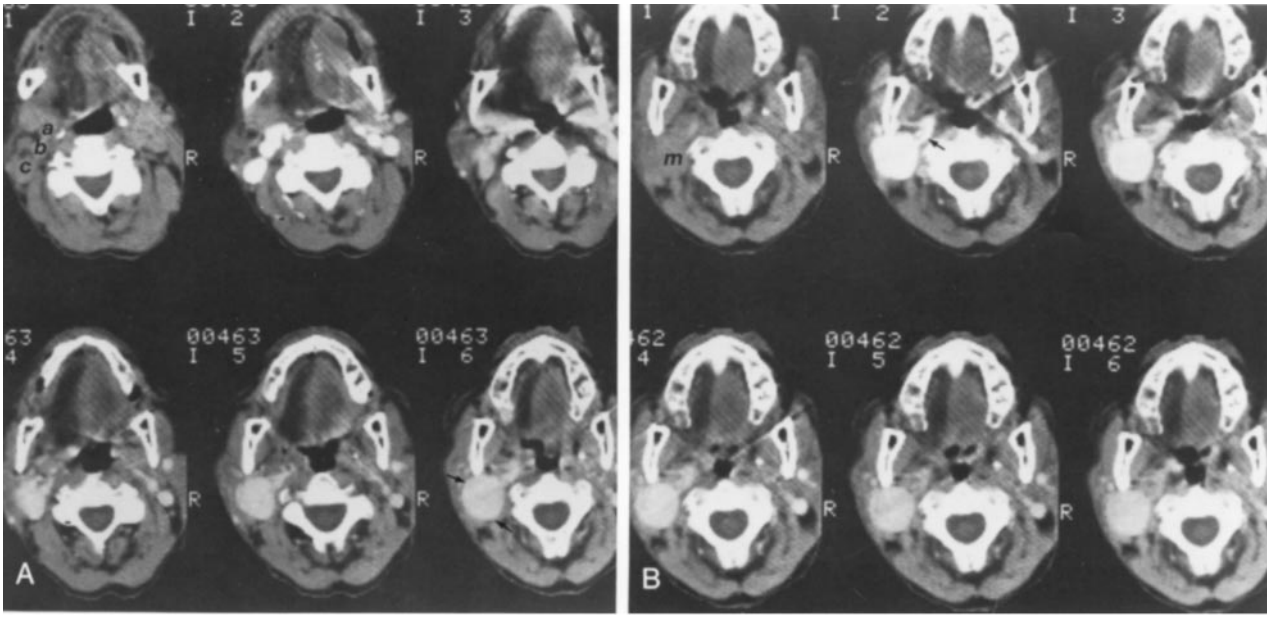


FIGURE 31–19. *A*, Glomus vagale tumor. Incremental dynamic computed tomographic (CT) scan with the first section obtained at the level of the tongue and extending superiorly (with automatic table incrementation) to the level of the superior alveolar ridge in this patient with a large right neck mass. Notice sequential enhancement of the external (a) and internal (b) carotid arteries and the internal jugular vein (c). A rounded enhancing mass is recognized on sections 5 and 6 (*arrows*), which show a prominent feeding vessel (see *B*). *B*, Glomus vagale tumor. Dynamic CT scan performed at the level of the right neck mass in the same patient in *A*. Note intense enhancement of a paravertebral mass (m) in the second section and rapid washout of contrast in the remainder of the sections. The glomus tumors often behave like an arteriovenous malformation in CT dynamic study. Note the large artery (most likely the ascending pharyngeal artery) (*arrow*) feeding this tumor. Note also the marked atrophy of the right side of the tongue in this patient with twelfth nerve palsy caused by superior extension of the tumor into the posterior cranial fossa and with associated destruction of the hypoglossal canal. Reproduced with permission from Mafee MF et al.[24]

T_2W MR scans. Chronic blood would appear hyperintense on T_1W, PW, and T_2W MR scans. Dense calcifications would appear hypointense on T_1W, PW, and T_2W MR scans, and proteinaceous fluid such as a thick mucosal secretion and exudate would appear hyperintense on T_1W, PW, and T_2W MR scans. Highly proteinaceous fluid may appear hypointense on T_2W MR images. Inspissated mucus may appear very hypointense on T_2W MR scans.

Positron Emission Tomography

Positron emission tomography (PET) using 2-^{18}F-fluoro-2-deoxy-D-glucose (FDG) has come to play a definite role in head and neck oncology.[28] Positron emission tomography depends on the altered metabolic activity of the tumoral tissue, of which increased glycolysis is a hallmark. 2-^{18}F-fluoro-2-deoxy-D-glucose, a glucose analog, is trapped preferentially within tumor cells with increased glycolysis, allowing for their detection. As a metabolic imaging modality, FDG-PET has been shown to have a significant impact in the diagnosis and management of head and neck malignancies. It is a powerful tool in the assessment of recurrence of malignancies and seems to be more effective and reliable than anatomic imaging techniques for the evaluation of head and neck tumors.

COMPUTED TOMOGRAPHY AND MAGNETIC RESONANCE IMAGING OF ANATOMIC SITES

Normal Computed Tomography and Magnetic Resonance Imaging Anatomy of the Face and Neck

A complete description of the anatomy of the face and neck is discussed in this volume. Emphasis here is placed on the structures that are most pertinent to understanding CT and MRI scans in the broad field of otorhinolaryngology.

Nasal Cavities, Paranasal Sinuses, Orbits, and Nasopharynx

Nasal Cavities The paranasal sinuses and nasal cavities are closely related in their development. They are present at birth; however, the sphenoid sinuses are too small to be seen, and the frontal sinuses develop later from the anterior ethmoid air cells. Ethmoid and maxillary sinuses are the only sinuses that are large enough at birth to be seen on imaging studies (CT, MRI) and to be clinically significant in rhinosinusitis. The framework of the external nose is composed of bones and hyaline cartilages. Its bony framework consists of the nasal bones, the frontal processes of the maxillae, and the nasal part of the frontal bone. The posterior nasal apertures or choanae are two oval openings, each measuring about 2.5 cm in the vertical and 1.25 cm in the transverse direction. The choanae are best evaluated on axial CT and MRI scans (Figure 31–20). The roof, floor, and medial (septal) and lateral walls of each half of the nasal cavity can be best evaluated on coronal CT and MR scans (Figure 31–21). The roof of the nasal cavity is narrow from side to side. The lateral wall of the nasal cavity is very irregular owing to the presence of nasal conchae. The nasal conchae are best evaluated on coronal CT scans, projecting downward and slightly medially, each forming the roof of a passage or meatus (inferior, middle, and superior), which communicates freely with the nasal cavity (see Figure 31–21). The inferior concha is an independent thin bone that articulates with the nasal surface of the maxilla and the perpendicular plate of the palatine bone. The

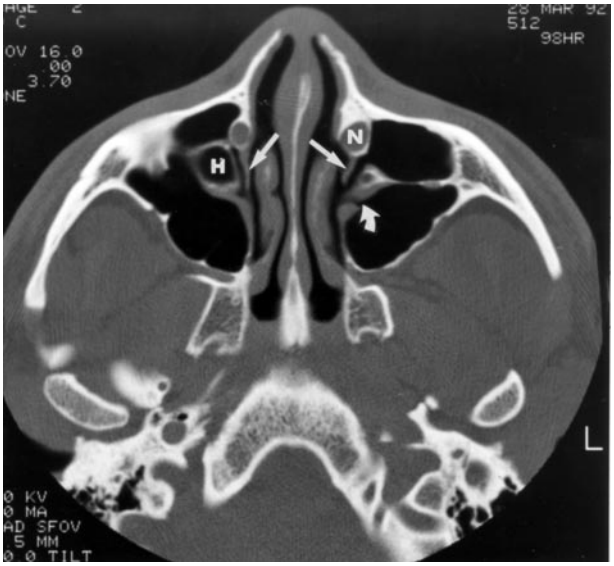


FIGURE 31–20. Axial computed tomographic scan shows the nasolacrimal canal (N), ostium of the maxillary sinus (*curved arrow*), Haller cells (H), and uncinate processes (*straight arrows*).

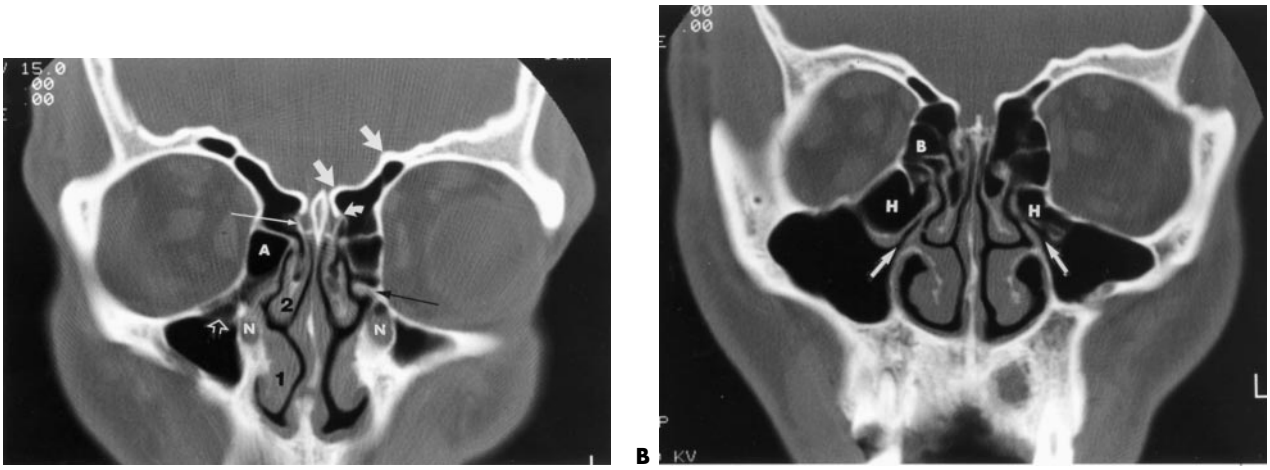


FIGURE 31–21. *A*, Coronal computed tomographic (CT) scan, same patient as in Figure 31–20, shows the inferior turbinate (1), middle turbinate (2), nasolacrimal ducts (N), uncinate process (*long black arrow*), Haller cell (*hollow arrow*), agger nasi cell (A), lateral lamella (*long white arrow*), fovea ethmoidalis (*double white arrows*), and canal for the ethmoidal artery (*curved arrow*). *B*, Coronal CT scan, same patient as in *A*, shows the bulla (B), Haller cells (H), and maxillary ostia (*arrows*), opening into the infundibulum.

inferior meatus receives the orifice of the nasolacrimal canal (see Figure 31–21, A). The middle and superior conchae are bony projections from the medial surface of the ethmoidal labyrinth (see Figure 31-21, A). The middle concha extends backward to articulate with the perpendicular plate of the palatine bone. The middle meatus receives the opening of the anterior and middle ethmoid air cells, frontal sinuses, and maxillary sinuses (Figure 31–21, A). The superior concha is a small curved lamina that lies above and behind the middle concha. The superior meatus receives the opening of the posterior ethmoidal air cells (Figure 31–22). A narrow interval, the sphenoethmoidal recess (see Figure 31–22), separates the superior concha from the anterior surface of the body of the sphenoid, through which the sphenoidal sinus opens into the nasal cavity (see Figure 31–22). A fourth concha (concha suprema) is

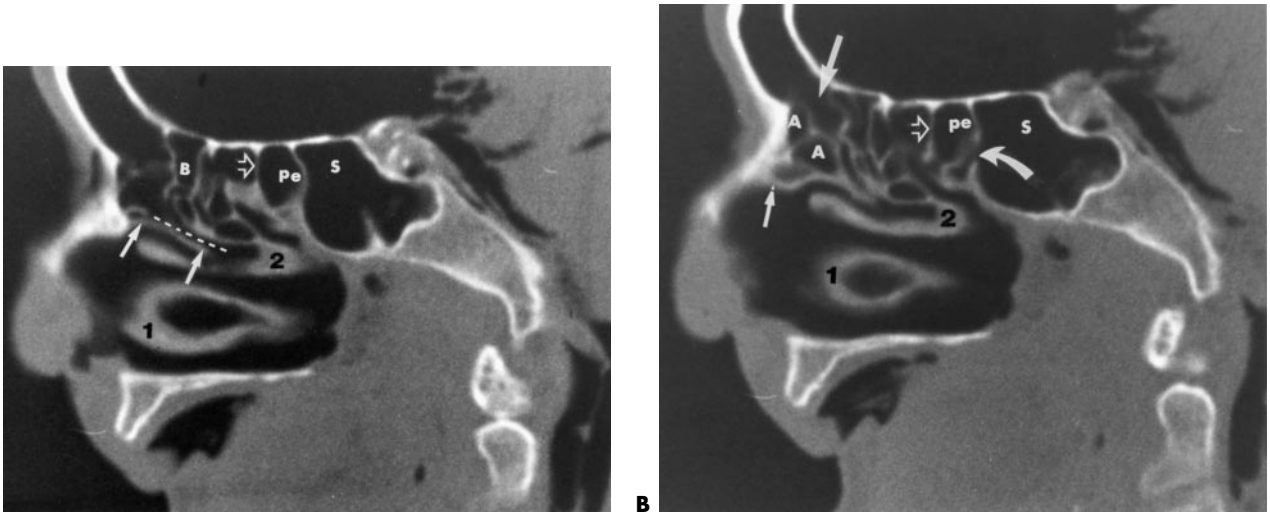


FIGURE 31–22. *A*, Sagittal computed tomographic (CT) scan shows the uncinate process of the ethmoid (*arrows*), infundibulum (*broken line*), bulla (B), basal lamella (*hollow arrow*), posterior ethmoid sinus (Pe), sphenoid sinus (S), and middle (2) and inferior (1) turbinates. *B*, Sagittal CT scan shows the inferior turbinate (1), middle turbinate (2), uncinate process (*small straight arrow*), agger nasi cell (A), frontal recess (*large straight arrow*), basal lamella (*hollow arrow*), posterior ethmoid sinus (pe), sphenoid sinus (S), and sphenoid sinus ostium (*curved arrow*).

often present on the medial surface of the ethmoidal labyrinth above and behind the posterior end of the superior concha. Immediately behind the superior meatus, the sphenopalatine foramen, which opens into the pterygopalatine fossa, pierces the lateral wall of the nasal cavity. It transmits the sphenopalatine artery, nasopalatine artery, and the nasopalatine and superior nasal nerves from the pterygopalatine fossa.

Maxillary Sinuses Maxillary sinuses begin to develop around the sixty-fifth day of gestation. The size of the sinus at birth is about 6 to 8 cm^3. Air can be seen on CT scans within the maxillary sinuses at or a few days after birth. Rapid expansion occurs from 7 to 18 years, related to eruption of the permanent teeth. They are the largest accessory air sinuses of the nose, and when fully developed, they occupy the body of the maxilla, and each is approximately 15 mL in volume (see Figures 31–20 and 31–21). The natural opening of the sinus is high above the floor and poorly placed for natural drainage. The middle meatus is of such a shape that pus running down from the frontal sinus or the anterior ethmoidal sinuses is directed by the hiatus semilunaris into the opening of the maxillary sinus, which may, in some cases, act as a secondary reservoir for pus discharged from these sinuses. Accessory ostia are present in 15 to 40% of cases, usually in the membranous medial sinus wall.

Ethmoid Sinuses Ethmoid cells begin to develop around the third to fourth month of fetal life, and there may be a few cells (four to five cells) present at birth. At birth, the size of the anterior ethmoid cells is approximately 5 mm high, 2 mm long, and 2 mm wide. The size of the posterior cells is 5 mm high, 4 mm long, and 2 mm wide. By the age of 12 years, the ethmoid air cells have almost reached their adult size (see Figure 31–22).

Because fat, fluid, and bone are natural contrast material, the orbit is the ideal structure for a highly detailed CT and MRI scan. A scan taken through the orbits shows the relationship of the ethmoid sinuses, sphenoid sinuses, pituitary fossa, cavernous sinuses, and temporal lobes with the orbit (see Figure 31–17). The cortical bones, which appear hyperdense on CT images (see Figure 31–17, A), are seen on MR as hypointense images because of little MR signal from them (see Figure 31–17, B). The bone marrow appears on T_1W MRI as a hyperintense image owing to the presence of fat. The air is seen as a dark image on both CT (low electron density) and MR (low proton density) images (see Figure 31–17).

The ethmoid sinuses are positioned between the orbits and comprise approximately 10 to 12 cells within the labyrinth of the ethmoid bone. The fine bone septa of the ethmoid labyrinth, the lamina papyracea, olfactory groove, cribriform plate and its lateral lamella, fovea ethmoidalis, canals for ethmoidal arteries, and maxillofacial bony structures will be best imaged by extended-bone window high-resolution CT (Figures 31–18, 31–22, and 31–23).[29,30] The anatomic landmarks of the ethmoid bone and ostiomeatal unit, such as the uncinate process, infundibulum, middle meatus-ethmoid complex, ethmoid bullae, middle turbinate, and cribriform plate (as well as the ostium of the maxillary sinuses), are best evaluated on coronal as well as axial and reformatted sagittal CT images (see Figures 31–20 and 31–21).[31–34] Just anterior to the superior attachment of the middle turbinate and anterior to the frontal recess (frontonasal duct) are the agger (ridge) nasi (a rudimentary turbinate) and agger nasi cells (see Figure 31–22). The agger nasi cells can invade the lacrimal bone or the ascending process of the maxilla. Just posterior and inferior to the agger nasi cells lies the ethmoidal uncinate process, a thin curved bar of bone from the lateral side of the ethmoidal labyrinth that forms a portion of the lateral nasal wall. It projects downward and backward and is subject to considerable variation in size. It ranges in height from 1 to 4 mm and is 14 to 22 mm long.[31] Anteriorly, it articulates with the lacrimal bone and ethmoidal process of the inferior nasal concha (see Figure 31–21). The superior edge of this process is free and forms the medial boundary of the hiatus semilunaris (see Figure 31–23). The ethmoidal infundibulum is a trough-shaped cavity (cleft) that is below the bulla ethmoidalis and above and lateral to the uncinate process. The infundibulum is best visualized on coronal CT scans (see Figure 31–21). In more than 50%, the infundibulum continues superiorly as the frontonasal duct into the frontal sinus.[32] The basal (ground) lamella, which separates the anterior and posterior air cells, can best be evaluated on axial as well as coronal CT scans. An air space (cleft) is usually found between the basal lamella and the bulla, which may extend superiorly to the bulla. This is called the sinus lateralis. This sinus lateralis, unlike the other anterior ethmoidal air cells that open into

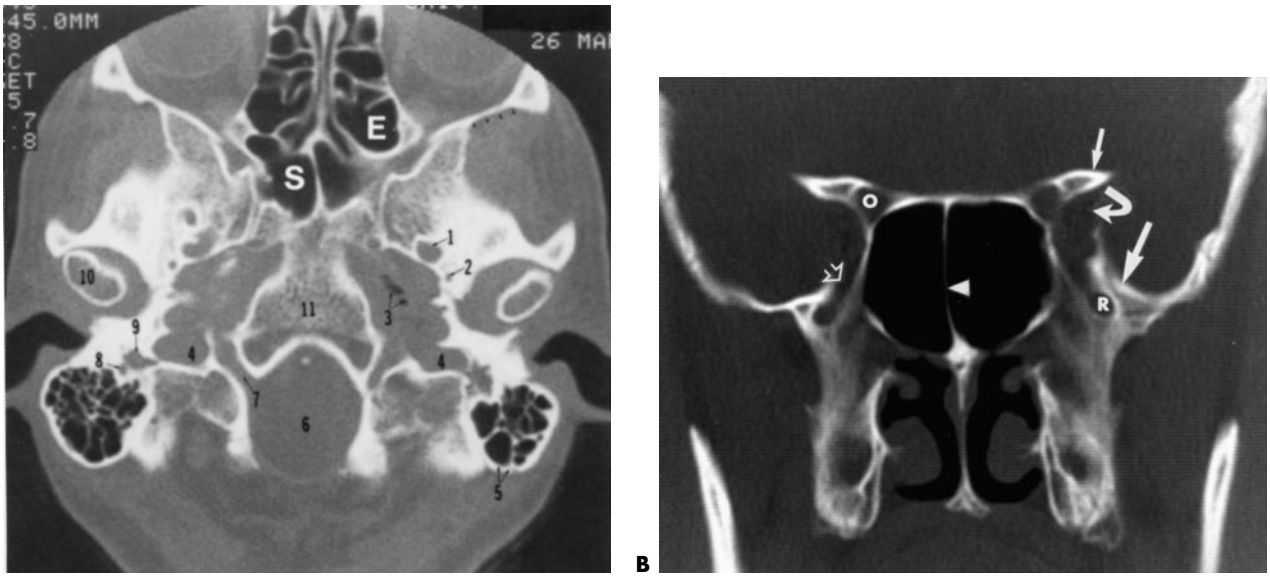


FIGURE 31–23. Normal base of the skull. *A*, High-resolution axial computed tomographic (CT) scan at the level of the clivus (11), using the extended-bone CT range technique showing the ethmoidal labyrinth (E), sphenoid sinus (S), posterior wall of the orbit (*arrowheads*), foramen ovale (1), foramen spinosum (2), air in the eustachian tube (3), jugular fossae (4), mastoid air cells (5), foramen magnum (6), hypoglossal canal (7), vertical portion of the facial nerve canal (8), lateral extension of the right jugular fossa (9), and mandibular condyle (10). *B*, Coronal CT scan shows the anterior clinoid (*small straight arrow*), superior orbital fissure (*curved arrow*), greater wing of the sphenoid (*large arrow*), foramen rotundum (R), intersphenoid septum (*arrowhead*), inferior orbital fissure (*hollow arrow*), and optic canal (O).

the infundibulum, may communicate with the frontal recess or open directly and independently into the middle meatus. Certainly, individual structural differences in ethmoidal and ostiomeatal complexes and other paranasal sinuses are to be expected, and the reader ought not be discouraged should certain illustrations in the literature fail to be identical with the images he or she will review for any individual patient. Certain anatomic variations are observed more commonly, and these are as follows: (1) concha bullosa, a pneumatized middle turbinate; (2) low position of the fovea ethmoidalis and dehiscence of the lateral lamella of the cribriform plate; (3) bulging of the optic canal into the posterior ethmoidal complex; and (4) deviation of the uncinate process. The superior edge of the uncinate process may elevate medially to obstruct the middle meatus and, more importantly, may deviate laterally to obstruct the infundibulum. Marked lateral deviation (commonly seen in patients with a hypoplastic maxillary sinus) or even fusion of the uncinate process to the medial orbital wall may endanger the orbit. (5) Haller cells are ethmoidal air cells extending along the medial floor of the orbit (infraorbital ethmoid air cells) (see Figure 31–21, B). (6) Onodi cells are posterior ethmoidal air cells, encroaching into the sphenoid sinus. (7) An asymmetric intersphenoid septum or septae are important because the posterior extension of this partition usually marks the location of the internal carotid artery groove (sulcus), which at times may be dehiscent, resulting in bulging of the internal carotid artery into the sphenoid sinus. Other anatomic variations include deviation of the nasal septum, paradoxical middle turbinate, uncinate process pneumatization (bulla), pneumatized superior turbinate, pneumatized cristal galli, congenital dehiscence or post-traumatic deformity of the medial wall or floor of the orbit, and exposed inferior orbital nerve within the maxillary sinus.

Frontal Sinuses Developmentally, the frontal sinus begins during the fourth month of fetal life in the region of the frontal recess. At birth, it is indistinguishable from the anterior ethmoid cells. Postnatal growth is slow. It is visible after the first or second year of life and is well developed by 7 to 8 years. Its size increases until the late teens. Each frontal sinus drains into the middle meatus of the nasal cavity via the frontonasal duct, best visualized on sagittal CT scans (see Figure 31–22). This communication between the

frontal sinus and nasal cavity is not strictly a duct but an internal channel positioned between the sinus and the anterior part of the middle meatus, referred to as the frontal recess. The natural ostium of the frontal sinus can be found at the superior-anterior end of the infundibulum. When the frontal recess air cells and agger nasi cells in this area are carefully removed, the ostium will come into view.

Sphenoid Sinuses Developmentally, the sphenoid sinuses begin during the third fetal month as evaginations of mucosa of the nasal cavity. At birth, the sinuses remain small and are little more than evaginations of the sphenoethmoid recess. After the fifth year, invasion of the sphenoid bone is more rapid, and by the age of 7, the sinus extends posteriorly to the level of the sella turcica. Further enlargement occurs after puberty. The pneumatization of sphenoid bone may be extensive. The sinuses may extend posteriorly to the clivus, anteriorly to the tuberculum sella and planum sphenoidale, and out into the anterior clinoids (see Figures 31–22 and 31–23). Lateral extension results in pneumatization of the base of the pterygoid processes as well as part of the greater wing of the sphenoid. Each sinus drains into the nasal cavity. The sphenoidal ostium drains into the superior meatus in the sphenoethmoidal recess, above the superior concha, well above the floor of the sinus. On average, the ostium measures 2 × 3 mm and lies 10 mm above the floor of the sinus.

Use of Computed Tomographic Scan for Computer-Aided Endoscopic Surgery Functional endoscopic sinus surgery has become a frequently performed operation in patients with chronic rhinosinusitis.[35,36] A preoperative CT scan is obtained in all patients according to the protocol of the specific system used. The CT protocol requires contiguous, 1 to 3 mm-thick, nonoverlapping, axial sections. Images are transferred to the computer-aided surgery system by means of an optical disk or storage tape. The CT scan is performed using fiducials; therefore, correlation of the patient's anatomy with corresponding points on the CT images is accomplished by the registration process.

Pathology of Paranasal Sinuses and Nasal Cavity

The inherently superior contrast resolution of CT and the three-dimensional effect of the combined axial and coronal sections have established CT as the method of choice in the staging of maxillofacial neoplasms and the evaluation of chronic and complicated inflammatory processes of the paranasal sinuses and face.[32–34] Magnetic resonance imaging has been used to evaluate maxillofacial anatomy as well as pathology. Familiarity with CT dictates that the MR evaluation complements and does not substitute for the CT evaluation. Magnetic resonance imaging is superior to CT for evaluation of the extent of soft tissue involvement and intracranial invasion by sinonasal tumors. A detailed discussion of the various pathologic entities of the paranasal sinuses and nasal cavity is limited by the scope of this chapter, although certain conditions will be described to illustrate the role of CT and MRI in the evaluation of patients with paranasal sinus and nasal diseases.

Conventional plain films still can be used to provide the screening studies in various pathologic conditions of the sinuses and in particular in maxillofacial trauma and give orientation and direction to further examinations using CT.[33,34] Developmental anomalies, such as a palatomaxillary cleft, are more clearly seen using CT (Figure 31–24). The appearance of lesions of the paranasal sinuses and face on CT scans usually does not provide sufficient evidence for a specific histologic diagnosis; however, cystic (Figure 31–25), cartilaginous, and osseous tumors are an exception (Figure 31–26). Fibrous dysplasia (FD) also frequently shows characteristic

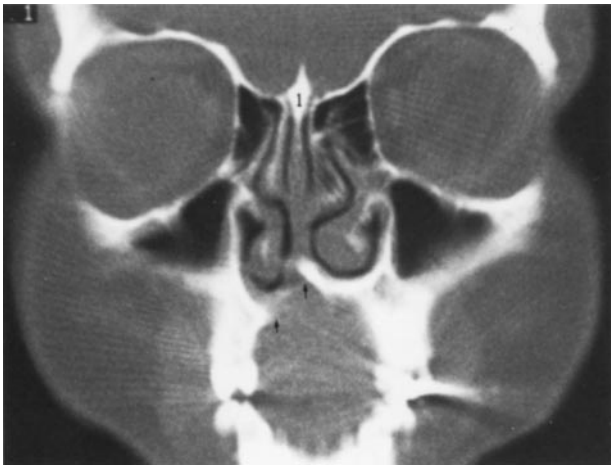


FIGURE 31–24. Coronal computed tomographic scan, showing a right palatomaxillary cleft (*arrows*). Note the crista galli (1), cribriform plate, roof of the ethmoid sinuses (fovea ethmoidalis and olfactory plates), lamina papyracea, and middle and inferior turbinates.

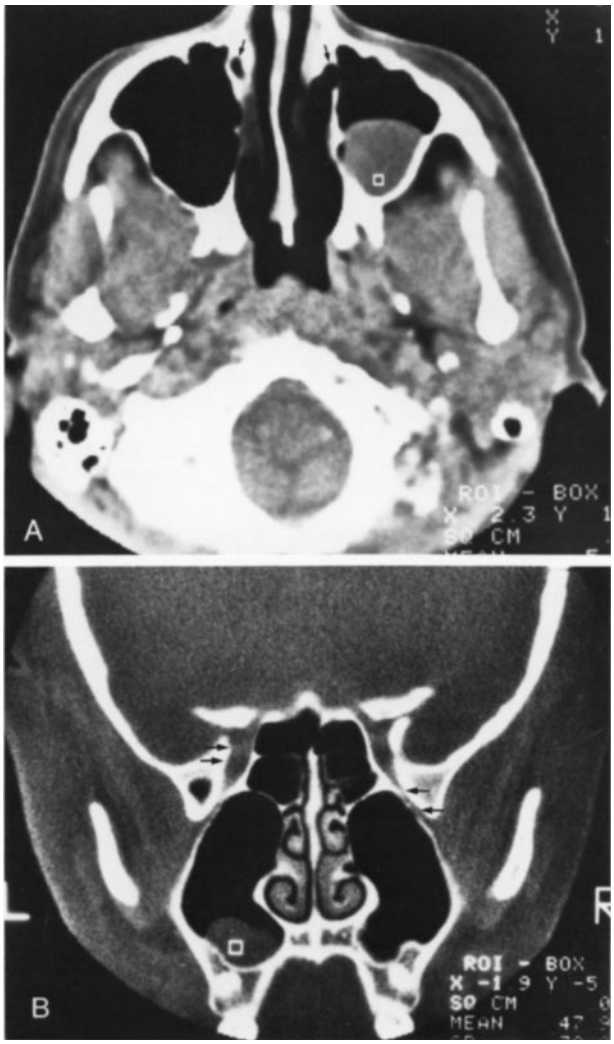


FIGURE 31–25. *A*, A computed tomographic (CT) scan showing a smoothly marginated soft tissue image with cursor measurement of water density (mean = 5) in the left maxillary sinus. This is characteristic of a retention cyst. Notice the nasolacrimal canals (*arrows*). *B*, Coronal CT scan showing a mucous cyst with cursor measurement of 47.9 HU on the floor of the left maxillary sinus. The mucous cysts usually show a high CT number compared with retention cysts. Note excellent bone details and visualization of the nasal turbinates, maxilloethmoid plate, and superior and inferior orbital fissures (*upper and lower arrows*).

islands of bone formation within a dense, rather uniform, stroma on a CT scan (Figure 31–27). Periosteal reaction and cortical break are not seen in FD.

Infection Conventional radiography is adequate for the diagnosis of acute sinusitis. Even though antibiotics have cut down on the incidence of complicated sinusitis with orbital involvement, it still occurs and may even be the first sign of sinus infection in children.[37] Infection can spread from the sinuses to the orbit by direct extension. It can also spread by way of numerous valveless communicating veins between the sinuses and the orbit.[37] In complicated sinusitis, CT is the best method to demonstrate the nature and the source of the problem. The orbital extension of sinusitis includes orbital periostitis, subperiosteal induration (phlegmon) (Figure 31–28), subperiosteal abscess (Figure 31–29), orbital cellulitis, orbital

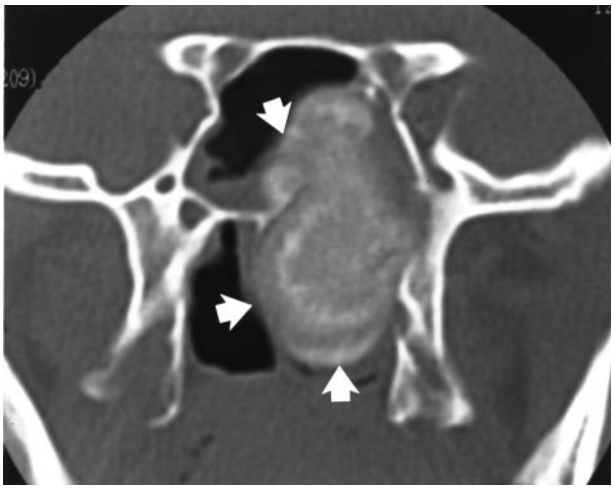


FIGURE 31–26. Ossifying fibroma. A coronal computed tomographic scan showing a large ossifying fibroma (*arrows*).

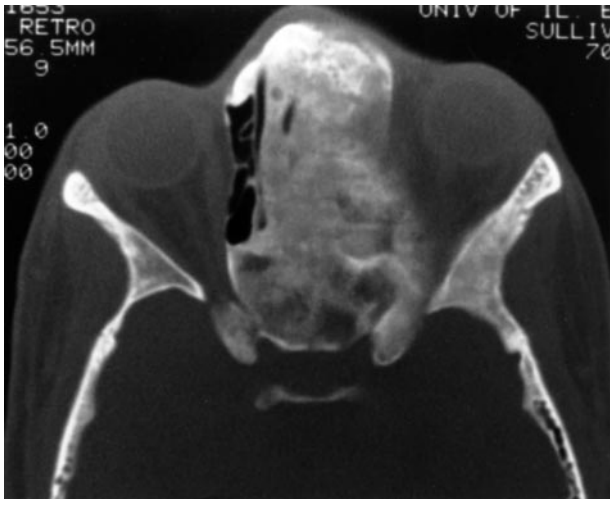


FIGURE 31–27. Fibrous dysplasia. Axial computed tomographic scan showing fibrous dysplasia involving ethmoid and sphenoid bones.

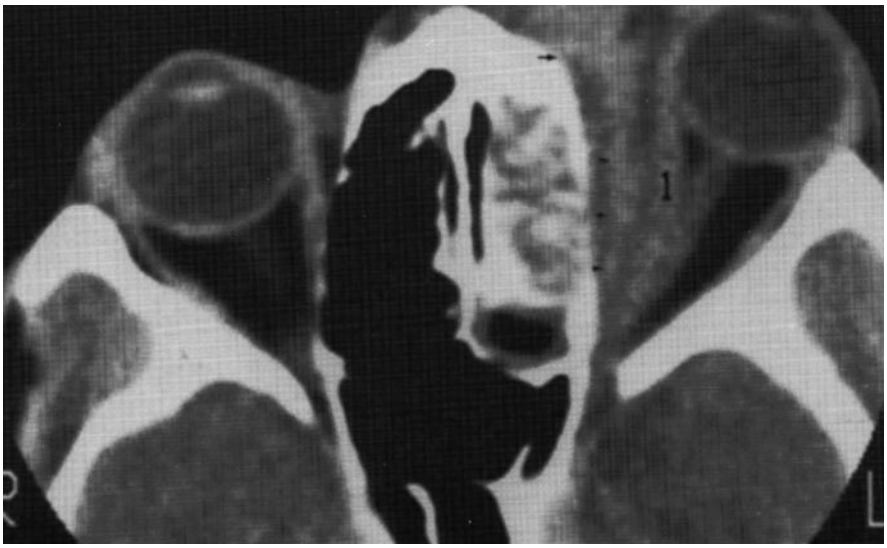


FIGURE 31–28. Ethmoid periostitis and subperiosteal phlegmon. Note soft tissue replacement of air in the left ethmoid labyrinth with thickening of the trabeculae. The left lamina papyracea is thickened and in some areas dehiscent. These findings are indicative of periostitis. Note soft tissue swelling (not yet leading to necrosis and abscess formation) in the left periorbita (*arrowheads*), resulting in displacement of the left medial rectus muscle (1). Notice proptosis of the left eye and marked periorbital edema and inflammation anterior to the orbital septum (*arrow*).

abscess, and ophthalmic vein thrombophlebitis (Figure 31–30).[37–41] Should the infection spread from the sinuses into the cranial cavity, one or more of the following complications may ensue: cavernous sinus thrombosis, meningitis, epidural (see Figure 31–30), subdural, or brain abscesses.[37–41] Intracranial complications are better evaluated by MRI than CT scanning (Figure 31–31).

Periostitis and osteomyelitis of the frontal bone can complicate frontal sinusitis.[37] An infection of the frontal sinus severe enough to involve the orbit may also extend through the posterior plate of the frontal sinus to involve the anterior cranial fossa (see Figure 31–30).[37] Orbital complications that result from sphenoid sinusitis can best be evaluated with CT and MRI scans (see Figure 31–31).

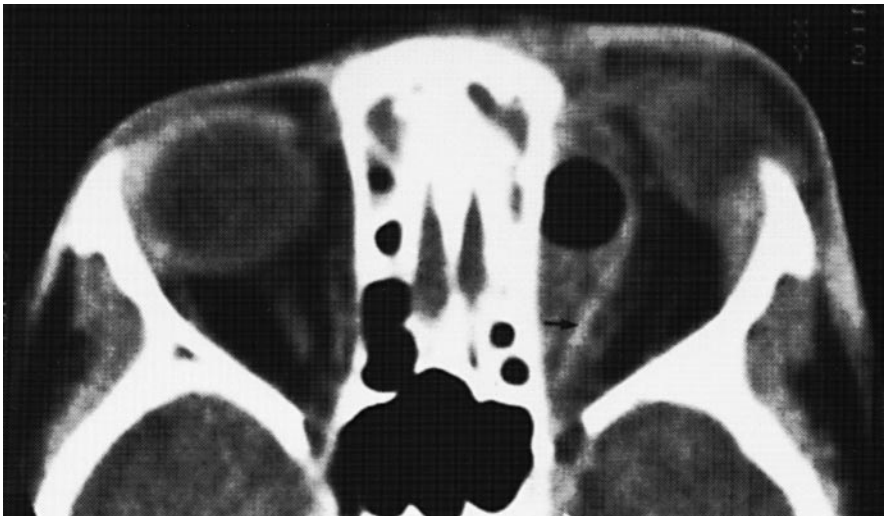


FIGURE 31–29. A computed tomographic scan showing proptosis of the left eye and a subperiosteal abscess. A large gas pocket is seen (*dark round area*) within the abscess. Note the displaced left medial rectus muscle (*arrow*).

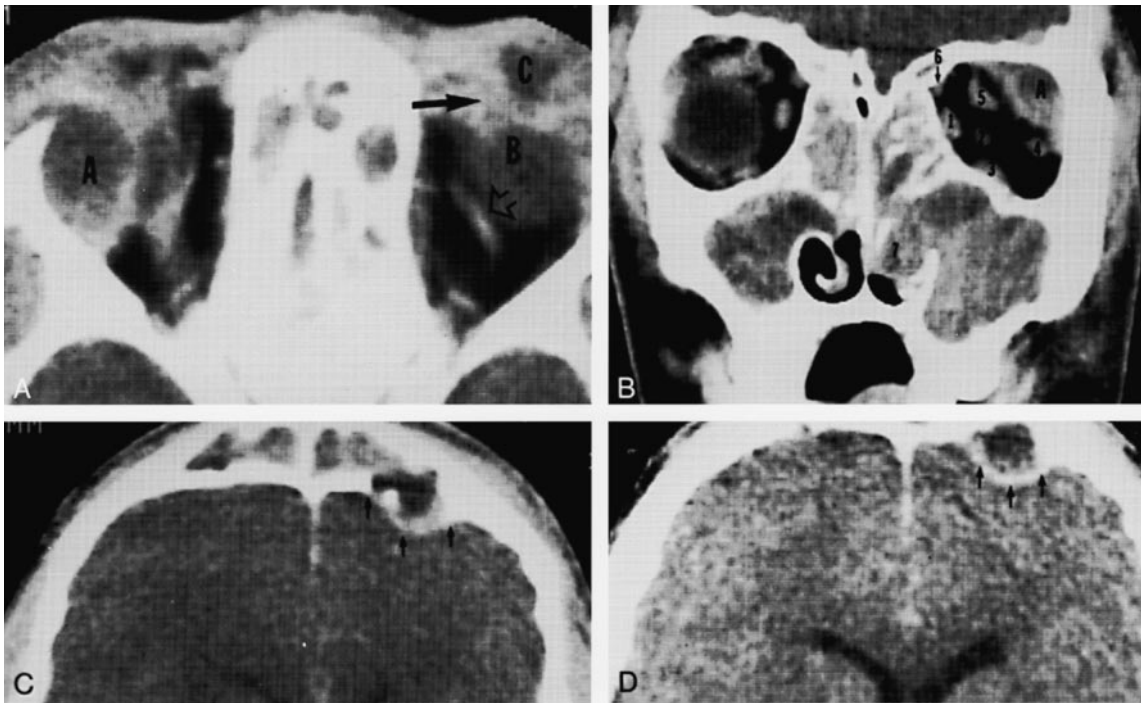


FIGURE 31–30. *A*, A computed tomographic (CT) scan showing increased density of the ethmoid labyrinth and bilateral retro-orbital soft tissue inflammation (cellulitis) (A and B). Note also a low-density image involving the left upper eyelid (C). The eyelid lesion (C) was drained, and 3 mL of pus were removed. *Hollow arrowhead* points to the superior ophthalmic vein, and the *arrow* points to the retroseptal granulation tissue. *B*, Postcontrast semicoronal CT scan of the same patient in *A* showing enhancing soft tissue densities (polyps and hyperplastic mucosa) involving the maxillary and ethmoid sinuses. Notice the polyp in the right nasal cavity (7) and periorbital cellulitis involving the right orbit (A). 1 = medial rectus muscle; 2 = optic nerve; 3 = inferior rectus muscle; 4 = lateral rectus muscle; 5 = superior rectus muscle; 6 = superior oblique muscle. *C*, A CT scan of the same patient in *A* and *B* showing increased density of the frontal sinuses with erosion of the posterior table of the left frontal sinus (*arrows*). *D*, Postcontrast CT scan obtained 5 mm superior to *C*, showing enhancement along the exposed and displaced dura (*arrows*).

Chronic inflammatory disease is often associated with sclerosis of the walls of the sinuses and ethmoidal trabeculae[41] (Figures 31–32 and 31–33). Fungi, particularly *Aspergillus*, may be considered possible causative agents. Fungal sinusitis can be divided into several categories: fungus ball formation, allergic fungal sinusitis, and acute and chronic sinusitis. "Fungal ball" refers to a saprophytic colonization of a cavity by fungus hyphae without invasion of adjacent tissue. It usually occurs in the lung (typically in a parenchymal cavity or an ectatic bronchus), the paranasal sinuses, nasal and auricular cavities, or the urinary tract.[42] *Aspergillus* is most commonly involved in the formation of fungus balls, although they may occasionally be caused by *Candida, Sporothrix*, or *Penicillium* or by bacteria, namely *Nocardia*.[42] Fungus balls are generally asymptomatic unless they are associated with bleeding. Fungus ball is seen as an area of increased density and calcification within the center of an involved sinus. Allergic fungal sinusitis can slowly expand and erode into surrounding structures. Atopy, asthma, and chronic illness requiring multiple courses of antibiotics are characteristic of this infection.[43] Neither fungus ball formation nor allergic fungal sinusitis has been associated with diabetes.[43]

Polyposis of the paranasal sinuses and nasal cavity (Figure 31–34) and mucocele of the maxillary (Figure 31–35), frontoethmoid (Figure 31–36), and sphenoid sinuses (Figure 31–37) are best evaluated by CT and MRI scans.[29,33,44,45] Mucoceles and benign tumors tend to expand the area of origin by virtue of their slow growth.[29,44] The gradual pressure, atrophy, and erosion of the bone by the enlarging soft tissue

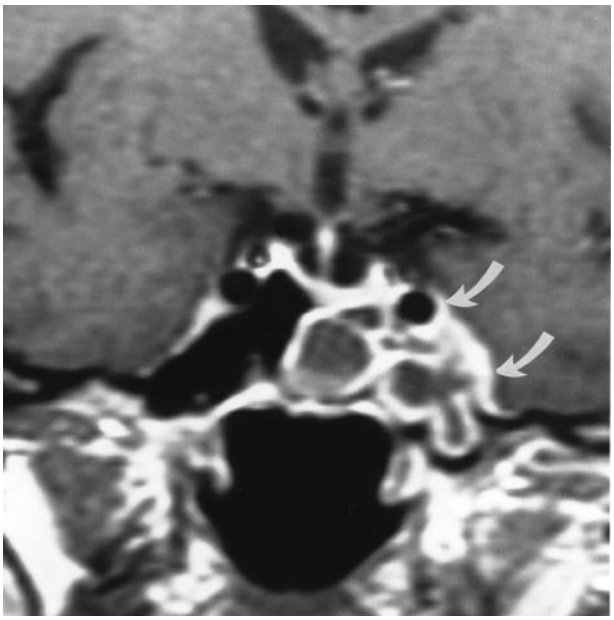


FIGURE 31–31. Aggressive aspergillosis, postcontrast T_1-weighted magnetic resonance scan showing left sphenoid sinusitis, extending into the left cavernous sinus (*arrows*).

mass of the mucocele and the expansile low-density appearance on CT (see Figures 31–36 and 31–37), with no enhancement after contrast infusion (except

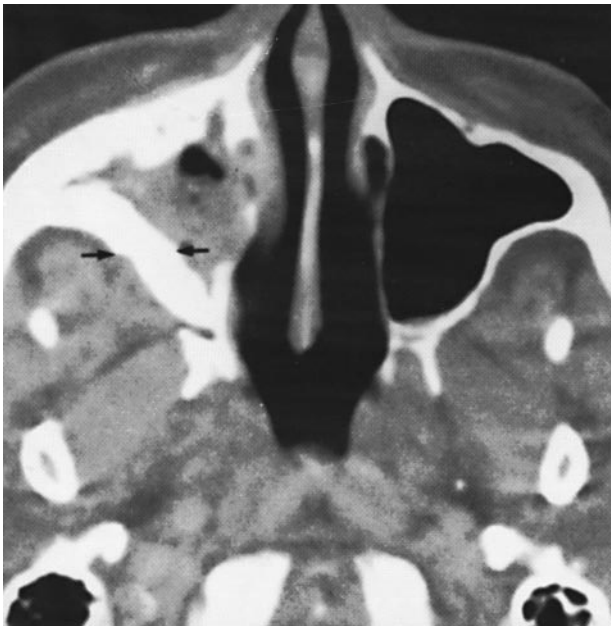


FIGURE 31–32. A computed tomographic scan showing marked hyperplastic mucosa with periosteal thickening (*arrows*) (osteoblastic sinusitis).

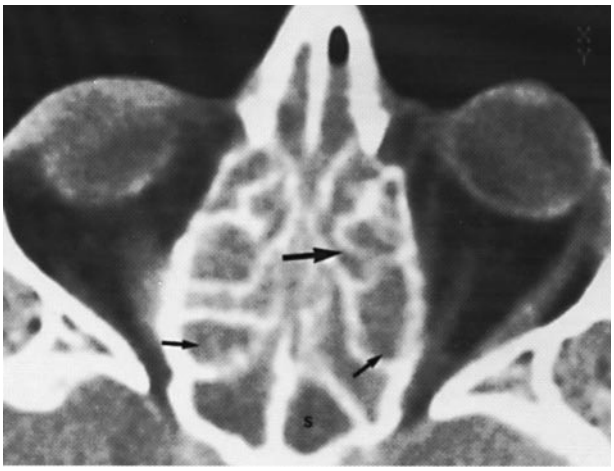


FIGURE 31–33. Chronic hyperplastic polypoid sinusitis. Postcontrast axial computed tomographic scan showing enhanced soft tissues in the ethmoid labyrinth (*arrows*), sphenoid sinuses (s), and nasal cavity in this patient with nasal and paranasal sinus polyposis.

around the inflamed capsule and peripheral calcifications), make the CT diagnosis of mucocele almost certain. Mucoceles may extend into the orbit or intracranially from the frontal, ethmoid, and sphenoid sinuses (see Figure 31–37). Computed tomography is the diagnostic method of choice for the diagnosis and management of the mucocele. Mag-

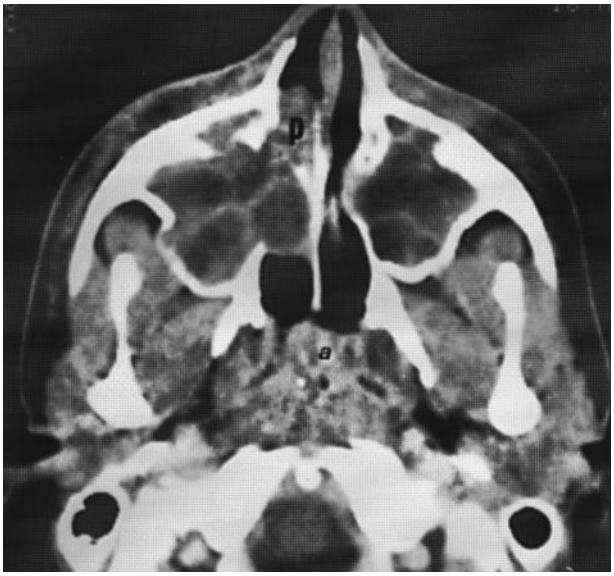


FIGURE 31–34. Nasal and paranasal sinus polyps. Postcontrast axial computed tomographic scan, showing polyp in the right nasal cavity (p) and enhancing soft tissue in the maxillary sinuses. Note the large adenoid (a).

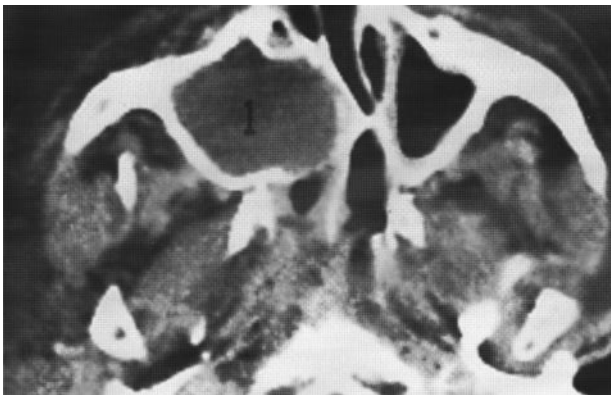


FIGURE 31–35. A computed tomographic scan showing expansion and replacement of the right maxillary sinus (1) by a soft tissue of low density, which is characteristic of a mucocele.

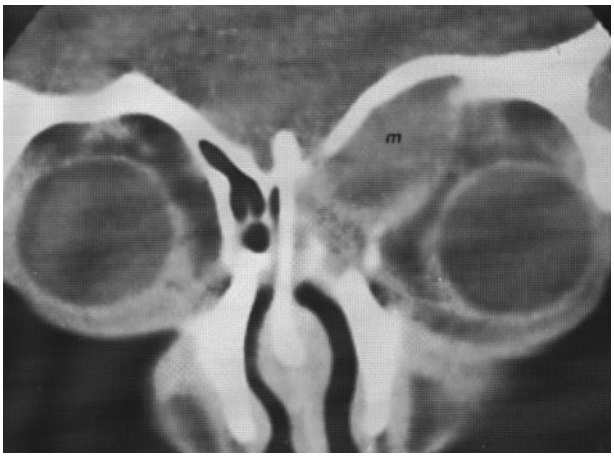
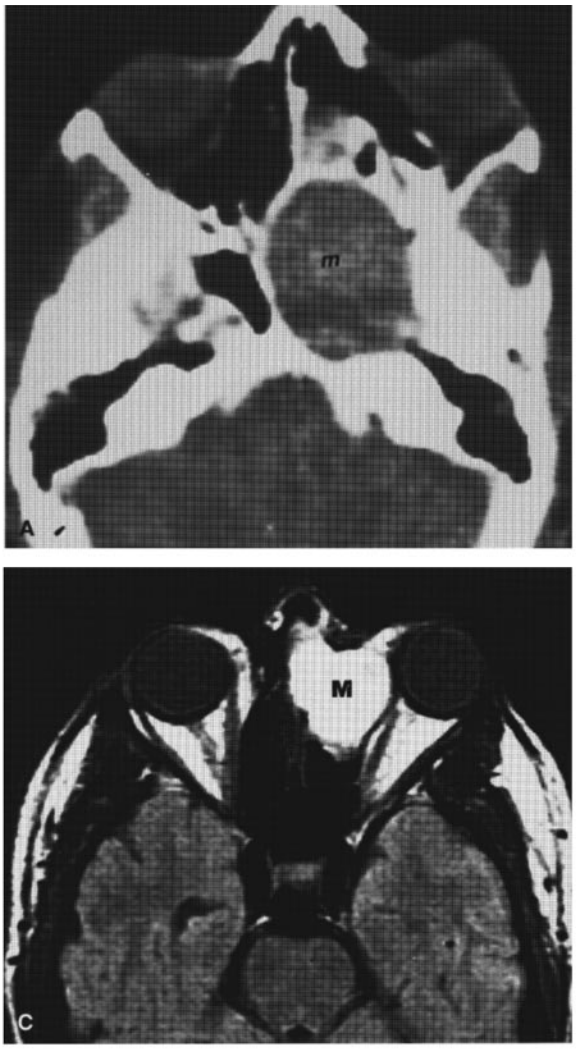


FIGURE 31–36. Mucocele of the ethmoid sinus. Post-contrast coronal computed tomographic scan showing expansion and replacement of the supraorbital recess of the left ethmoid sinus by a soft tissue of relatively low density (m). There is erosion and lateral displacement of the lamina papyracea and the superiomedial rim of the left orbit.

netic resonance imaging can provide valuable information regarding intracranial extension of a mucocele. The signal characteristics of mucoceles are variable according to the protein concentration of the secretions. Mucoceles are most often hyperintense on both T_1W and T_2W MR images (see Figure 31–37). Some mucoceles may appear hypointense on T_2W MR images, owing to high proteinaceous content. Calcifications and inspissated mucus can easily be missed by MRI. At times, a sinus filled with inspissated mucus can appear on MRI as a perfectly aerated and normal sinus. Magnetic resonance imaging should be used as a complementary study to evaluate mucoceles. At present, MRI is the best imaging method for postoperative investigation of the obliterated frontal sinus.[46]

Tumors of the Paranasal Sinuses Computed tomography and MRI have made significant contributions to the imaging evaluation of tumors of the paranasal sinuses and other head and neck tumors.[44–49] Computed tomography and MRI have also been valuable in the management of malignant tumors, particularly in the planning of portals for radiation therapy.[50]

Conventional radiography is adequate for the diagnosis of paranasal osteoma. Computed tomography, however, is indicated when there is orbital extension or other complications (Figure 31–38). Fibro-osseous lesions of the head and neck are best evaluated using CT scan. With MRI, serious mistakes may be made in the evaluation of the fibro-osseous lesions, particularly by a less experienced radiologist. Malignant tumors of the nasal cavity and paranasal sinuses account for only 0.2 to 0.8% of all malignant neoplasms and only 3% of all tumors occurring in the head and neck.[45,47–49] Approximately 50 to 65% of malignant sinonasal tumors arise within the maxillary sinuses, 10 to 25% in the ethmoid sinuses, 0.1 to 4% in the frontal and sphenoid sinuses, and 15 to 30% in the nasal cavity.[45] Malignant tumors of the nasal cavity and paranasal sinuses destroy bone and invade the adjacent soft tissue structures (Figures 31–39 and 31–40). Extension into the orbit, pterygopalatine fossa (Figure 31–41, A), and cranial cavity can best be demonstrated by contrast-enhanced CT or MRI (Figure 31–41). On CT scans, osteogenic or chondrogenic sarcomas of the sinonasal cavities demonstrate irregular islands of tumor bone formation, irregular periosteal reaction, and marked bone destruction[45] (Figure 31–42, A). On the other hand, osteoclastoma or giant cell tumor produces bone destruction as well as bone expansion (Figure 31–42, B). Early subperiosteal extension of a neoplastic condition is best evaluated by CT or MRI scans. Computed tomography remains the study of choice for osseous and chondrogenic lesions such as osteoma, osteoid osteoma, osteoblastoma, ossifying fibroma, chondroblastoma, osteogenic and



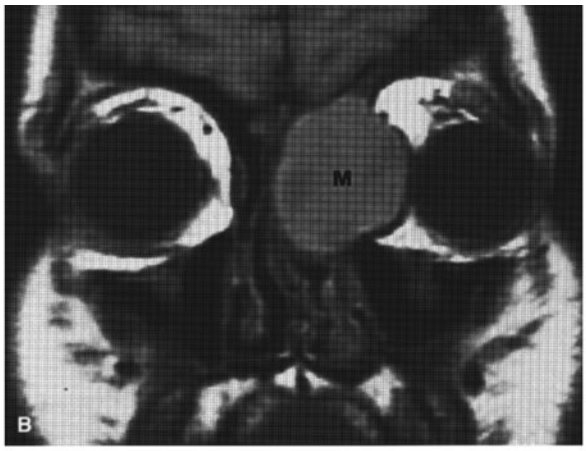


FIGURE 31–37. *A*, A computed tomographic scan showing expansion and replacement of the left sphenoid sinus by a soft tissue of low density, which is characteristic of a mucocele (m). *B*, Mucocele of the left ethmoid sinus. T_1-weighted magnetic resonance coronal (MR) scan shows an expansile rather hyperintense mass (M) involving the left ethmoid sinus. *C*, Proton-weighted axial MR scan, the same patient as in *B*, showing that the mass (M) is markedly hyperintense.

chondrogenic sarcomas, and developmental conditions such as FD.[33,45] Although CT is more specific in the diagnosis of osteogenic and chondrogenic sarcomas, MRI is more sensitive for the determination of the extent of their soft tissue components as well as the presence of subtle or obvious intracranial spread.[33] Magnetic resonance imaging is superior to CT in differentiating inflammatory conditions from neoplastic processes.[45,47] Most inflammatory lesions are hyperintense on T_2W MR images as opposed to most malignant tumors and lymphoreticular proliferative, myeloproliferative, and chronic granulomatous disorders.[45, 47] Most tumors of the sinonasal cavities are not as hyperintense as the surrounding inflammation and retained secretions; therefore, MRI plays an important role in the mapping and staging of these tumors. The intraorbital and intracranial complications of sinus surgery for

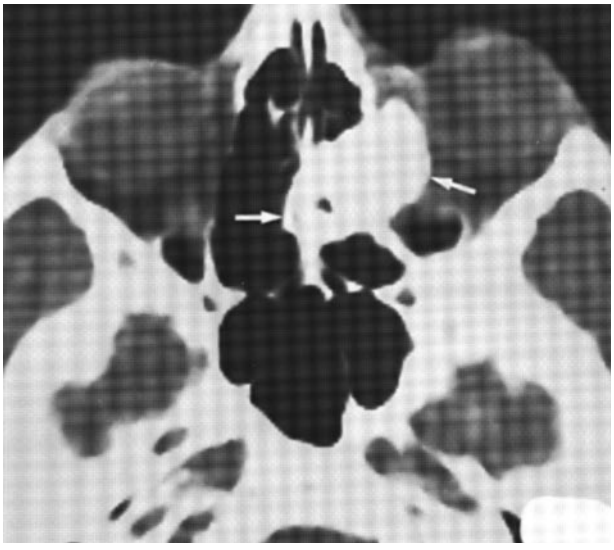


FIGURE 31–38. A computed tomographic scan showing a left ethmoid sinus osteoma extending into the left orbit (*arrows*).

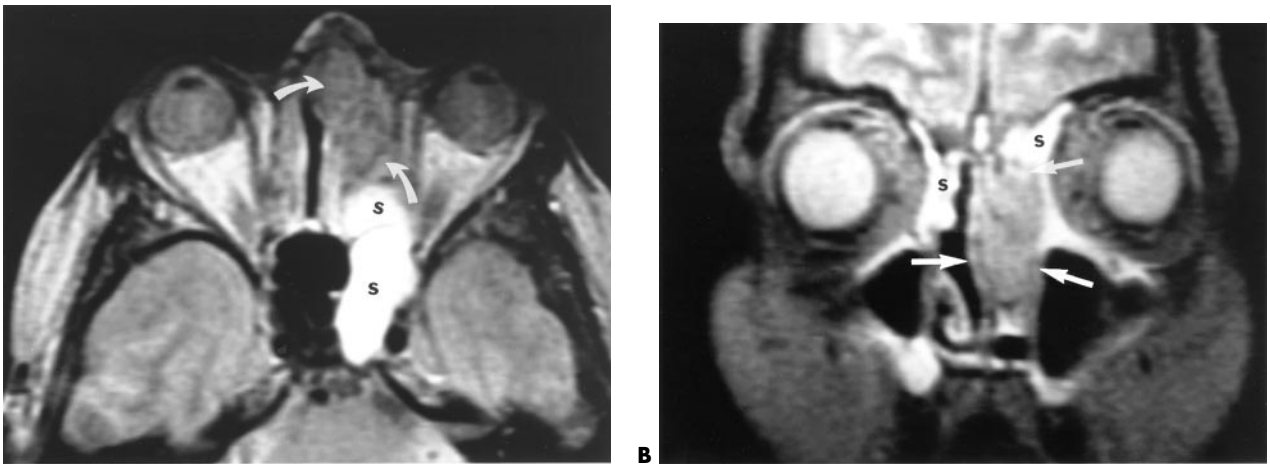


FIGURE 31–39. Squamous cell carcinoma of the sinonasal cavities. *A*, Axial proton-weighted magnetic resonance (MR) scan showing a large mass (*arrows*) involving the left nasal cavity and ethmoid complex. Note mucosal secretion (S) in the left posterior ethmoid sinus and left sphenoid sinus. *B*, Coronal T_1-weighted MR scan, showing a large mass (*arrows*). Note mucosal secretion (S) in the ethmoid air cells.

tumors or inflammatory conditions are best demonstrated by MRI.

Trauma Conventional plain film examination can be used for radiologic diagnosis of the isolated maxillofacial fracture.[51] In complex maxillofacial trauma, CT provides the most cost-effective imaging modality.[51] Current-generation spiral CT scanners provide excellent bone detail. The orbit is superbly suited to CT scanning. Orbital findings in trauma are well documented.[51] The fracture fragments and foreign bodies projecting into the orbit and fat and muscle herniation projecting into the ethmoid or maxillary sinuses are readily identified together with any hematoma and/or pneumo-orbit (Figure 31–43). Computed tomography is the procedure of choice for evaluation of the position of the Silastic implant for the repair of orbital floor fracture. The fracture of the lamina papyracea, which is difficult to image conventionally, can be imaged using axial and coronal sections.

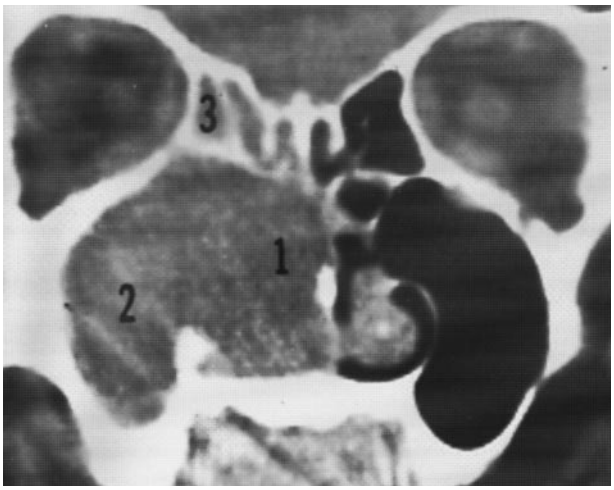


FIGURE 31–40. Carcinoma of the nasal cavity. Postcontrast coronal computed tomographic scan showing a large tumor in the left nasal cavity (1) extending into the left maxillary sinus (2) and left ethmoid sinus (3). Note destruction of the medial wall of the left maxillary sinus and partial destruction of the maxilloethmoidal plate.

The intracranial and intraorbital extension of complex maxillofacial fractures, shattering craniofacial fractures, and medial and lateral orbital wall and orbital apex fractures can be imaged using CT scanning in a manner not previously possible.[51] The axial dimension of posterior maxillary, alveolar, and pterygoid plate fractures and the posterior extensions of the Le Fort group of fractures are particularly suited to axial CT imaging (Figure 31–44).[30] Displacement and rotation of the malar bone and bone fragments in tripod fractures can best be demonstrated with axial CT (Figure 31–45). Sphenoid sinus fractures can be evaluated using axial and particularly direct coronal section and sagittal reformatted images. The coronal CT sections should be taken only in the injured patient who is physically stable and is without cervical spine injury. In a sphenoid sinus fracture, opacification of the sinus

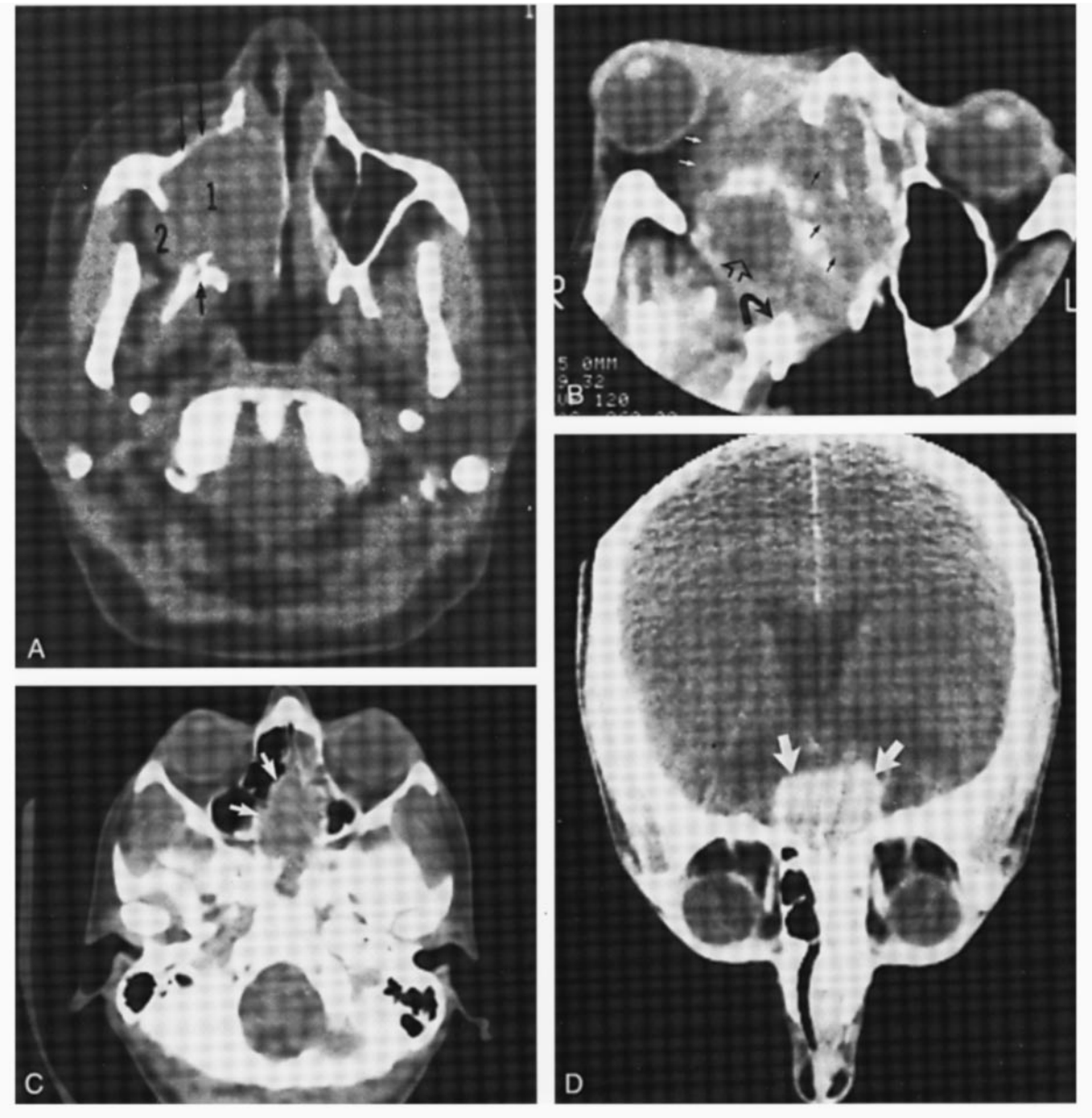


FIGURE 31–41. *A*, Carcinoma of the right maxillary sinus. Postcontrast axial computed tomographic (CT) scan showing a large mass involving the right maxillary sinus (1) and extending into the right nasal cavity. There is erosion of the anterior wall of the maxillary antrum (*arrows*). Note destruction of the posterior wall of the right maxillary sinus and extension of the tumor into the right infratemporal fossa (2). Note the tumor in the right pterygopalatine fossa with partial erosion in the pterygoid portion of the right sphenoid bone (*arrow*). *B*, Rhabdomyosarcoma; axial CT scan shows a large mass in the right ethmoid sinus (*black arrows*) with extension into the right orbit (*white arrows*). Note extension of the tumor in the superior portion of the right maxillary sinus (*hollow arrow*) with tumor seen in the pterygopalatine fossa (*curved arrow*). *C*, Esthesioneuroblastoma; axial CT scan shows a mass in the left ethmoid sinus (*arrows*) with opacity of the sphenoid sinuses. *D*, Coronal CT scan of the same patient as in *C* showing markedly enhanced mass (*arrows*) in the anterior cranial fossa caused by extension of the tumor along the cribriform plate into the cranium.

Continued on next page

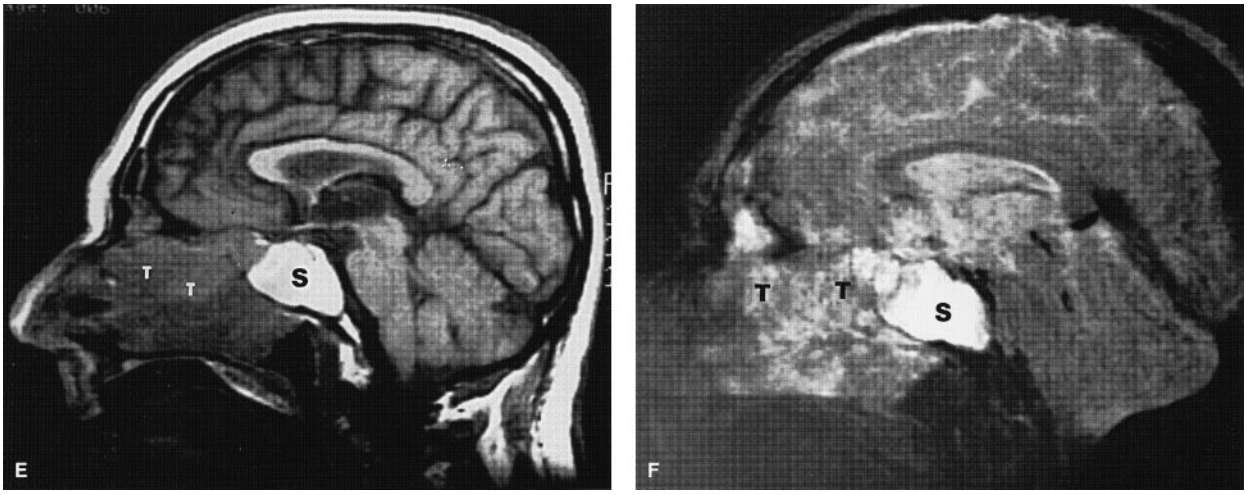


FIGURE 31–41 continued. *E*, Adenocarcinoma of the nasal cavity and ethmoid sinus. T_1-weighted sagittal magnetic resonance (MR) scan shows a large tumor (T) within the nasal cavity and ethmoid cells. Note the hyperintensity of the retained secretion in the sphenoid sinus (S). *F*, T_2-weighted sagittal MR scan of the same patient as in *E*. The tumor (T) appears hyperintense. The retained fluid remains hyperintense within the sphenoid sinus (S).

with blood or CSF can be readily recognized with a CT scan.

Three-dimensional reconstruction of CT images is a well-established imaging modality that has been investigated in various clinical settings. Perhaps one of its most useful applications is in trauma. Rotational abnormalities are better appreciated on three-dimensional images (see Figure 31–45).

Cerebrospinal Fluid Leak and Pneumocephalus

In trauma patients, the recognition of free air within the cranial cavity and cisternal and ventricular system is the result of a dural and subarachnoid tear. This can occur at any number of sites: the posterior wall of the frontal sinus, the roof of the ethmoid labyrinth, the cribriform plate, the walls of the sphenoid sinus, the middle and posterior cranial fossae, and beyond. Localization of the fracture site and

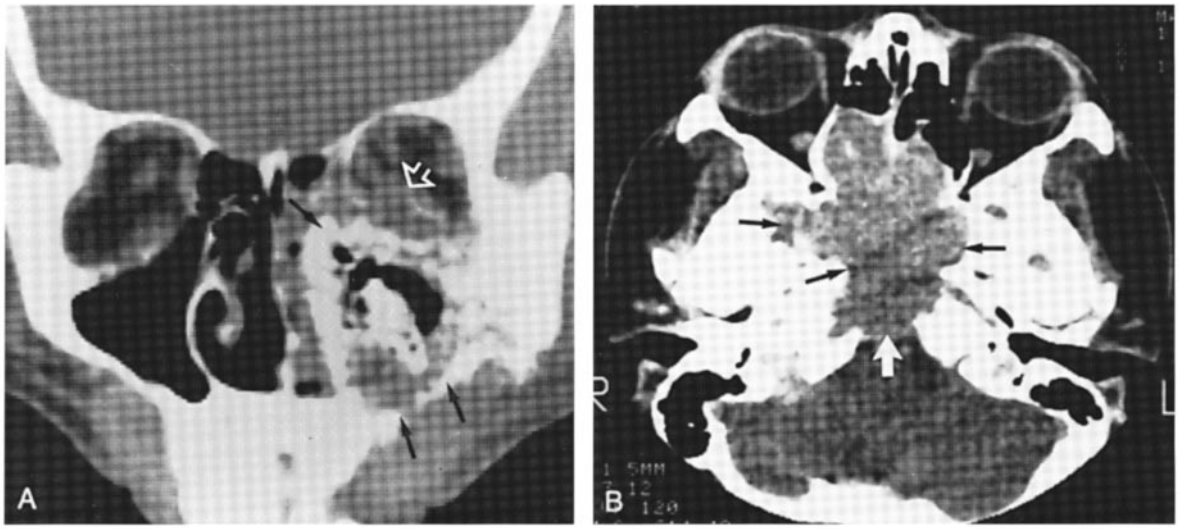
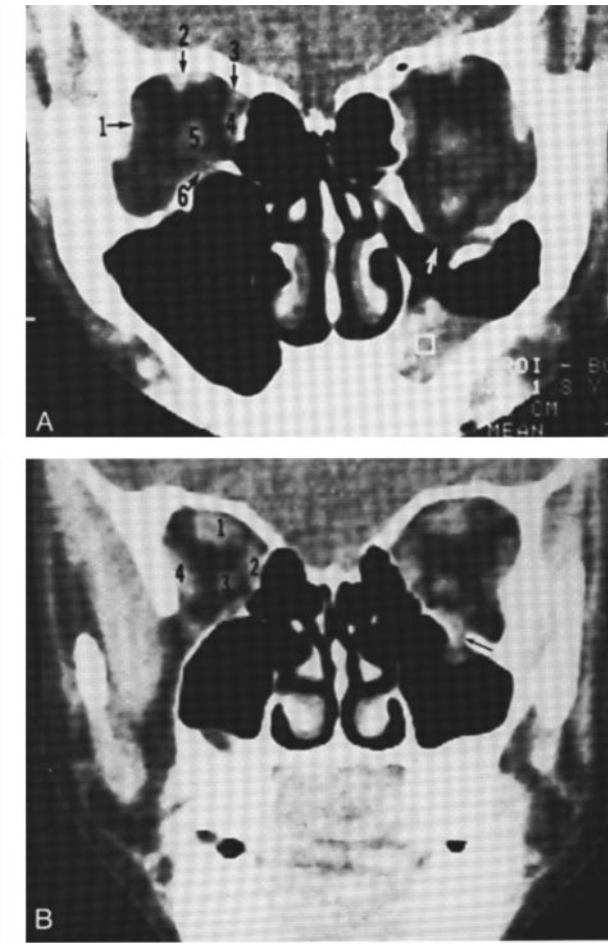


FIGURE 31–42. *A*, Osteogenic sarcoma computed tomographic (CT) scan showing a large maxillary tumor with irregular islands of bone formation (*arrows*) and with marked bone destruction. Notice invasion of the floor of the orbit (*hollow arrow*). *B*, Giant cell tumor of the sphenoid bone; axial CT scan shows a destructive tumor of the sphenoid bone (*black arrows*) with extension into the posterior ethmoid cells. Notice destruction of the clivus (*white arrow*) and exposed dura.



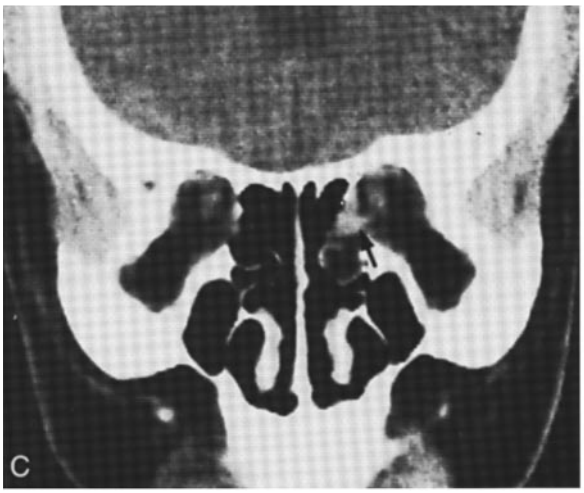


FIGURE 31–43. *A*, Coronal computed tomographic (CT) scan showing a "blowout fracture" of the right orbital floor with displacement of the periorbital fat and fractured fragments into the right maxillary sinus (*arrow*). Notice fluid in the right maxillary sinus with a cursor measurement of 36, which is indicative of hemorrhage. 1 = lateral rectus muscle; 2 = superior rectus muscle; 3 = superior oblique muscle; 4 = medial rectus muscle; 5= optic nerve; 6 = inferior rectus muscle. *B*, Coronal CT scan showing a fracture of the floor of the right orbit (*arrow*) with the inferior rectus muscle partially hooked over the fracture site. 1 = superior rectus muscle; 2 = medial rectus muscle; 3 = optic nerve; 4 = lateral rectus muscle. *C*, Coronal CT scan showing a fracture of the right medial orbital wall with entrapment of the medial rectus muscle (*arrow*).

bone fragments in axial and, if possible, direct coronal display, using thin sections (1.5 to 3 mm) in critical areas is mandatory. Contrast examination with intrathecal, water-soluble, iodinated contrast material is not indicated at this stage since most acute CSF leaks will close spontaneously or with operative fracture reduction alone.[30] In persistent CSF leaks, all such fistulae should undergo CT cisternography (Figure 31–46).[52,53] The examination must be carried out during a period of active CSF rhinorrhea or otorrhea for effective imaging results.[52,53] Other examinations, such as a radionuclide CSF scan, may not be needed.[30] When the leak is large, MRI may be sufficient to make the correct diagnosis. Aggressive arachnoid granulations, causing CSF leak, may be detected by MR following IV administration of gadolinium contrast material. There will be local enhancement at the site of arachnoid granulation.

Nasopharynx The nasopharynx is a cuboidal structure whose anterior limits are the choanae. The roof is attached to the base of the skull at the clivus and slopes downward to become the posterior pharyngeal wall, which overlies the atlas and its related ligaments and muscles. The inner surface of the pharyngeal musculature (pharyngeal constrictor muscle) is lined by the pharyngobasilar fascia. The pharyngobasilar fascia, or pharyngeal aponeurosis, encircles the space between the upper border of the superior constrictor muscle and the skull base. The eustachian tube opens into the nasopharynx by piercing the pharyngobasilar fascia. The most prominent anatomic landmark on the lateral wall of nasopharynx on each side is the cartilaginous medial end of the eustachian tube (torus tubarius) (Figures 31–47 and 31–48). Between the torus and the posterior wall is the fossa of Rosenmüller, a cleft-like space whose apex reaches the anterior margin of the carotid canal (see Figures 31–47 and 31–48).[54,55]

Computed tomography and MR examinations of the nasopharynx delineate the air-containing cavity of the nasopharynx, the mucosal surfaces, and the deep

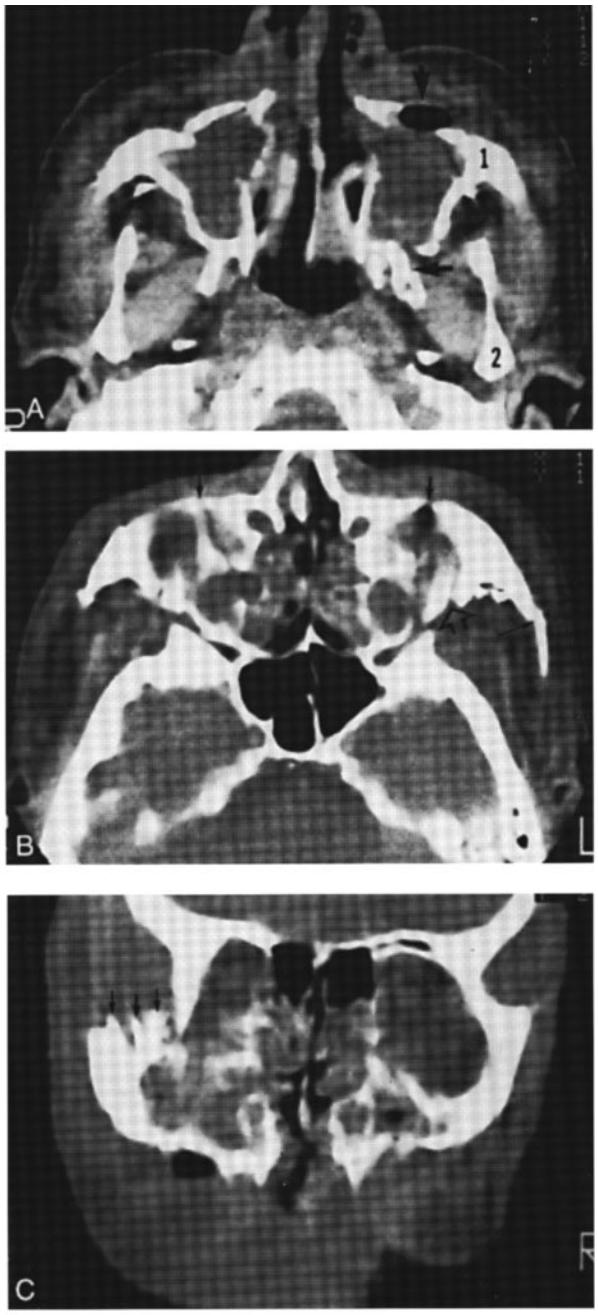


FIGURE 31–44. A complex maxillofacial fracture. Axial computed tomographic (CT) scan showing fractures of the anterior and medial walls of both maxillary sinuses. Note marked soft tissue swelling of the face, air in the subcutaneous tissue (*arrow*), and opacification of the maxillary sinuses caused by hematoma. Fracture of the left zygoma (1), fracture of the left pterygoid plates (*posterior arrow*), and normal mandible (2). *B*, Axial CT scan obtained 1 cm superior to *A*. Note bilateral fractures of the inferior orbital rims (*anterior arrows*) extending along the zygomaticomaxillary suture on the left and involving the posterior wall of the left maxillary sinus (*open arrow*). Note fractures of the left zygomatic arch (*arrow*), the floor of the right orbit, and the posterior wall of the right maxillary sinus. *C*, Coronal CT scan of the same patient in *A* and *B*. Note separation of the left frontozygomatic suture (*arrows*) and lateral displacement of the left fractured malar bone. Note fractures of the anterior wall of both maxillary antra and shattered fractures of the ethmoid sinuses and floors of both orbits.

fascial planes (Figures 31–47 and 31–48). The anatomic landmarks of the nasopharynx and the muscles affecting eustachian (pharyngotympanic) tube function, the tensor and levator palatini and the stylopharyngeus, are better depicted on MR images (see Figure 31–47, A). The cartilaginous end of the eustachian tube bulges into the airway, producing the mucosal prominence known as the torus tubarius (see Figure 31–48, A).

The pharyngeal recesses or fossae of Rosenmüller extend posteriorly and superiorly to the tori (see Figure 31–48). The pharyngeal openings of the eustachian tubes are located on the anterior and inferior surfaces of the tori (see Figure 31–48). Scans taken using a modified Valsalva's maneuver show distention of both the eustachian tubes and lateral recesses. On axial and coronal CT and MR images obtained during respiration, air is rarely seen more than 3 to 5 mm within the pharyngeal end of the eustachian tube. The tympanic end of the eustachian tube can easily be seen in thin (1 to 1.5 mm) axial CT scans. The intrapharyngeal muscles (levator and tensor veli palati and salpingopharyngeus) lie between the pharyngobasilar fascia and the mucosa. The pharyngobasilar fascia is not seen as a separate discrete structure on CT images. The levator veli palatini is the major contributor to intrapharyngeal muscle density seen on CT.[54,55] The muscle belly produces a characteristic oval image just deep to the mucosa posterior to the torus (see Figure 31–47, B). The tensor veli palatini opens the eustachian tube and can be imaged on MR as a distinct muscle bundle (see Figure 31–47, B).

The prevertebral musculature (mainly the longus capitis muscles) is identifiable readily on CT and MRI scans (see Figures 31–47 and 31–48). The parapharyngeal space (PPS), filled with loose fibrofatty tissues, lies between the pharyngeal muscles and the lateral and medial pterygoid muscles (see Figure 31–48, B). This space is subdivided by a consistent layer of fascia into prestyloid and poststyloid

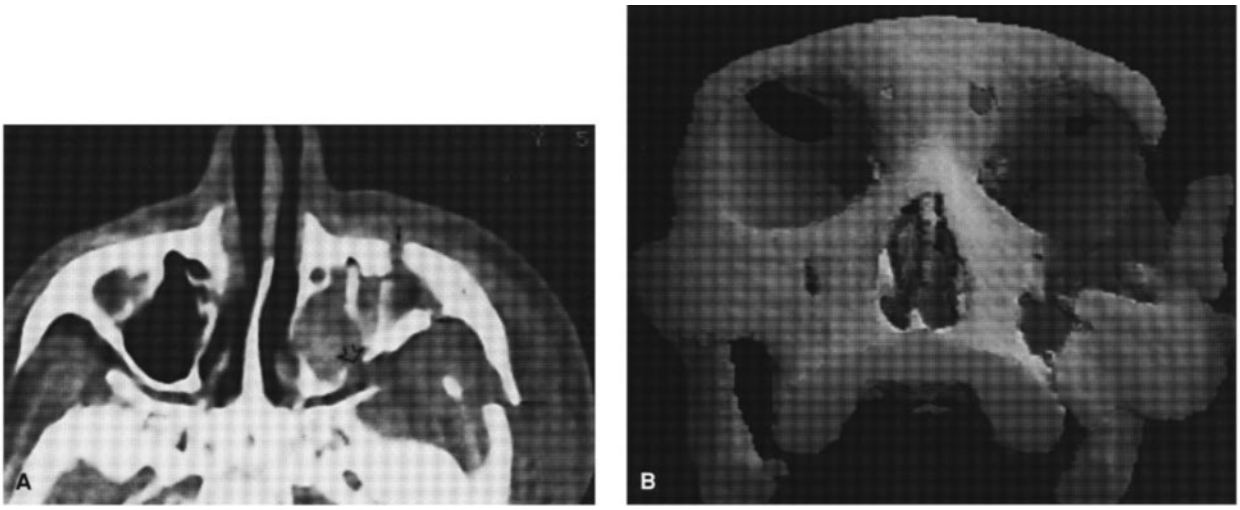


FIGURE 31–45. *A*, Malar fracture. Axial computed tomographic (CT) scan showing fractures involving the inferior orbital rim and the left malar bone, which extend along the left zygomaticomaxillary suture (*small arrow*). Notice fracture of the left zygomatic arch, fracture of the posterior wall of the left maxillary sinus (*open arrowhead*), and bone fragments within the left maxillary antrum with associated left antral submucosal hematoma. *B*, Malar fracture. Three-dimensional image reformation from regular CT scans showing the fractures and displacement of the malar bone.

compartments. The relationship between the medial pterygoid fascia, representing the medial wall of the masticator space, and the fascia of the tensor veli palatini, representing the medial wall of the prestyloid PPS, allows better preoperative differentiation and perhaps more accurate diagnosis. Specifically, tumors that arise in the prestyloid PPS are most often of salivary gland origin.[56] We have seen neurogenic tumors and hemangiopericytoma in the prestyloid PPS. Tumors that arise in the masticator space are unlikely to be of salivary gland origin. Tumors that arise in the poststyloid PPS are most often of neurogenic or chemoreceptor (paraganglion) origin. The fat content of the PPS allows one to identify it as hypodense (dark) on CT scan (see Figure 31–47, A) and hyperintense (bright) on T_1W and PW MR images (see Figure 31–47, B).

FIGURE 31–46. Cerebrospinal fluid rhinorrhea. Coronal computed tomographic scan after intrathecal injection of metrizamide showing contrast-enhanced subarachnoid space of the anterior cranial fossa (*upper arrow*) just above the crista galli. Note the fracture of the left cribriform plate (*lower arrow*) and the presence of metrizamide (high density) within the ethmoid labyrinth and the air metrizamide fluid level in the left maxillary sinus.

Squamous cell carcinomas account for 80% of nasopharyngeal cancers. The remaining nasopharyngeal tumors will prove to be lymphomas (including plasmacytoma and Burkitt's lymphoma), chloroma, chordoma, and a variety of rare conditions, including adenocarcinomas, melanoma, rhabdomyosarcoma, schwannoma, lipoma, liposarcoma, Kaposi's sarcoma, and metastasis. Magnetic resonance imaging and CT scans of the nasopharynx are indispensable in patients with clinically diagnosed nasopharyngeal disease in whom information concerning the extent of disease is necessary (Figures 31–49 through 31–53) and in those patients in whom a search is being made for an unknown primary tumor of the head and neck (see Figure 31–51). Hypertrophied and normal lymphoid tissue is seen commonly within the nasopharynx.[54] Ade-

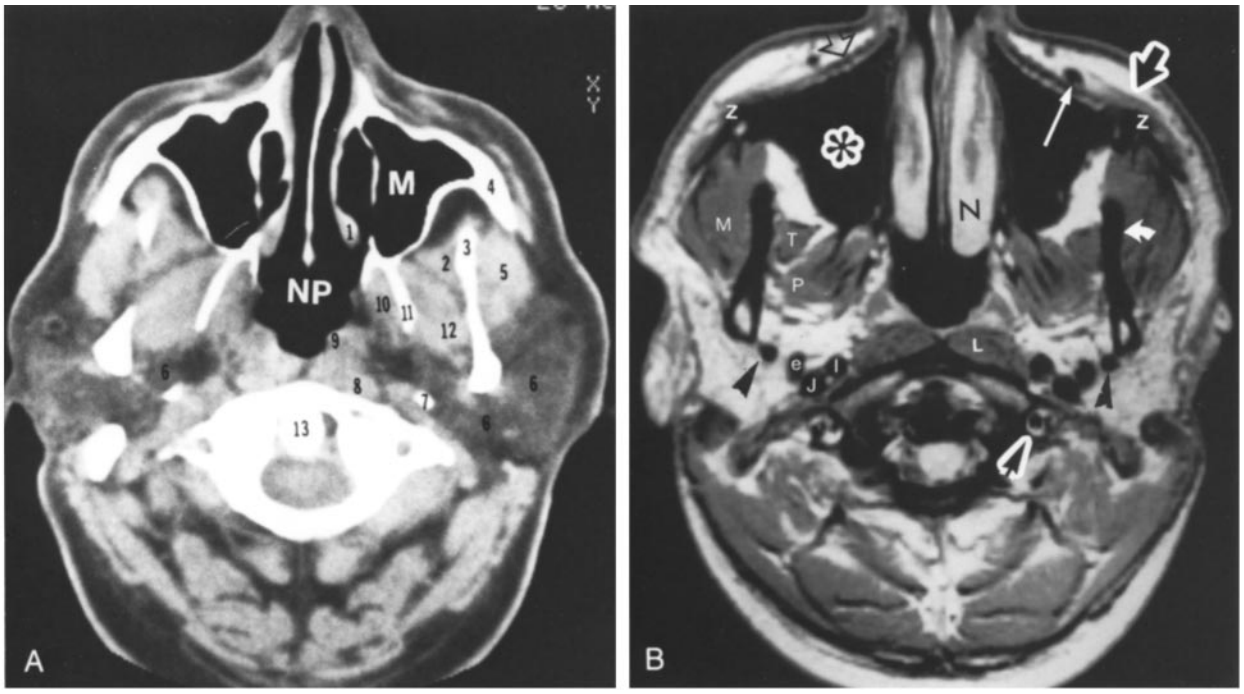


FIGURE 31–47. Normal nasopharynx, maxillary sinus, and infratemporal fossa. *A*, Axial computed tomographic scan through the midportion of the maxillary sinuses (M) and upper part of the odontoid process (13) showing the nasal septum, inferior turbinate (1), nasopharynx (NP), temporalis muscle (2), coronoid process (3), zygoma (4), masseter muscle (5), parotid gland (6), styloid process (7), longus capitis muscle (8), torus tubarius (9), tensor veli palatini muscle (10), lateral pterygoid plate (11), and lateral pterygoid muscle (12). The density seen behind the torus is mainly caused by the levator veli palatini muscle. The tensor veli palatini is rarely seen as a distinct muscle bundle. Part of the density seen medial to the lateral pterygoid plate is caused by the tensor veli palatini muscle (10). Medial to the deep lobes of the parotid glands (6), the symmetric low-density parapharyngeal spaces are seen. Posterior to the styloid process are seen the densities of the petrostyloid neurovascular structures. *B*, Axial proton-weighted magnetic resonance scan showing the quadratus labii superioris muscle (*open arrow*), obicularis oculi and zygomaticus muscles (*white and black arrows*), angular vein (*white arrow*), zygomatic bone (Z), maxillary antrum (*), nasal turbinate (N), masseter muscle (M), pterygoid muscles (P), mandible (*curved arrow*), retromandibular vein (*arrowhead*), vertebral artery (*white and black arrows*), longus colli muscle (L), external (e) and internal (I) carotid arteries, internal jugular vein (J), and temporalis muscle (T).

noidal tissue is usually symmetric and has a lobulated pattern. Benign lymphoid tumor of the nasopharynx is limited to the mucosa and is smooth in outline. Any extension into the deeper plane with obliteration of the facial planes should be considered evidence of a more aggressive lesion[54,55] (Figures 31–49, 31–50, 31–54, and 31–55). Extension of the nasopharyngeal carcinoma into the nasal cavity (see Figure 31–49), infratemporal fossa, orbit, and cranial cavity is best demonstrated by contrast-enhanced CT examination or MRI.[55] Nasopharyngeal carcinomas and, in particular, lymphoepitheliomas are infiltrative. Lymphomas, which are the second most common malignant nasopharyngeal tumors, are bulky and often respect the fascial planes. In children, rhabdomyosarcoma of the nasopharynx as well as lymphoepithelioma (see Figure 31–52) and lymphoma should always be included in differential diagnosis of a nasopharyngeal mass. Lymphoepitheliomas or the so-called Schmincke tumors refer to undifferentiated nonkeratinizing tumors that have numerous lymphocytes. These tumors can be seen in young adults and children (see Figure 31–52). We have seen them in children as young as 5 to 12 years of age. A dynamic CT or a dynamic MR study should be used to differentiate lesions such as juvenile angiofibroma (see Figure 31–54) and for any nasopharyngeal hypervascular lesion such as extension of the glomus tumor along the eustachian tube into the nasopharynx (see Figure

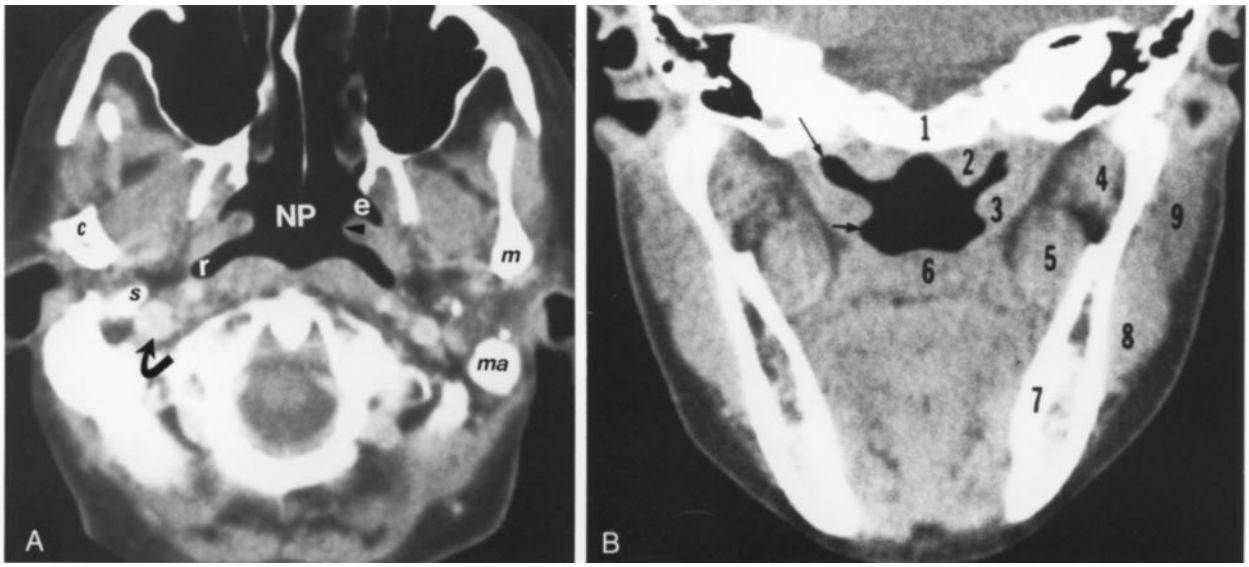


FIGURE 31–48. *A*, Normal nasopharynx. Axial computed tomographic (CT) scan through the midnasopharynx (NP) with the patient instructed to blow against a closed mouth. Note the torus tubarius (*arrowhead*) and extension of air into the eustachian tube (e) (anterior to tori) and Rosenmüller's fossae (r) (posterior to tori). Note enhancement of the retrostyloid vascular structures (internal jugular vein) (*curved arrow*) after intravenous iodine-contrast infusion. The upper part of the parotid gland is seen posterior to the mandible (m) and anterior to the mastoid process (ma). s = styloid process; c = mandibular condyle. *B*, Semicoronal CT section of the nasopharynx showing the clivus (1), longus capitis muscle (2), torus tubarius (3), lateral (4) and medial (5) pterygoid muscles, soft palate (6), mandible (7), masseter muscle (8), and parotid gland (9). Note the eustachian tube opening (*lower arrow*) just inferior to the tori (3) and superior extension of Rosenmüller fossa (*upper arrow*).

31–55) and hypervascular metastases such as metastatic hypernephroma and thyroid carcinoma.

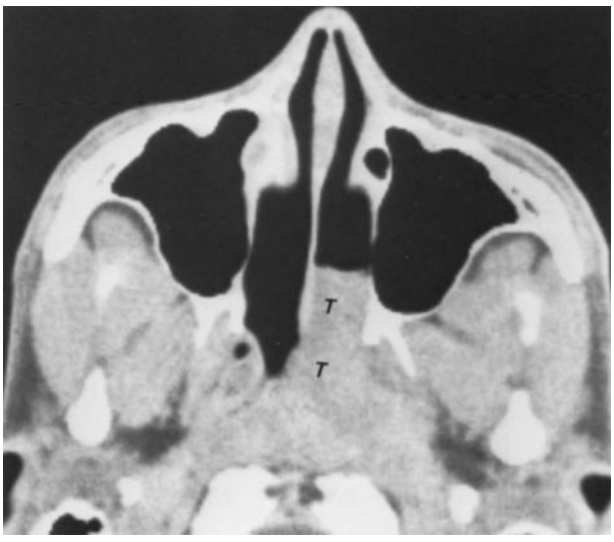


FIGURE 31–49. Nasopharyngeal carcinoma. Axial computed tomographic scan of the nasopharynx showing a large soft tissue tumor (T) occupying the left side of the nasopharyx and extending into the left nasal cavity.

Benign lesions of the nasopharynx such as inflammatory processes (Figure 31–56), Tornwaldt's cyst, parapharyngeal cyst, polyps, papillomas, juvenile angiofibromas, neuromas, fibromas, and lipomas are all well suited for MRI and CT.[54,55]

Orbits

Because fat, fluid, and bone are natural contrast material, the orbits are the ideal structures for highly detailed CT and MRI scans. Scans taken through the orbits show the relationship of the ethmoid sinuses, sphenoid sinuses, pituitary fossa, cavernous sinuses, and temporal lobes with the orbits (see Figure 31–17). The cortical bones, which appear hyperdense on CT images (see Figure 31–17), are seen on MR as hypointense images (see Figure 31–17). The bone marrow appears on T_1W scans as hyperintense owing to the presence of fat. The air is seen as a dark image on both CT and MRI scans. The optic nerves within the orbits can be visualized with both CT and MRI. The intracanalicular portion of the optic nerve is best seen on MRI (see Figure 31–17). The lateral orbital wall separates the orbit from the middle cra-

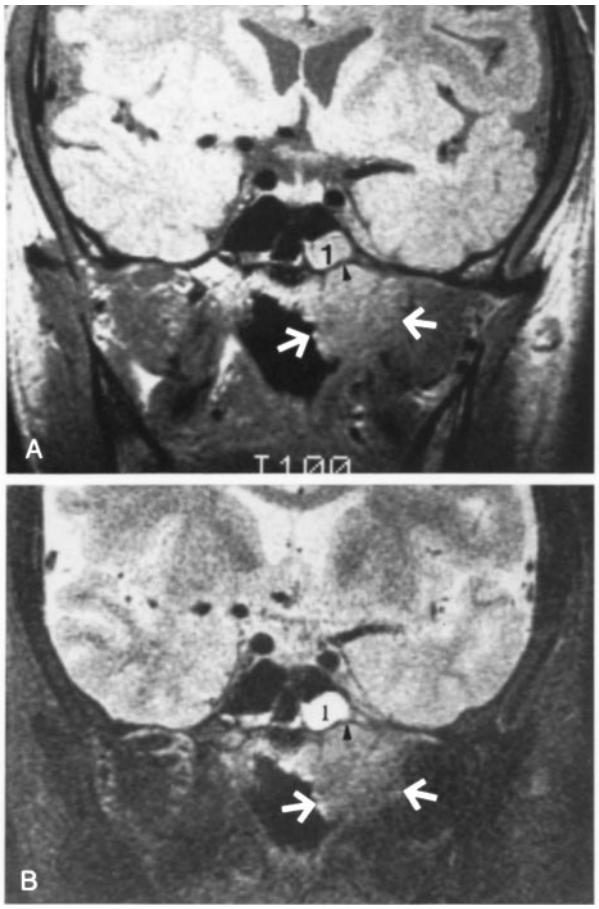


FIGURE 31–50. Squamous cell carcinoma of the nasopharynx. Proton-weighted (*A*) and T_2-weighted (*B*) magnetic resonance (MR) scans show the tumor (*arrows*) and a retention cyst (1) in the left sphenoid sinus. Note the difference between the signal of the tumor and the cyst in the T_2-weighted MR scan. The bone of the base of the skull adjacent to the tumor is not involved (*arrowhead*).

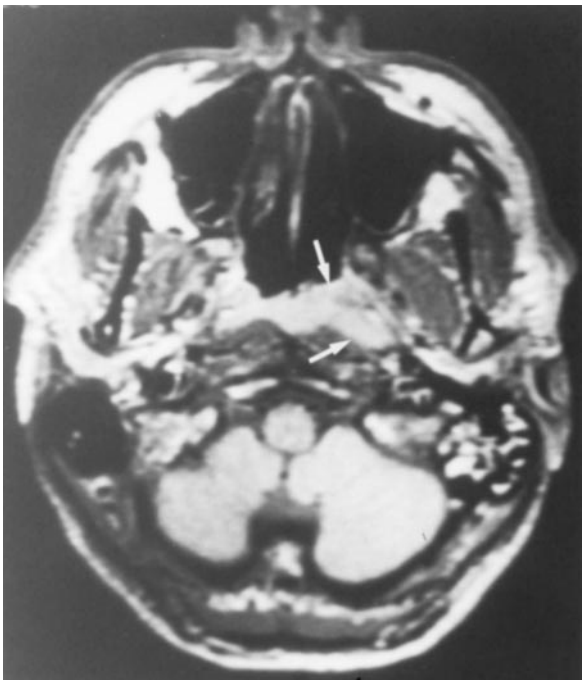


FIGURE 31–51. Nasopharyngeal carcinoma. Proton-weighted magnetic resonance scan showing an infiltrative process involving the left Rosenmüller's fossa (*arrows*).

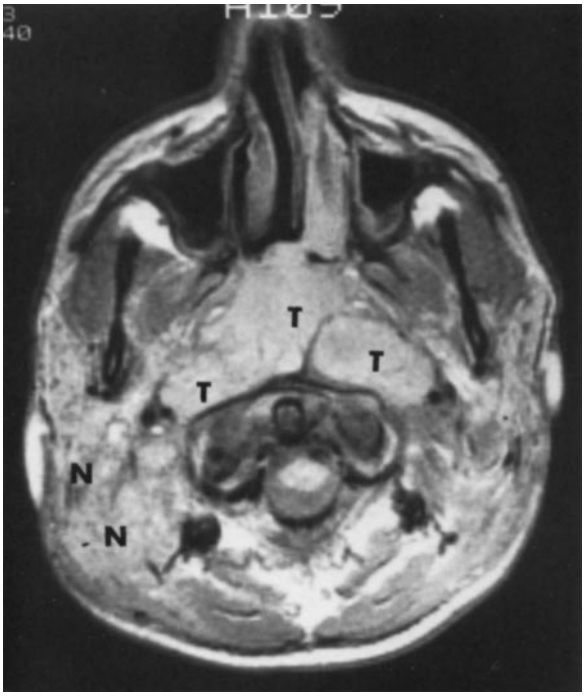


FIGURE 31–52. Poorly differentiated carcinoma with lymphoid stroma in a 16-year-old boy with a 4-month history of nasal stuffiness, fever, and neck masses. Proton-weighted magnetic resonance scan shows massive nasopharyngeal tumor (T) with large cervical lymph nodes (N).

nial and temporal fossae. Computed tomography and MRI are most valuable to evaluate orbital extension of sinonasal pathology (see Figures 31–38, 31–40, and 31–41). Orbital extension of nasopharyngeal tumors, pterygopalatine lesions, and infratemporal fossa pathology can be best evaluated by MRI, particularly on postgadolinium fat-suppressed T_1W MR scans.

Infratemporal Fossa

The boundaries and content of the infratemporal fossa or zygomatic fossa were once difficult to image using radiographic techniques before the advent of

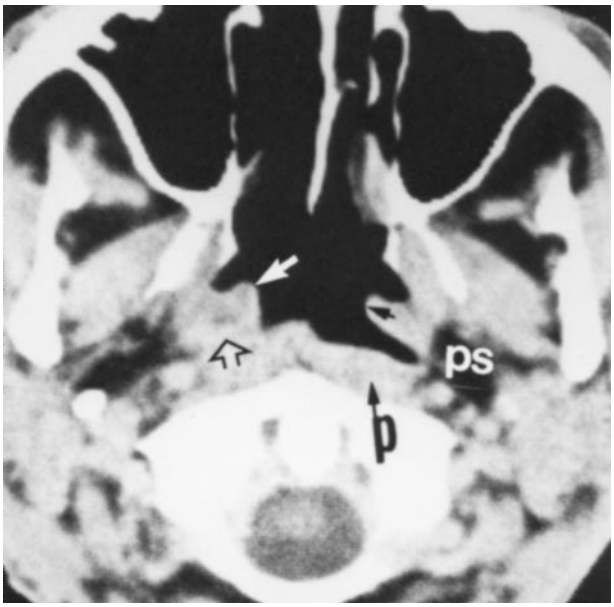


FIGURE 31–53. Clinically unrecognized nasopharyngeal tumor. Notice the normal torus tubarius on the left (*arrowhead*) and anterior to that the normal eustachian tube opening and posterior to that the normal Rosenmüller's fossa and parapharyngeal space (ps). Notice enlargement of the right torus tubarius (*white arrow*) and early infiltration in the adjacent portion of the parapharyngeal space by a soft tissue mass (*hollow arrow*). Biopsy of this lesion revealed carcinoma. p = left paravertebral muscle, longus capitis.

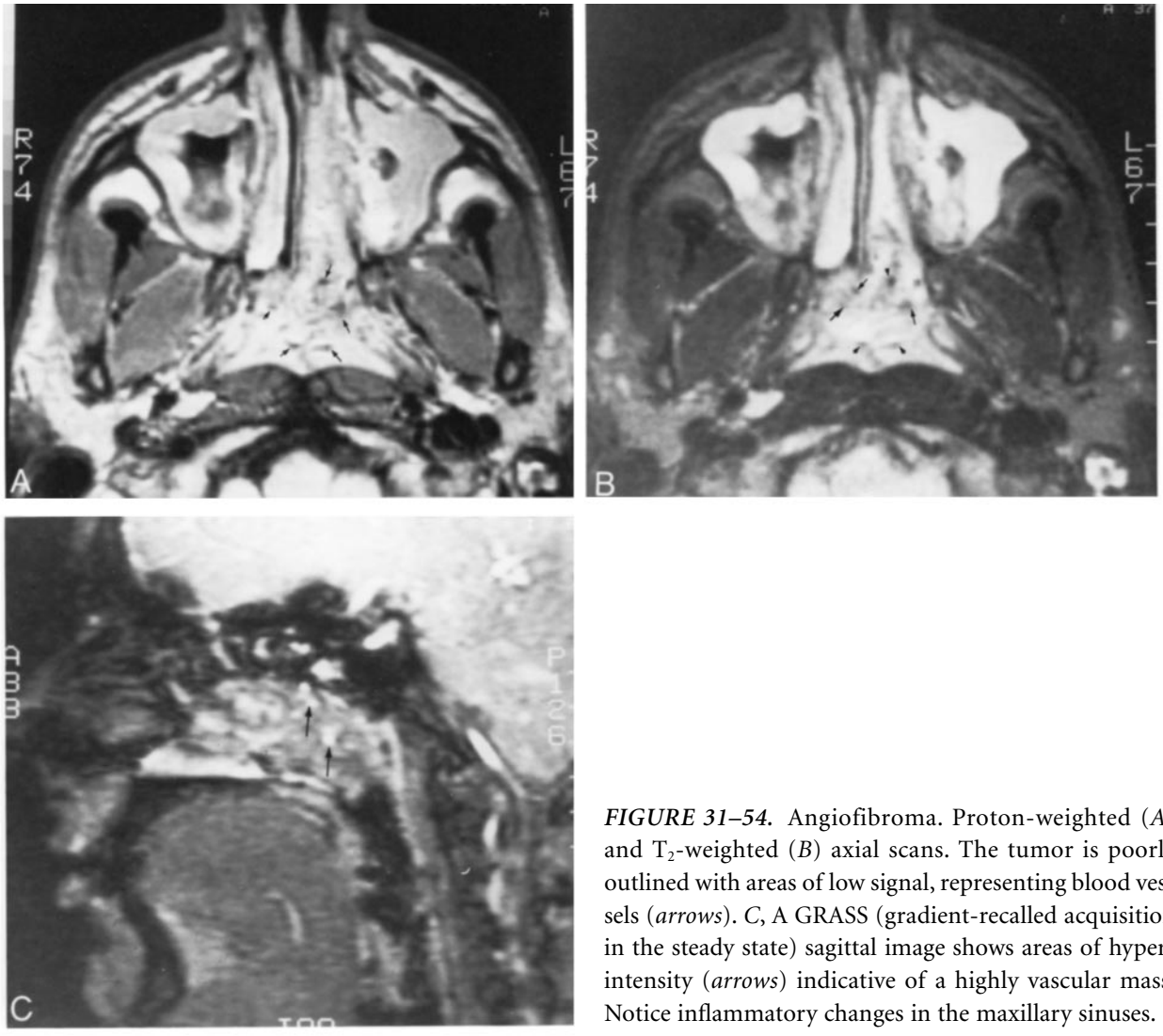


FIGURE 31–54. Angiofibroma. Proton-weighted (*A*) and T_2-weighted (*B*) axial scans. The tumor is poorly outlined with areas of low signal, representing blood vessels (*arrows*). *C*, A GRASS (gradient-recalled acquisition in the steady state) sagittal image shows areas of hyperintensity (*arrows*) indicative of a highly vascular mass. Notice inflammatory changes in the maxillary sinuses.

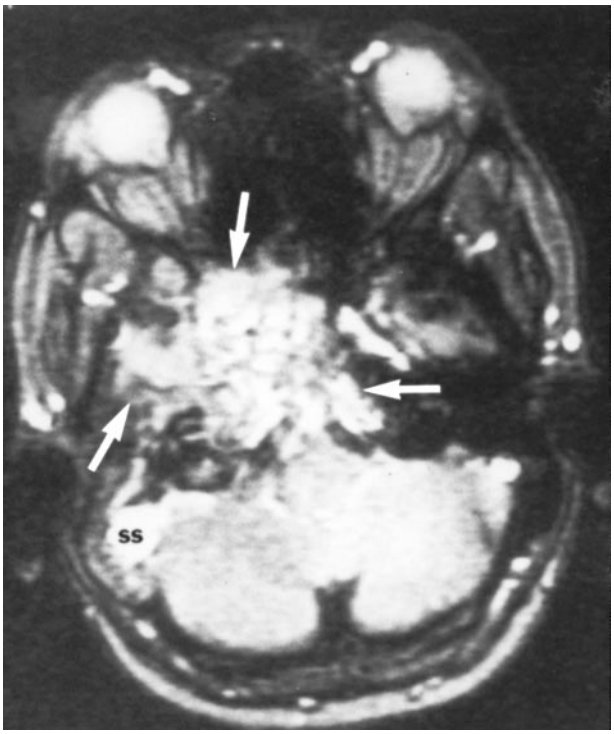


FIGURE 31–55. Diffuse glomus tumor extension in the nasopharynx. Dynamic magnetic resonance (MR) image shows an extensive mass (*arrows*) in the base of the skull with extension into the nasopharynx. With this MR dynamic GRASS (gradient-recalled acquisition in the steady state) technique, vascular tumors appear bright. Notice the hyperintensity of the vasculatures. ss = sigmoid sinus.

CT scanning. They can be evaluated in detail by CT and MRI in axial, coronal, and direct or reformatted sagittal sections (see Figure 31–16, B). The infratemporal fossa is an irregularly shaped cavity that lies below the infratemporal crest of the greater wing of the sphenoid, which is at a level with the upper border of the zygomatic arch. It is medial and deep to the zygomatic arch and is bound anteriorly by the infratemporal surface (posterior wall) of the maxilla and ridge from its zygomatic process (Figures 31–47 and 31–48). It lies lateral to the PPS[54–56] and is bound medially by a line from the styloid process to the lateral pterygoid plate[55] and laterally by the zygomatic arch and the articular tubercle of the temporal bone. Above the level of the zygomatic arch, the infratemporal fossa is continuous with the temporal fossa. Inferiorly, the infratemporal fossa has no floor but is continuous with the space external to the buccinator muscle and buccal space (Figure 31–57).[55] This space contains most of the mandible; pterygoid, masseter, part of the temporalis muscles; deep lobe of the parotid gland; internal maxillary vessels; mandibular and maxillary nerves; and pterygoid plexus of veins.[55–56] The foramen ovale, foramen spinosum, and alveolar canal open into it.[56]

Pterygomaxillary Fissure

Along the superior and medial parts of the infratemporal fossa are two fissures that meet at right angles.

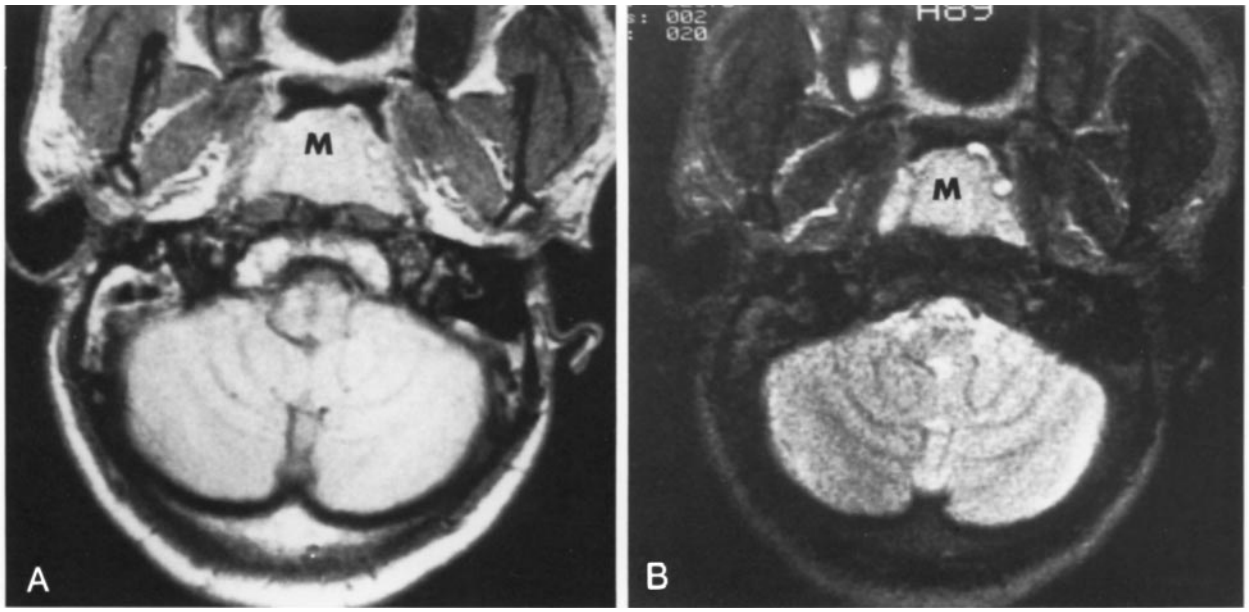


FIGURE 31–56. Scleroma of the nasopharynx. Proton-weighted (*A*) and T_2-weighted (*B*) axial magnetic resonance scans show a mass (M) in the nasopharynx.

The horizontal limb is the inferior orbital fissure, which is shown best in coronal and axial CT sections (see Figure 31–23), and the vertical limb is the pterygomaxillary fissure, which is formed by the divergence of the maxilla from the pterygoid process of the sphenoid and is shown best in sagittal (see Figure 31–16, B) and horizontal (axial) CT and MRI images. The pterygomaxillary fissure connects the infratemporal fossa with the pterygopalatine fossa (see Figure 31–16, B) and transmits the terminal part of the internal maxillary artery and veins.[55,56] The pterygopalatine (sphenomaxillary or sphenopalatine) fossa is a small, pyramidal space at the angle of the junction of the inferior orbital fissure and pterygomaxillary fissure at the apex of the orbit (see Figure 31–16, B). It is bound by the infratemporal (posterior) surface of the maxilla, the vertical plate of the palatine bone, and the base of the pterygoid process (see Figure 31–16, B). It communicates with the nasal cavity by the sphenopalatine foramen, with the infratemporal fossa by the pterygomaxillary fissure, and with the orbit by the inferior orbital fissure. The sphenopalatine foramen transmits the sphenopalatine vessels and the superior nasal

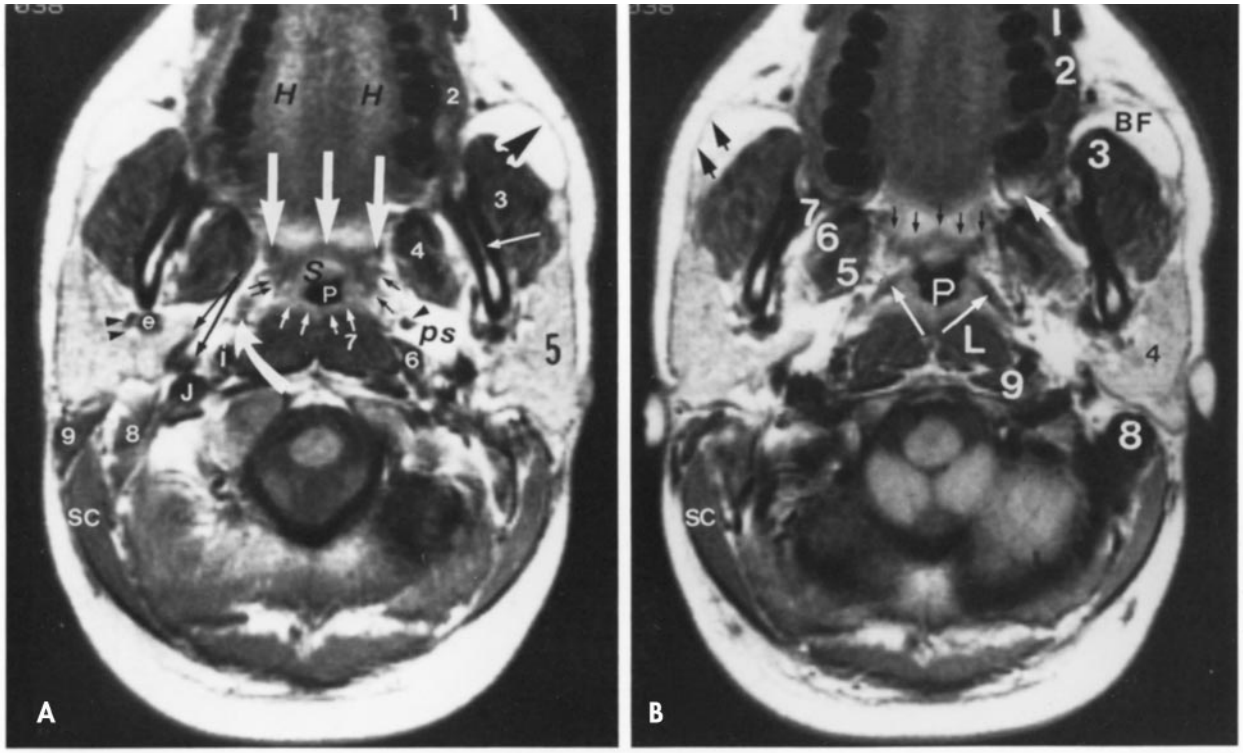


FIGURE 31–57. All magnetic resonance images (*A* to *E*) are proton density 2000/20 of the same normal subject. *A*, This section passes through the lower part of the alveolar process of the maxilla and upper part of the ramus of the mandible (*white arrow*) and shows the following structures: quadratus labii superioris (1), buccinator (2), masseter (3), and internal pterygoid (4) muscles; parotid gland (5); rectus capitis anterior (6), longus capitis (7), rectus capitis lateralis (8), digastric (9), splenius capitis (sc), and superior constrictor (*small white and black arrows*) muscles; pharynx (P); soft palate (S); parapharyngeal space (ps); pharyngeal vein (*black arrowhead*); retromandibular vein (*black arrowheads*); external (e) and internal (I) carotid arteries; internal jugular vein (J); pharyngopalatinus muscle (*large white arrows*); hard palate (H); and platysma muscle (*white and black arrowheads*). Note part of the levator veli palatini muscle (*curved white arrow*). *B*, This section is 3 mm superior to *A* and passes through the midportion of the alveolar process of the maxilla, upper end of the coronoid process of the mandible, and mastoid portion of the temporal bone (8). The following structures are seen: quadratus labii superioris (1), buccinator (2), and masseter (3) muscles; parotid gland (4); medial (5) and lateral (6) pterygoid and temporalis (7) muscles; mastoid process (8); rectus capitis anterior (9), longus capitis (L), splenius capitis (sc), levator veli palatini (*white arrows*), and pharyngopalatinus (*small black arrows*) muscles; nasopharynx (P); buccopharyngeal space (*white arrow*); buccal fat (BF); and platysma muscle (*black arrows*).

Continued on next page

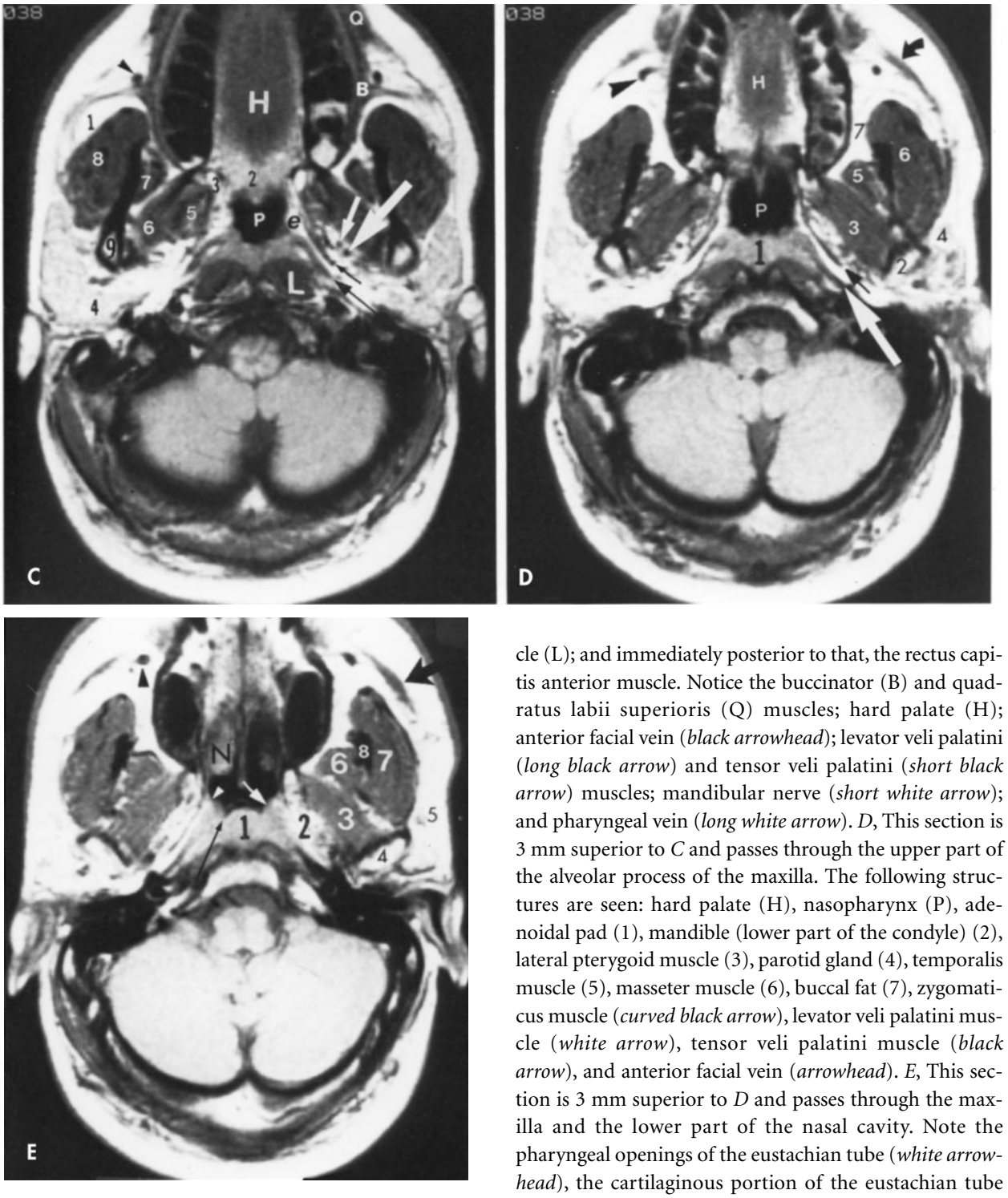


FIGURE 31–57 continued. *C*, This section is 3 mm superior to *B* and passes through the upper part of the alveolar process of the maxilla and upper part of the mandibular ramus. The following structures are seen: buccal fat (1); soft palate (2); parapharyngeal fat (3); parotid gland (4); medial (5) and lateral (6) pterygoid, temporalis (7), and masseter (8) muscles; mandible (9); eustachian tube (e); nasopharynx (P); longus capitis muscle (L); and immediately posterior to that, the rectus capitis anterior muscle. Notice the buccinator (B) and quadratus labii superioris (Q) muscles; hard palate (H); anterior facial vein (*black arrowhead*); levator veli palatini (*long black arrow*) and tensor veli palatini (*short black arrow*) muscles; mandibular nerve (*short white arrow*); and pharyngeal vein (*long white arrow*). *D*, This section is 3 mm superior to *C* and passes through the upper part of the alveolar process of the maxilla. The following structures are seen: hard palate (H), nasopharynx (P), adenoidal pad (1), mandible (lower part of the condyle) (2), lateral pterygoid muscle (3), parotid gland (4), temporalis muscle (5), masseter muscle (6), buccal fat (7), zygomaticus muscle (*curved black arrow*), levator veli palatini muscle (*white arrow*), tensor veli palatini muscle (*black arrow*), and anterior facial vein (*arrowhead*). *E*, This section is 3 mm superior to *D* and passes through the maxilla and the lower part of the nasal cavity. Note the pharyngeal openings of the eustachian tube (*white arrowhead*), the cartilaginous portion of the eustachian tube (torus tubarius) (*white arrow*), the pharyngeal recess or Rosenmüller's fossa (*long black arrow*), adenoidal pad (1), parapharyngeal space (2), lateral pterygoid muscle (3), mandible (condyle) (4), parotid gland (5), temporalis muscle (6), masseter muscle (7), coronoid process (8), zygomaticus muscle (*short black arrow*), anterior facial vein (*black arrowhead*), and nasal turbinate (N). Reproduced with permission from Mafee MF et al.[5]

and nasopalatine nerves. Five foramina open into the pterygopalatine fossa, including the foramen rotundum, pterygoid (vidian) canal, and sphenopalatine foramen, which, in turn, lead to the pterygopalatine canal (see Figures 31–16, B, and 31–23). The fourth foramen, the pharyngeal canal, which transmits the pharyngeal nerve and artery, is a potential communication between the nasopharynx and the fossa. The fifth foramen, the pterygopalatine canal, is placed inferiorly at the junction of the anterior and posterior walls and leads into the greater palatine canal. The pterygopalatine canal transmits the palatine vessels and nerves, serving as a potential communication with the oral cavity. The fossa contains the maxillary nerve, pterygopalatine ganglion, and terminal part of the internal maxillary artery.[55,56] Many of the superficial and deep muscles of the face are imaged on CT and MRI scans at the level of the infratemporal fossa and nasopharynx (see Figures 31–47 and 31–48). The lateral pterygoid muscle is a short, thick muscle, somewhat conical in shape, which extends almost horizontally between the infratemporal fossa and the condyle of the mandible (see Figure 31–47). It has its origins in the lateral surface of the lateral pterygoid plate and lateral surface of the greater wing of the sphenoid bone and inserts into the articular disk of the temporomandibular joint and the condyle of the mandible (see Figures 31–47 and 31–48).[55,56] The main mass of the medial pterygoid muscle lies deep and inferior to the lateral pterygoid muscle (see Figures 31–47 and 31–48) and arises from the medial surface of the lateral pterygoid plate, palatine bone, and maxilla and inserts into the angle and ramus of the mandible (see Figure 31–48).[55,56] In axial CT and MRI sections taken at the level of the low nasopharynx and soft palate, the medial pterygoid muscle occupies a position on the inside of the ramus of the mandible, similar to that of the masseter on the outside (see Figure 31–57). Both pterygoid muscles can be visualized in a coronal section taken at the level of the infratemporal fossa (see Figure 31–48, B). The masseter is a thick, somewhat quadrilateral muscle that arises from the zygomatic process of the maxilla and the zygomatic arch and inserts into the angle and lateral surface of the ramus and the lateral surface of the coronoid process of the mandible.[55,56] In axial and coronal sections, the masseter is seen as a thick, somewhat quadrilateral muscle, anterior and inferior to the parotid gland, occupying a position on the outside of the ramus of the mandible (see Figures 31–47, 31–48, and 31–57). The temporalis is a broad, radiating muscle, arising from the whole of the temporal fossa.[55,56] It passes inferiorly deep to the zygomatic arch and inserts into the medial surface, apex and anterior border of the coronoid process, and anterior aspect of the ramus of the mandible (see Figures 31–48 and 31–57).[55,56] The temporalis muscle is seen anterolateral to the lateral pterygoid muscle and pterygoid process as it inserts on the coronoid process. Symmetric, low-density, fascial-fatty planes are seen on CT scans between the lateral pterygoid and infratemporal muscles. They are continuous with the symmetric fat pad just anterior to the infratemporal muscles and posterior to the posterior wall of the maxillary sinus. Obliteration of these fascial-fatty planes is usually indicative of extension of a neoplastic or inflammatory process into them (Figure 31–58).

Parapharyngeal Space

The PPS is an irregular, predominantly fat-filled space within the suprahyoid neck (see Figures 31–47, 31–48, and 31–57). Various synonyms for PPS have been used, including lateral PPS, pharyngomasticatory space, pterygopharyngeal space, and pharyngomaxillary space. The PPS is subdivided into the prestyloid and retrostyloid compartments. Fascia that extends from the styloid process to the tensor veli palatini muscle crosses posteriorly in the PPS fat and separates the PPS into those two spaces. The carotid sheath is considered anatomically to be the posterior compartment of the PPS. Infection and both benign and malignant tumors of the infratemporal fossa (Figure 31–58) and PPS are manifested by obliteration or distortion of the normal soft tissue planes and/or by a mass effect.[56,57] Tumors and inflammatory processes in these areas are difficult to assess clinically (Figures 31–59, 31–60, and 31–61). However, they are identified readily by MRI and CT. Tumors of the PPS are rare, accounting for only 0.5% of head and neck neoplasms.[55] Neuromas of the vagus, trigeminal, sympathetic chain (Figures 31–62 and 31–63, A) and hypoglossal nerves and neurofibromas that arise within the PPS present as rounded masses bulging into the nasopharynx.[55,56] The benign nature of the tumor is often predicted by an evaluation of CT and MRI scans, which show mass effect but not obliteration or invasion of the fascial planes (see Figure 31–59). Eighty percent of PPS tumors are benign, and 20% are malignant. The most common

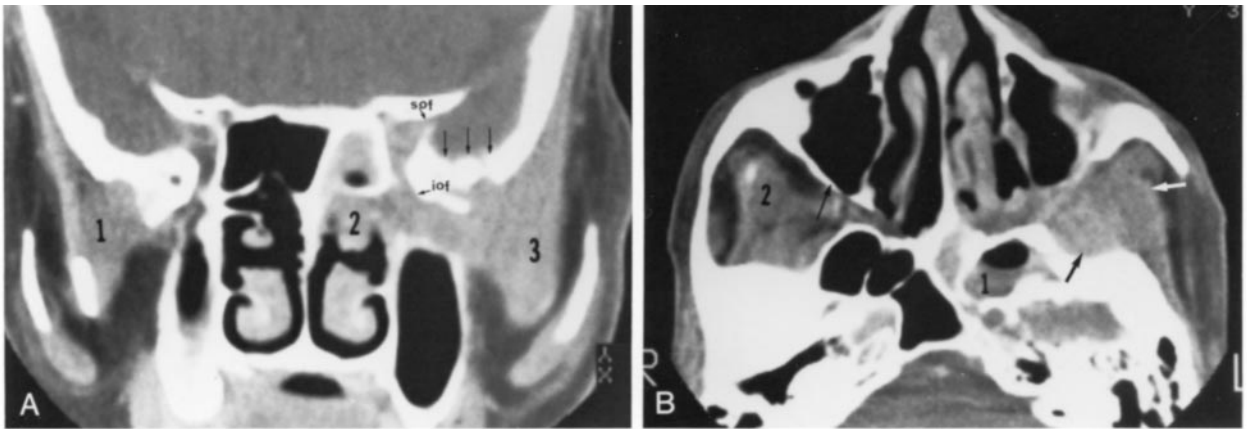


FIGURE 31–58. *A*, Mucormycosis. Semicoronal computed tomographic (CT) scan showing a soft tissue density in the left nasal cavity (2) and left sphenoid sinus, with bone destruction along the superior aspect of the lateral wall of the left nasal cavity and soft tissue infiltration along the inferior orbital fissure (iof) and into the apex of the left orbit along the superior orbital fissure (sof) (compare with the density of the right inferior and superior orbital fissures). There is also involvement of the left infratemporal fossa (3) with an irregularity of the bone of the left middle cranial fossa (*arrows*) caused by osteomyelitis. *B*, Mucormycosis. Postcontrast axial CT scan of the same patient in *A* showing soft tissue density in the left sphenoid sinus (1) and left nasal cavity, with extension into the left infratemporal fossa causing obliteration of the deep fascial plane (*white and black arrows*). Compare with the normal right infratemporal fossa (2). Note also involvement of the posterior portion of the left maxillary sinus as compared with the right side (*arrow*).

primary tumors are salivary gland neoplasms that originate from the deep lobe of the parotid gland or from ectopic minor salivary gland tissue. These tumors are located in the prestyloid PPS. On CT and MRI, these tumors displace the internal carotid artery posteriorly (Figures 31–59, 31–60, and 31–61). Neurogenic tumors and paragangliomas are the most common tumors of poststyloid PPS. Unlike prestyloid PPS masses, these retrostyloid PPS masses result in displacement of the internal carotid artery anteriorly (Figures 31–62, 31–63, and 31–64). Tumors of the mandible are usually confined to the bone; however, extension deep into the adjacent infratemporal fossa may occur. Extension of the nasal, maxillary,

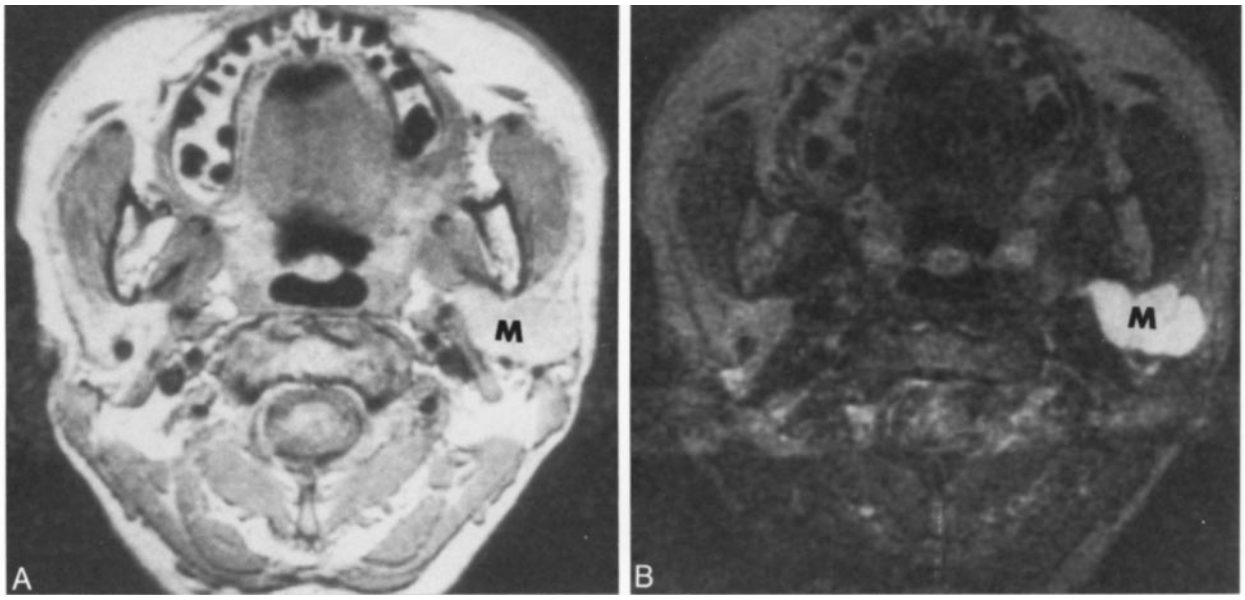


FIGURE 31–59. Pleomorphic adenoma. Proton-weighted (*A*) and T_2-weighted (*B*) axial magnetic resonance scans show a well-defined mass (M) involving the left parotid gland.

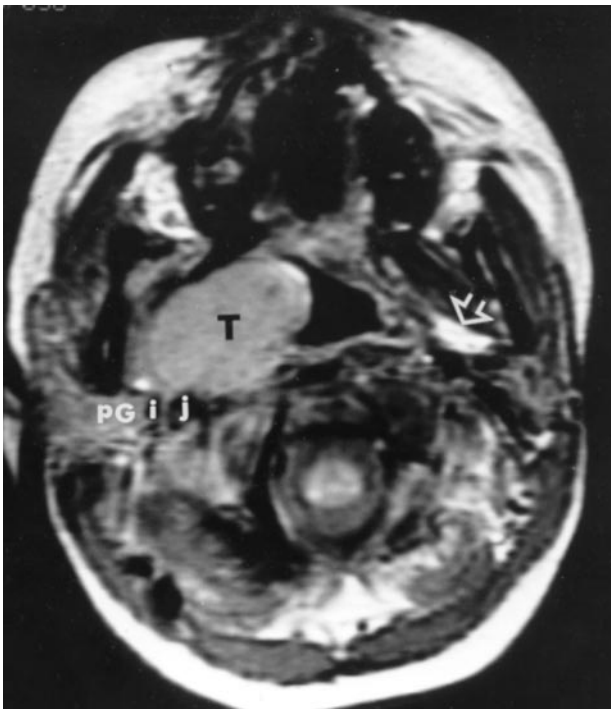


FIGURE 31–60. Parapharyngeal adenocarcinoma. Proton-weighted axial magnetic resonance scan shows a large tumor (T) involving the right parapharyngeal space. Notice the normal left parapharyngeal space (*arrow*). i = internal carotid artery; j = internal jugular vein; PG = parotid gland.

ethmoid, and sphenoidal lesions into the infratemporal fossa and the pterygopalatine fossa can be delineated clearly by MRI and CT scans.

Carotid Sheath The carotid sheath, forming a tubular investment for the carotid artery, internal jugular vein, and vagus nerve, is within the retrostyloid PPS,[55,56] abutting but separated from the infratemporal fossa by the "styloid diaphragm," an aponeurotic sheet of the pharyngobasilar fascia passing from the pharyngeal wall to the styloid process and its muscles.[55] Intravenous iodine-contrast material is given to define the internal carotid artery and internal jugular vein from the density of cranial nerves IX to XII, the sympathetic trunk, normal lymph nodes, and the density of three muscles arising on the styloid process (ie, stylohyoideus, stylopharyngeus, and styloglossus) (see Figure 31–48, A).

Tumors of the carotid space are often neurogenic or paraganglionic in nature (see Figure 31–61). Invasion of the carotid artery by metastatic nodal disease can be demonstrated by CT and MRI. Both techniques may not be reliable for carotid artery invasion, particularly if the artery is not completely surrounded by invading tumor.

Base of the Skull

The base of the skull is an anatomic region that can be best evaluated using CT. The combination of CT and MRI provides maximum information concerning hard tissue and soft tissue structures. Part of the anatomy of the base of the skull has been reviewed in the sections on the nasal and paranasal cavities. Many of the important structures of the skull base are illustrated in Figures 31–16 and 31–21 to 31–23. The unique ability of CT to image both bone and soft tissue structures with excellent tissue-contrast discrimination (resolution) makes it ideal to study the base of the skull and temporal bones. Both CT and MRI provide an accurate assessment of the total extent of disease and precise identification of intraorbital or intracranial extension of the tumor or inflammatory process. Computed tomography is superior to MRI for a detailed bony process (see Figures 31–16 and 31–23), and MRI is superior to CT for a detailed soft tissue process. In many instances, they are complementary to each other. Use of the

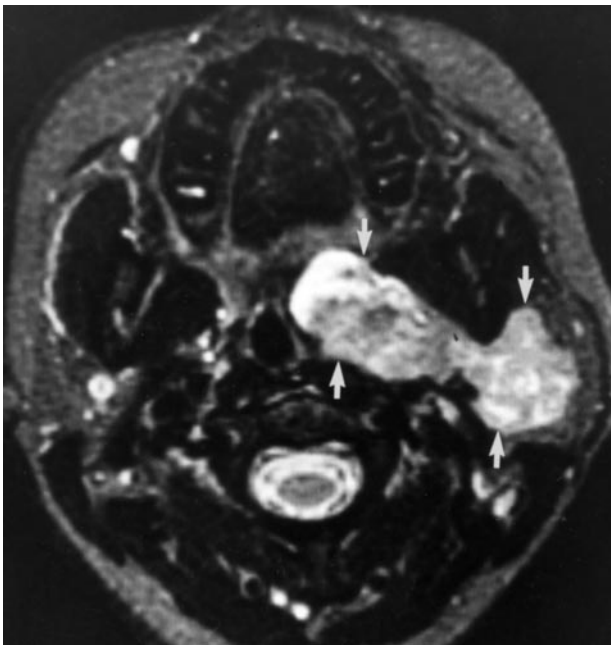


FIGURE 31–61. Malignant tumor of the parotid gland. T_2-weighted axial magnetic resonance scan shows a large mass involving the deep lobe of the left parotid gland, extending into the parapharyngeal space (*arrows*). This was reported as being compatible with liposarcoma.

extended-bone CT range (scale) high-resolution scan is a major imaging advantage and provides the best diagnostic bone detail available to date (see Figures 31–16 and 31–21 to 31–23).

Applications of Magnetic Resonance Imaging and Computed Tomography in Pathologic Conditions of the Skull Base

With MRI and CT, the diagnosis and management of patients with disease involving the skull base, nasopharynx, and head and neck lesions have been advanced significantly.[55–60] Primary and secondary tumors of the base of the skull cause expansion, destruction, or sclerotic reaction of the bones of the skull base. The base of the skull is an anatomic region where a variety of congenital, inflammatory, neoplastic, and traumatic lesions may develop. Computed tomography and MRI are the most sensitive methods of detecting various lesions of the skull base (Figure 31–65). In general, skull base lesions can be divided into three categories: (1) intracranial lesions that involve the skull base; (2) primary or secondary osseous lesions, as well as fractures; and (3) extracranial lesions that involve the skull base. The intracranial lesions that affect the skull base may result from congenital anomalies (see Figure 31–4, A); developmental inclusion cysts (epidermal, dermoid) (Figure 31–66); craniopharyngiomas; neuronteric cysts, neoplasms such as meningiomas, chordomas, aggressive pituitary adenomas, and neuromas (Figure 31–67); and vascular lesions such as aneurysms. Intra-axial lesions of the

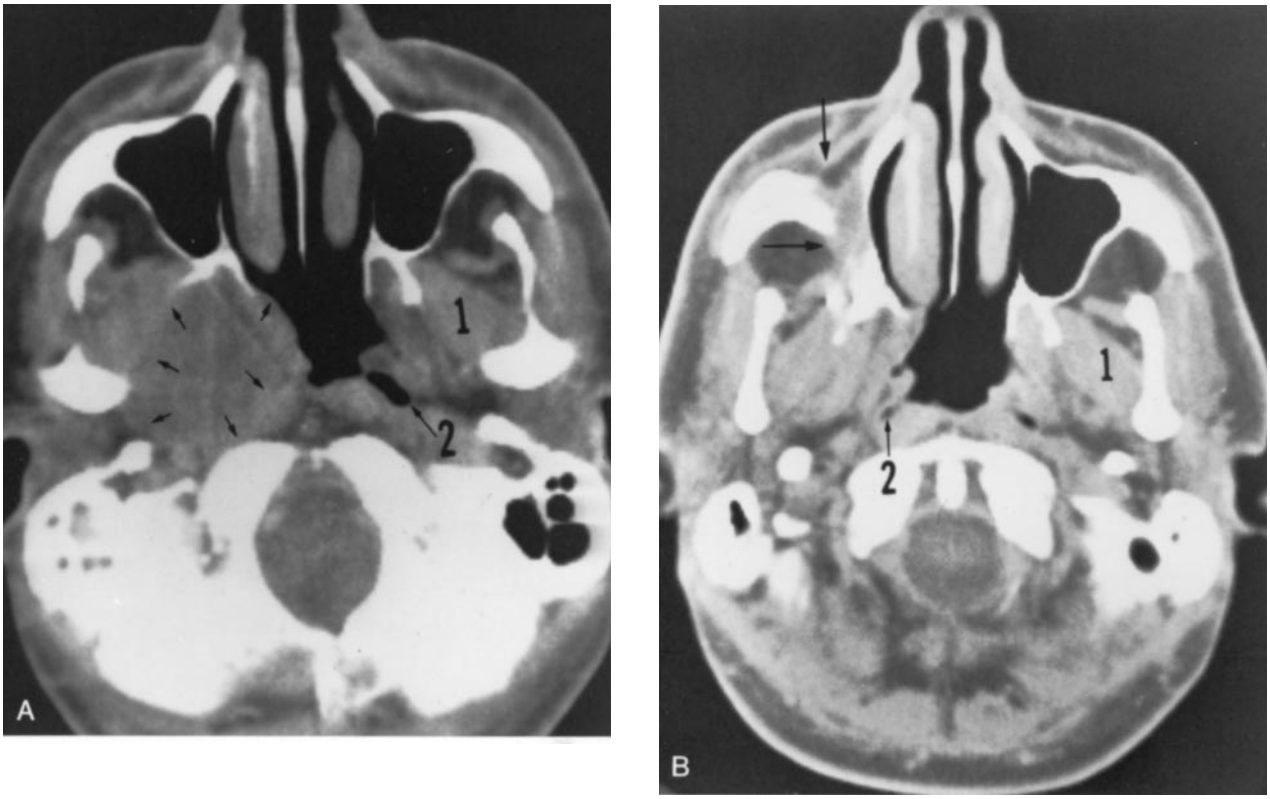


FIGURE 31–62. *A*, Paranasopharyngeal neurofibroma. Axial computed tomographic (CT) scan in a patient with a 12-year history of recurrent right serous otitis media showing a large, smoothly marginated mass (*arrows*) in the right parapharyngeal region that is protruding into the nasopharynx. Note the normal Rosenmüller's fossa on the left (2) and its obliteration on the right. Note also expansion of the pterygoid fossa without bone destruction as well as distortion of the right pterygoid muscle, all of which indicate the benign and slow-growing nature of the lesion. 1 = left lateral pterygoid muscle. *B*, Parapharyngeal neurofibroma. Axial CT scan of the same patient in *A* following extended transantral removal of the right parapharyngeal tumor. The superficial landmarks of the right side of the nasopharynx (effaced in *A*) are symmetric with those on the left and appear normal in this CT scan. Notice extension of air in the right Rosenmüller's fossa (2) and right eustachian tube and the symmetric appearance of lateral pterygoid muscles (1) and parapharyngeal spaces. Note postsurgical changes of the maxillary sinus (*arrow*).

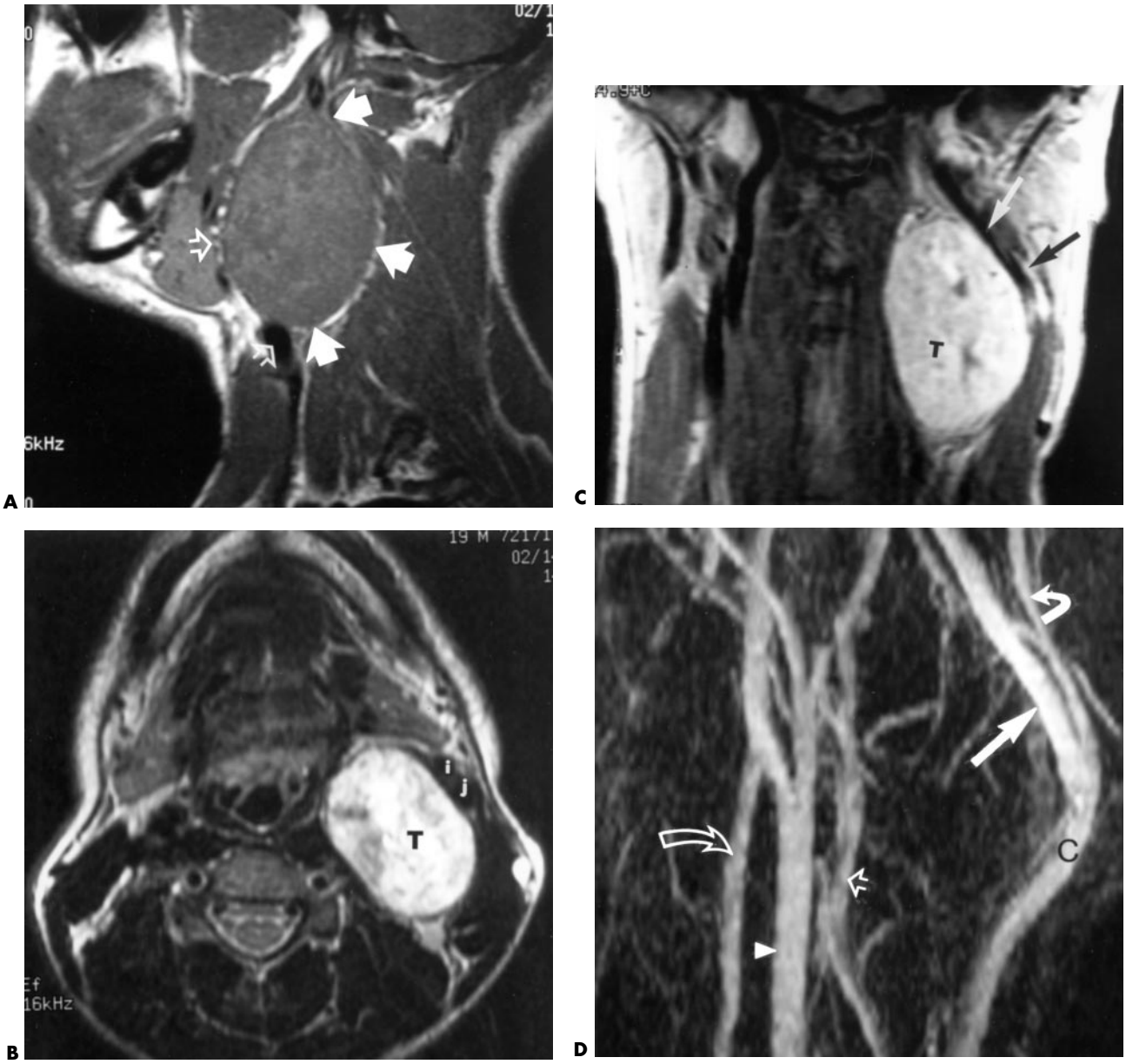


FIGURE 31–63. *A*, Parapharyngeal schwannoma. T_1-weighted sagittal magnetic resonance scan shows a large mass (*large arrows*), resulting in anterior displacement of the internal carotid artery (*hollow arrows*). *B*, T_2-weighted axial magnetic resonance (MR) scan of the same patient in *A* shows marked hyperintensity of the tumor (T). i = internal carotid artery; j = internal jugular vein. *C*, Coronal MR scan of the same patient in *A* and *B* shows the tumor (T) and laterally displaced carotid artery (*arrows*). *D*, Two-dimensional time-of-flight magnetic resonance angiogram showing laterally displaced common (C), internal (*straight arrow*), and external (*curved arrow*) carotid arteries. Note the left vertebral artery (*hollow arrow*), right common carotid artery (*arrowhead*), and right vertebral artery (*curved hollow arrow*).

brain do not affect the skull base. Primary osseous lesions of the skull base include congenital cholesteatoma (epidermoid) (see Figure 31–66), cholesterol granuloma (hematic cyst) (Figure 31–68), osteoma, osteoid osteoma, osteoblastoma, osteoclastoma (giant cell tumor), intraosseous hemangioma, fibrous dysplasia, chordoma, chondroma, chondromyxoid tumor chondrosarcoma, osteogenic sarcoma, lymphoma including plasmacytoma, Langerhans' histiocytosis, hemangiopericytoma, aggressive fibromatosis, and metastasis. Metastatic lesions are most commonly from primary tumors of the lung, breast, prostate, thyroid, and kidney; leukemia; and neuroblastoma (children).

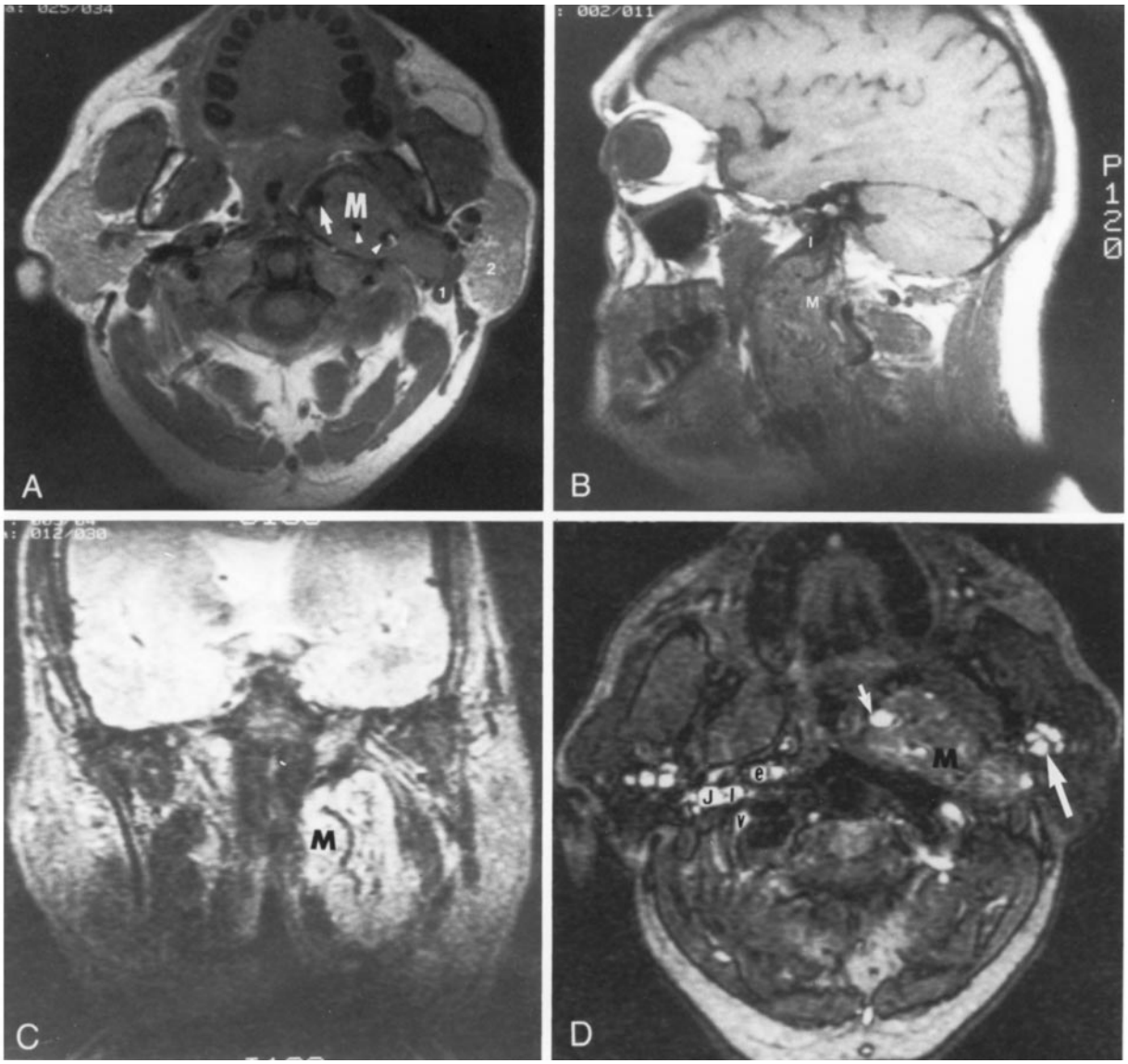


FIGURE 31–64. Carotid body tumor. *A*, Proton-weighted axial magnetic resonance (MR) scan shows a large mass (M) displacing the internal carotid artery (*arrow*) toward the parapharyngeal space. Note prominent vessels within the mass (*arrowheads*). 1 = posterior belly of the digastric muscle; 2 = parotid gland. *B*, Sagittal T_1-weighted MR scan showing the carotid body tumor mass (M). Note marked anterior displacement of the internal carotid artery (I) and serpiginous low-density vascular structures within the mass. *C*, T_2-weighted coronal MR scan showing a hyperintense mass (M) with several tortuous low-density images representing prominent vessels within the mass. The location of the mass as seen in sagittal and coronal MR scans is high, raising the possibility of a glomus vagale rather than a carotid body tumor. *D*, A dynamic GRASS (gradient-recalled acquisition in the steady state) axial MR image showing the tumor (M). Note the hyperintensity of the arteries and veins in this section. e = external carotid artery; I = internal carotid artery; J = internal jugular vein; v = vertebral artery. Note the retromandibular veins (*large arrow*) and displaced left internal carotid artery (*small arrow*). Reproduced with permission from Mafee MF et al.[5]

The extracranial lesions commonly affecting the skull base include carcinoma of the nasopharynx; tumors and infections of the sinonasal cavities (see Figure 31–40); necrotizing otitis externa, glomus complex tumors, and carcinoma of the temporal bone (Figure 31–69); hemangiopericytoma; juvenile angiofibroma; neurogenic tumors; and metastatic retropharyngeal and skull base nodal deposits.

Tumors and inflammatory lesions of the paranasal sinuses, orbits, or nasal cavities, which

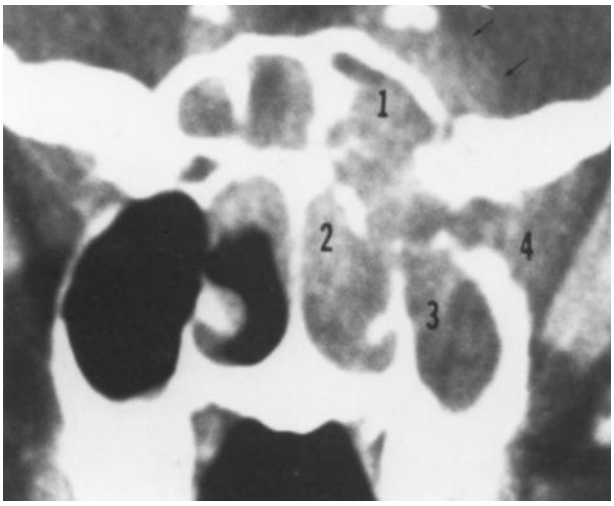


FIGURE 31–65. Nasopharyngeal carcinoma. Postcontrast semicoronal computed tomographic scan of the nasopharynx showing a large nasopharyngeal tumor (2) with extension into the sphenoid sinus (1), maxillary sinus (3), and infratemporal fossa (4). There is also tumor extension into the apex of the orbit and along the cavernous sinus. Note extensive irregular bone destruction of the floor of the sphenoid sinus.

may involve the skull base, include mucoceles (see Figure 31–37), carcinoma (see Figure 31–41), esthesioneuroblastoma (see Figure 31–41), lacrimal gland tumors (adenoid cystic carcinoma), rhabdomyosarcoma (see Figure 31–41), lymphoma, Wegener's granulomatosis, and Langerhans' histiocytosis. In adults, the most common lesion of the sinuses that invades the skull base is squamous cell carcinoma. In younger age groups, lymphoma is more common. In patients with diabetes and those who are immunocompromised, opportunistic fungus infections, such as mucormycosis and aspergillosis, involve the sinonasal cavities and rapidly invade the skull base (see Figure 31–58). Rhinocerebral mucormycosis is a well-recognized disorder in patients with diabetic ketoacidosis, but about half of the patients with diabetes who have this infection do not have ketoacidosis. The infection can be acute, subacute, or chronic. *Aspergillus* species are the most common cause of chronic fungal sinusitis. Involvement of the sphenoid sinus may extend to adjacent structures such as the orbit, cavernous sinuses, and temporal lobe, either directly or by retrograde phlebitis, with relative sparing of bone.

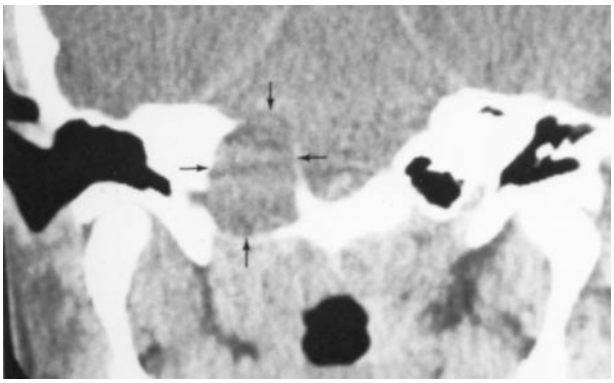


FIGURE 31–66. Primary cholesteatoma (epidermoid cyst). Contrast-enhanced coronal computed tomographic scan showing a nonenhanced low-density image in the region of the right petrous apex (*arrows*). Note moderate enhancement only in the peripheral portion (capsule) of the lesion. Lack of enhancement, relatively low density, its expansile nature, and smooth border of destroyed bone are almost characteristic of an epidermoid cyst. Reproduced with permission from Mafee MF et al.[58]

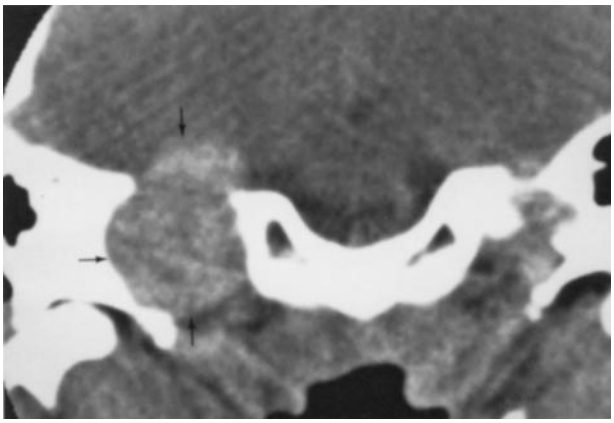


FIGURE 31–67. Schwannoma of the jugular fossa. Contrast-enhanced coronal computed tomographic scan showing a moderately enhanced soft tissue mass in the region of the right enlarged and eroded jugular fossa (*arrows*). The extra- and intracranial components of the tumor are clearly visualized. Note the smooth border of the destroyed bone similar to that in Figure 31–66, which is indicative of the slow-growing nature of the lesion as opposed to irregular bone destruction seen with the malignant process in Figure 31–70. Reproduced with permission from Mafee MF et al.[58]

Benign lesions such as primary cholesteatoma (see Figure 31–66) and cholesterol granuloma (see Figure 31–68) involving the base of the skull demonstrate characteristic bone expansion, with CT showing a low-density image (because of desquamated debris and cholesterol content) and no enhancement after intravenous injection of iodinated contrast

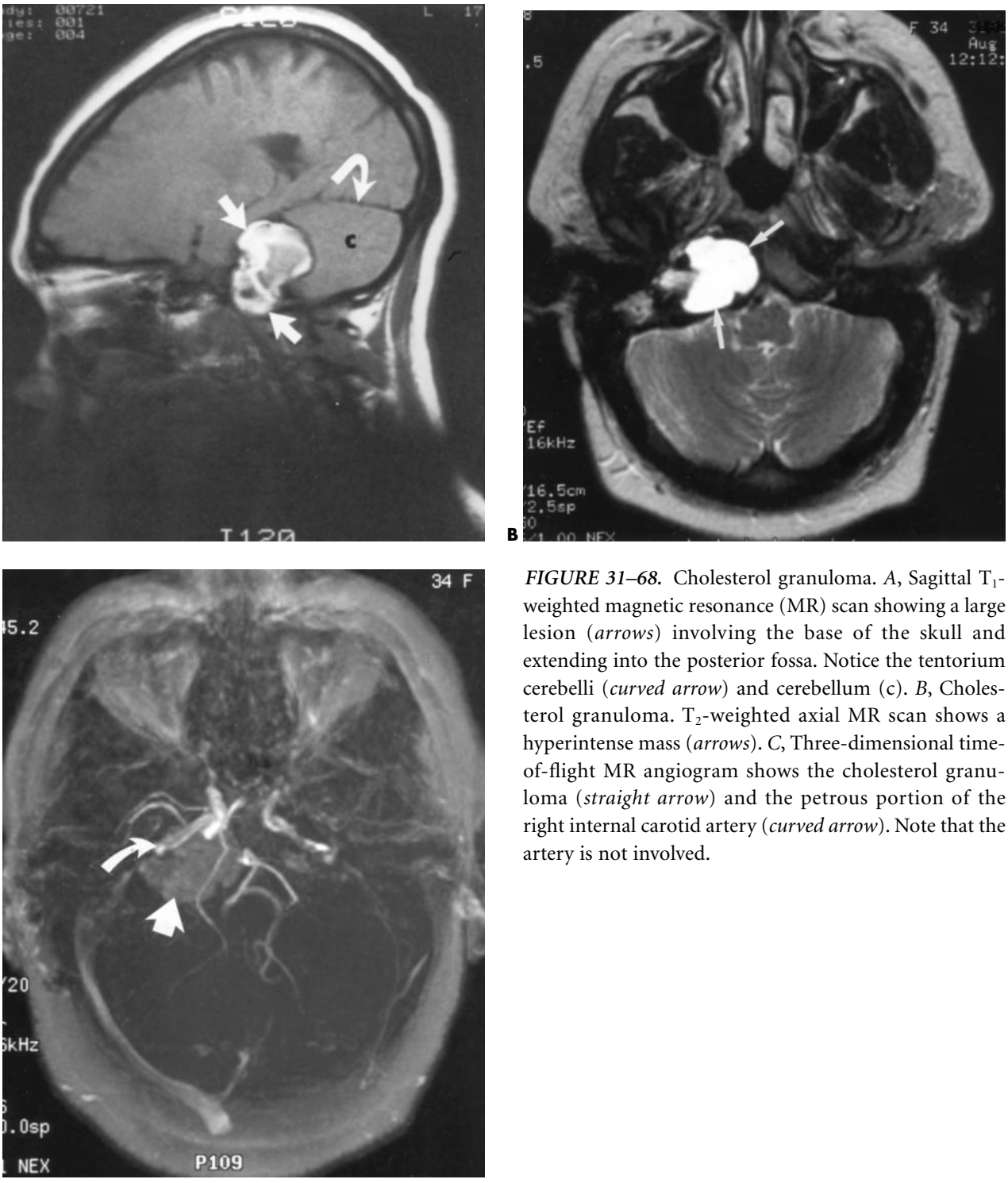
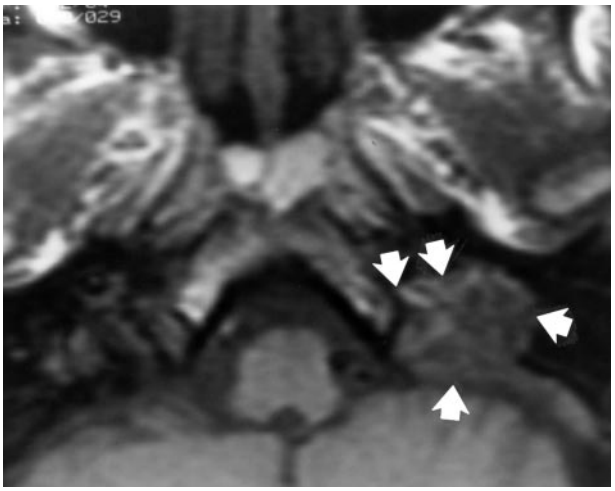
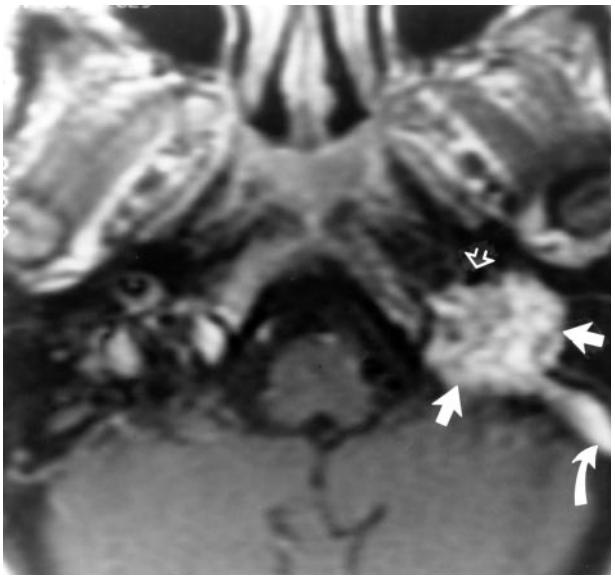


FIGURE 31–68. Cholesterol granuloma. *A*, Sagittal T_1-weighted magnetic resonance (MR) scan showing a large lesion (*arrows*) involving the base of the skull and extending into the posterior fossa. Notice the tentorium cerebelli (*curved arrow*) and cerebellum (c). *B*, Cholesterol granuloma. T_2-weighted axial MR scan shows a hyperintense mass (*arrows*). *C*, Three-dimensional time-of-flight MR angiogram shows the cholesterol granuloma (*straight arrow*) and the petrous portion of the right internal carotid artery (*curved arrow*). Note that the artery is not involved.

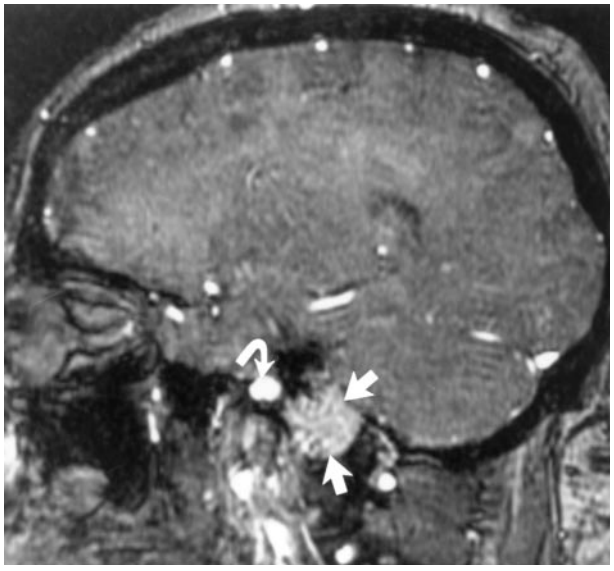
medium (see Figure 31–66).[58–60] Although cholesteatomas may not be differentiated from cholesterol granulomas on CT scans, they often can be easily differentiated by MRI. On T_1W images, epidermoid cysts (cholesteatomas) are seen as a hypointense or isointense image relative to brain and appear hyperintense to brain on T_2W MR images. Cholesterol granulomas, however, appear as a hyperintense image on both T_1W and T_2W MR scans (see Figure 31–68, B). Cholesterol granulomas may be homogeneous or heterogeneous. The heterogeneity of signal is usually owing to hemosiderin debris.[60] Primary



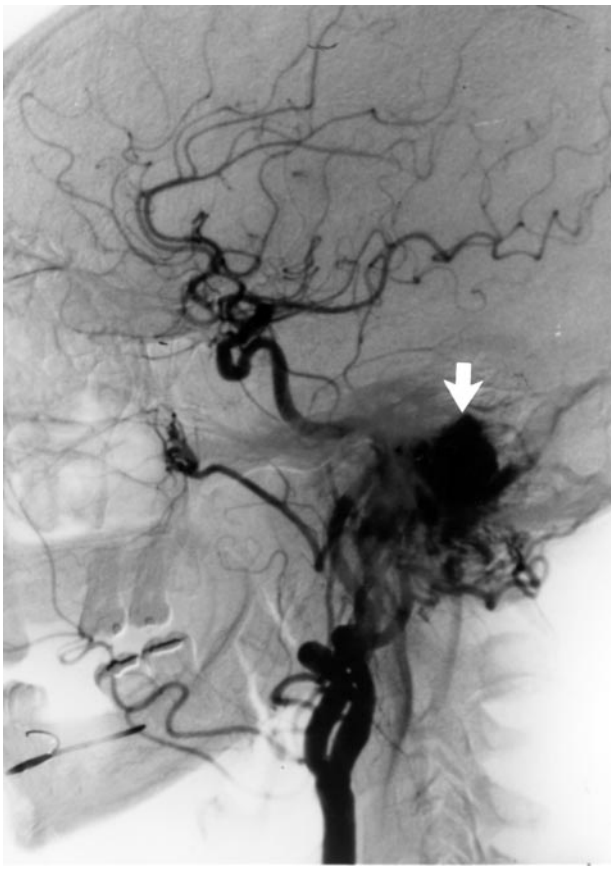
A



B



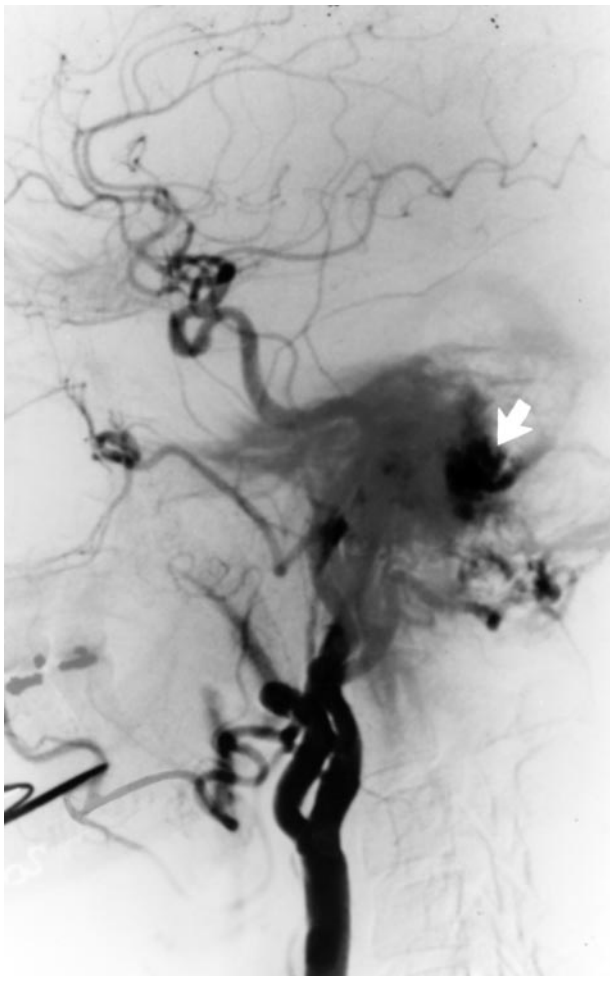
C



D

FIGURE 31–69. Glomus jugulare tumor. *A*, T_1-weighted axial precontrast magnetic resonance (MR) scan shows a large mass (*arrows*) in the left jugular fossa. *B*, T_1-weighted axial postcontrast MR scan shows marked enhancement of the glomus tumor (*straight arrows*). Notice the normal internal carotid artery (*hollow arrow*) and sigmoid sinus (*curved arrow*). *C*, A GRASS (gradient-recalled acquisition in the steady state) image shows tumor vascularity (*straight arrows*) and normal internal carotid artery (*curved arrow*). *D*, Standard angiogram shows marked vascularity (*arrow*) of another patient with a glomus jugulare tumor.

Continued on next page



E

FIGURE 31–69 continued. *E*, Postembolization angiogram of the same patient in *D* shows decreased vascularity of the tumor bed.

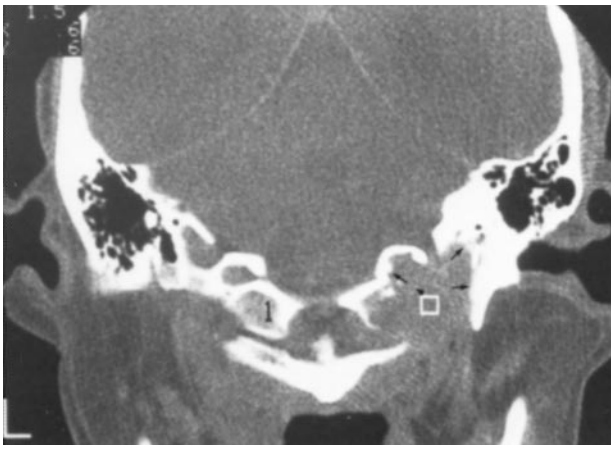


FIGURE 31–70. Liposarcoma of the base of the skull. Contrast-enhanced coronal computed tomographic scan showing extensive bone destruction of the right jugular fossa and occipital condyle (*arrows*). Note erosion of the right occipital condyle as compared with the left one (1).

cholesteatoma (epidermoid) can be differentiated from arachnoid cysts using FLAIR or diffusion-weighted (DW) MR pulse sequence. Cholesteatomas appear hypertense on FLAIR and DW MR images, whereas arachnoid cysts appear hypointense. Unlike cholesteatomas and cholesterol granulomas, lesions such as jugular fossa schwannoma and glomus jugulare tumors, beside causing bone destruction, will demonstrate moderate to marked enhancement after intravenous infusion of contrast medium (see Figure 31–69). On T_1W MR scans, schwannomas are seen as a hypointense image to brain and become hyperintense on T_2W MR scans. Glomus tumors also have a similar MR appearance; however, they usually show signal void (dark) areas because of their increased vascularity. The intracranial extension of a base of the skull lesion is best demonstrated on postcontrast-enhanced MRI. Destruction of the jugular fossa and hypoglossal canal is demonstrated in Figure 31–70 in a patient with liposarcoma of the base of the skull who presented with twelfth nerve palsy. The same appearance may be present on CT of a patient with metastatic carcinoma or lymphoma, including multiple myeloma. As far as the glomus complex tumors are concerned, CT reveals the bone destruction and total extent of the disease and gives accurate assessment of middle ear, intracranial, extracranial, and nasopharyngeal extension.[23,55] Dynamic CT has been valuable in the diagnosis and differentiation of glomus complex tumors from other pathologic entities such as meningioma and metastatic and recurrent neck tumors.[22,55] The MRI features of glomus jugulare tumors and their pre- and postembolization angiographic characteristics are illustrated in Figure 31–71.

Trauma Computed tomography is the modality of choice for fracture of the skull base. The site of the leak in CSF rhinorrhea or otorrhea can be detected most accurately using CT water-soluble contrast cisternography (see Figure 31–46).

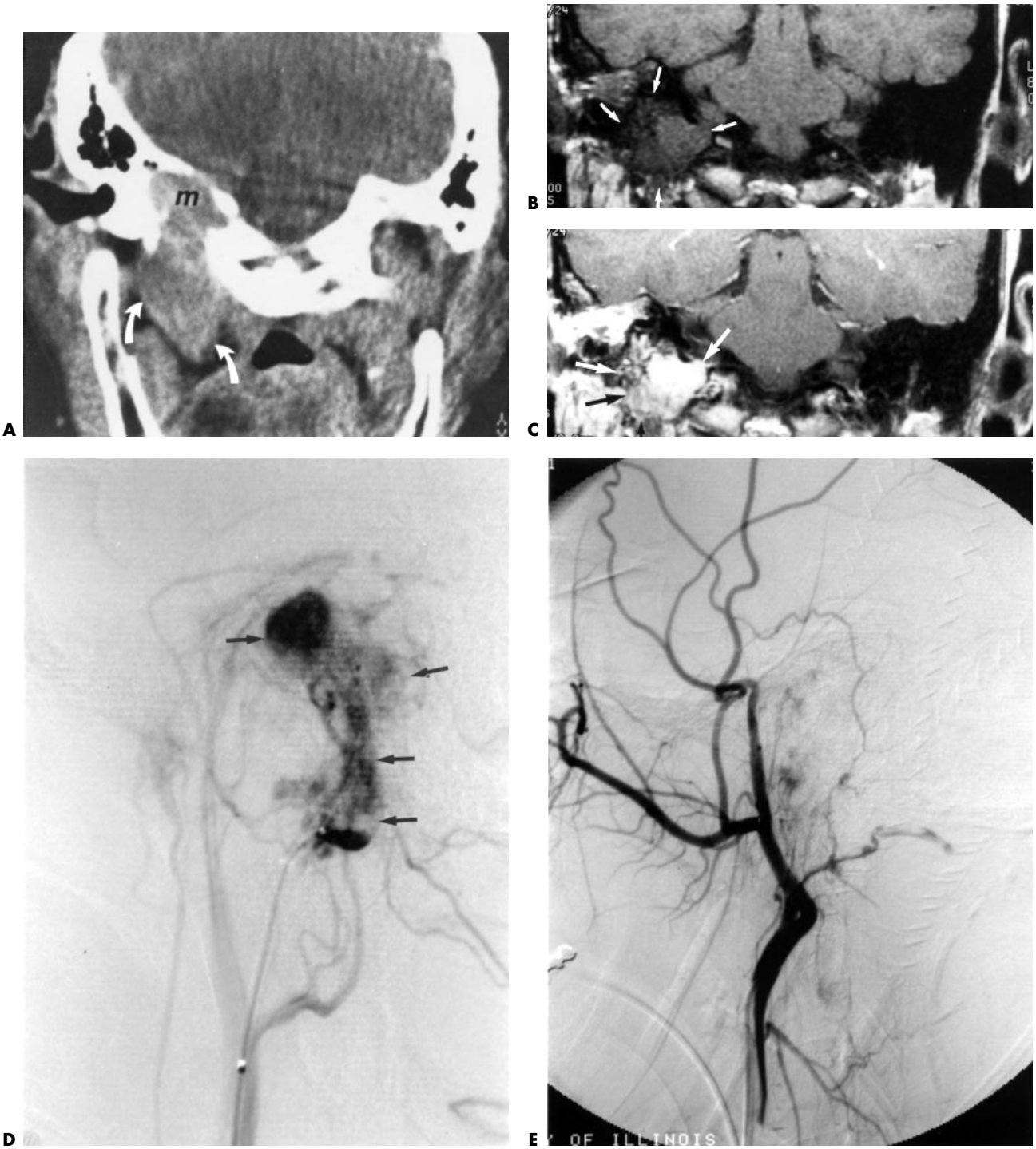


FIGURE 31–71. *A*, Coronal computed tomographic scan showing a soft-tissue mass (m) in the region of the right enlarged and eroded jugular fossa. Note extension of the tumor along the carotid sheath down to the level of the soft palate (*curved arrows*). *B*, Coronal T_1-weighted magnetic resonance (MR) scan shows a large mass in the left jugular fossa (*arrows*). *C*, Coronal post–gadolinium-diethylenetriamine pentaacetic acid T_1-weighted MR scan shows marked enhancement of the lesion (*arrows*). *D*, Standard angiogram shows marked vascularity of the tumor (*arrows*). *E*, Postembolization angiogram shows a marked decrease in tumor vascularity. At surgery, the tumor was completely resected, and the patient recovered with no neurologic deficit. (Courtesy of Arvind Kumar, MD, University of Illinois at Chicago Medical Center.)

REFERENCES

1. Lauterbur PC. Image formation by induced local interactions: examples employing nuclear magnetic resonance. Nature 1973;242:190–1.
2. McCullough EC. Basic physics of x-ray computed tomography. In: Krobkin M, Newton TH, editors. Computed tomography. St. Louis: CV Mosby; 1977.
3. Hounsfield GN. Picture quality of computed tomography. AJR Am J Roentgenol 1976;127:3–9.
4. Mafee MF, Rasouli F, Spigos DG, et al. Magnetic resonance imaging in the diagnosis of nonsquamous tumors of the head and neck. Otolaryngol Clin North Am 1986;19:523–36.
5. Mafee MF, Compos M, Raju S, et al. Head and neck, high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.
6. Mafee MF, Valvassori GE, Kumar A, et al. Tumors and tumor-like conditions of the middle ear and mastoid: role of CT and MRI: an analysis of 100 cases. Otolaryngol Clin North Am 1988;21:349–75.
7. Gorter CJ. Paramagnetic relaxation. Physica 1936;3: 503.
8. Rabi I, et al. A new method of measuring paramagnetic moment. Physiol Rev 1938;53:318.
9. Gorter CJ, Broer LJF. Negative results of an attempt to observe nuclear magnetic resonance in solids. Physica 1942;9:591.
10. Bloch F, et al. Nuclear induction. Physiol Rev 1946; 69:127.
11. Purcell EM, Torrey HG, Pound RV. Resonance absorption by nuclear magnetic moments in a solid. Physiol Rev 1946;69:37.
12. Rosen BR, Brady TS. Principles of nuclear magnetic resonance for medical application. Semin Nucl Med 1983;13:308–18.
13. Pykett IL, Newhouse JH, Buonanno FS, et al. Principles of nuclear magnetic resonance imaging. Radiology 1982;143:157–68.
14. Gore JC, Emery EW, Orr JS, et al. Special articles. Medical nuclear magnetic resonance imaging. I. Physical principles. Invest Radiol 1981;16:269–74.
15. Hawkes RC, Holland GN, Moore WS, et al. Nuclear magnetic resonance imaging: an overview. Radiography 1980;551:253–5.
16. Luyten PR, den Hollander JA, et al. Observation of metabolites in the human brain by MR spectroscopy. Radiology 1986;161:795–8.
17. Mafee MF. Magnetic resonance imaging and its simplifications for ophthalmology. In: Reinecke RD, editor. Ophthalmology annual. New York: Raven Press. 1989. p. 193–264.
18. Mafee MF. MRI and in-vivo proton H-1 spectroscopy of the lesions of the globe. Semin Ultrasound CT MRI 1988;9.
19. Smith SL. Nuclear magnetic resonance imaging. Anal Chem 1985;57:597–608.
20. Wehrli FW, Macfall JR, Newton TH. Parameters determining the appearance of NMR images. In: Newton TH, Potts DV, editors. Modern neuroradiology. Vol II. San Anselmo (CA): Clavadel Press; 1983.
21. Pisaneschi MJ, Mafee MF, Samii M. Applications of magnetic resonance angiography in head and neck pathology. Otolaryngol Clin North Am 1995;28: 543–61.
22. Mafee MF. Dynamic CT and its application to otolaryngology-head and neck surgery. J Otolaryngol 1982;11:307–18.
23. Shugar MA, Mafee MF. Diagnosis of carotid body tumors by dynamic computerized tomography. Head Neck Surg 1982;4:518–21.
24. Mafee MF, Valvassori GE, Shugar MA, et al. High resolution and dynamic sequential computed tomography. Use in the evaluation of glomus complex tumors. Arch Otolaryngol 1983;109:691–6.
25. Damadian R, Zaner K, Hor D, Dimaio T. Human tumors by NMR. Physiol Chem Phys 1973;5: 381– 402.
26. Mafee MF, Peyman GA, Grisolano JE, et al. Malignant uveal melanoma and simulating lesions: MR imaging evaluation. Radiology 1986;160:773–80.
27. Mafee MF, Peyman GA, Peace JH, et al. Magnetic resonance imaging in the evaluation and differentiation of uveal melanoma. Ophthalmology 1987;94: 341–8.
28. Chisin R, McCapinlac HA. The indications of FDG-PET in neck oncology. Radiol Clin North Am 2000; 38:999–1012.
29. Hesselink JR, New PF, Davis KR, et al. Computerized tomography of the paranasal sinuses and face: parts I and II. J Comput Assist Tomogr 1978;2:559–76.
30. Noyek AM, Kassel EE, Gruss JS, et al. Sophisticated CT and complex maxillofacial trauma. Laryngoscope 1982;92:1–17.
31. Mafee MF, Chow JM, Meyers R. Functional endoscopic sinus surgery: anatomy, CT screening, indications, and complications. Am J Radiol 1993;160: 735–44.

32. Chow JM, Mafee MF. Radiologic assessment preoperative to endoscopic sinus surgery. Otolaryngol Clin North Am 1989;22:691–701.
33. Mafee MF. Modern imaging of paranasal sinuses and the role of limited sinus computerized tomography; consideration of time, cost and radiation. Ear Nose Throat J 1994;8:532–46.
34. Mafee MF. Imaging methods for sinusitis. JAMA 1993;269:2808.
35. Amon JB. Computer-aided endoscopic sinus surgery. Laryngoscope 1998;108:949–61.
36. Loehrl TA, Toohill RJ, Smith TL. Use of computer-aided surgery for frontal sinus ventilation. Laryngoscope 2000;110:1962–7.
37. Hawkins DB, Clark RW. Orbital involvement in acute sinusitis. Lessons from 24 childhood patients. Clin Pediatr 1977;16:464–71.
38. Mafee MF. Imaging of the orbit. In: Valvassori GE, Mafee MF, Carter BL, editors. Imaging of the head and neck. New York: Georg Thieme Verlag; 1995. p. 158–215.
39. Eustis HS, Mafee MF, Walton C, Mondorca J. MR imaging and CT of orbital injections and complications in acute rhinosinusitis. Radiol Clin North Am 1998;36:1165–83.
40. Dobben GD, Philip B, Mafee MF, et al. Orbital subperiosteal hematoma, cholesterol granuloma and infection: evaluation with MR imaging and CT. Radiol Clin North Am 1998;36:1185–200.
41. Mafee MF, Carter BL. Nasal cavity and paranasal sinuses. In: Valvassori GE, Mafee MF, Carter BL, editors. Imaging of the head and neck. New York: Georg Thieme Verlag; 1995. p. 248–328.
42. Gori F, Nesi G, Pedemonte E. *Aspergillus* fungus balls on the mitral valve. N Engl J Med 2001;344:310–1.
43. Brown RB, Lau SK. Case 3-2001, the New England Journal of Medicine. N Engl J Med 2001;344:286–93.
44. Som PM, Shugar JM. The classification of ethmoid mucoceles. J Comput Assist Tomogr 1980;4: 199–203.
45. Mafee MF. Nonepithelial tumors of the paranasal sinuses and nasal cavity: role of CT and MR imaging. Radiol Clin North Am 1993;31:75–90.
46. Weber R, Draf W, Kratzch B, et al. Modern concepts of frontal sinus surgery. Laryngoscope 2001;111: 137–46.
47. Chow JM, Leonetti JP, Mafee MF. Epithelial tumors of the paranasal sinuses and nasal cavity. Radiol Clin North Am 1993;31:61–75.
48. Mafee MF, Pai E, Philip B. Rhabdomyosarcoma of the orbit: evaluation with MR imaging and CT. Radiol Clin North Am 1998;36:1215–27.
49. Wenig BM, Mafee MF, Ghosh L. Fibro-osseous, osseous, and cartilaginous lesions of the orbit and paraorbital region: correlative clinicopathologic and radiographic features, including the diagnostic role of CT and MR imaging. Radiol Clin North Am 1998;36:1241–59.
50. Chernak ES, Rodriquez-Antunez A, Jelden GL, et al. The use of computed tomography for radiation therapy treatment planning. Radiology 1975;117: 613–4.
51. Som PM, Brandwein M. Sinonasal cavities: inflammatory diseases, tumors, fractures, and postoperative findings. In: Som PM, Curtin HD, editors. Head and neck imaging. 3rd ed. St. Louis: CV Mosby: 1996. p. 51–276.
52. Drayer BP, Wilkins RH, Boehnke M, et al. Cerebrospinal fluid rhinorrhea demonstrated by metrizamide cisternography. AJR Am J Roentgenol 1977; 129:149–51.
53. Manelfe C, Cellerier P, Sobel D, et al. Cerebrospinal fluid rhinorrhea: evaluation with metrizamide cisternography. AJR Am J Roentgenol 1982;138:471–6.
54. Mancuso AA, Bohman L, Hanafee WN, Maxwell D. Computed tomography of the nasopharynx: normal and variants of normal. Radiology 1980;137: 113–21.
55. Mafee MF. Nasopharynx, parapharyngeal space and base of skull. In: Valvassori GE, Mafee MF, Carter BL, editors. Imaging of the head and neck. New York: Georg Thieme Verlag; 1995. p. 332–63.
56. Curtin HD. Separation of the masticator space from the parapharyngeal space. Radiology 1987;163: 195–204.
57. Som PM, Braun IF, Shapiro MD, et al. Tumors of the parapharyngeal space and upper neck: MR imaging characteristics. Radiology 1987;104:823–9.
58. Mafee MF, Valvassori GE, Dobben GD. Role of radiology in surgery of the ear and skull base. Otolaryngol Clin North Am 1982;15:723–54.
59. Mafee MF, Aimi K, Valvassori GE. Computed tomography in the diagnosis of primary tumors of the temporal bone. Laryngoscope 1984;94:1423–30.
60. Mafee MF, Kumar A, Heffner DK. Epidermoid cyst (cholesteatoma) and cholesterol granuloma of the temporal bone and epidermoid cysts affecting the brain. Neuroimag Clin North Am 1994;3:561–78.

CHAPTER 32

Allergic Rhinitis

Jayant M. Pinto, MD, Robert M. Naclerio, MD

Allergic rhinitis, the most common of all atopic diseases, is an important public health problem. It affects up to 20% of the adult population in the United States.[1] It is the sixth most prevalent chronic condition in the United States, and its prevalence is increasing.[2] In an otolaryngologist's practice, the prevalence of allergic rhinitis in patients coming to the office exceeds that of the general population and is in the range of 50%. Although it is a benign chronic disease of the upper airway, quality-of-life studies demonstrate that allergic rhinitis causes significant impairment of function, exceeding that of heart disease and asthma.[3,4] In the United States, the public health impact is significant, with approximately $1.5 billion spent annually on office visits and medications. Diminished functional capacity owing to allergic rhinitis alone causes the loss of 3.5 million work days and 2 million school days each year.[5] These numbers are augmented when one considers related disorders such as asthma and sinusitis, which are thought to be affected by allergic rhinitis.

Evidence supporting an increase in the incidence of allergic rhinitis comes from studies of Swedish military conscripts, with skin testing used to document the existence of atopy.[6] The prevalence of allergic rhinitis was 4.4% in this 18-year-old Swedish population in 1971 and increased to 8.1% by 1981. Similar trends have been observed in studies performed in the 1990s.[7,8] The cause of the increasing prevalence of allergic rhinitis is not clear. Because the increase has occurred within a decade, it does not represent a genetic shift in the population. Pollution may play a role by increasing nasal responsiveness to environmental allergens. Common outdoor pollutants include sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and ozone, whereas indoor pollutants include sidestream cigarette smoke. Exposure to high levels of SO_2, NO_2, and ozone leads to bronchial hyperresponsiveness and potentiates the pulmonary response to allergen, perhaps converting asymptomatic to symptomatic asthmatic patients.[9,10] Similarly, skin test–positive subjects who lack clinical symptoms of allergic rhinitis may be converted into symptomatic patients by exposure to environmental pollutants. Another purported role for pollution is that of an adjuvant in the development of immunoglobulin (Ig)E antibodies, which are central to the allergic reactions.[11] Studies in both animals and humans suggest that diesel exhaust particles enhance IgE production by a variety of mechanisms.[12] Evidence against the role of outdoor pollutants comes from a German study comparing the prevalence of rhinitis and asthma in genetically similar populations in a West and an East German city.[13] Despite higher levels of SO_2 and particulate matter in the air in the East German city, the incidence of rhinitis and asthma was considerably lower than in the comparison city in West Germany.

Another potential explanation for the apparent increased prevalence of allergic rhinitis involves the role of indoor allergens. Evaluation of specific IgE antibodies suggests that increased rates of sensitization occur more frequently to common indoor allergens such as dust mites and household pets than to outdoor allergens such as grass pollens. This study suggested that "Western"-style housing with better weatherproofing and decreased ventilation may be increasing the indoor allergen burden, leading to an increased rate of sensitization.[14] In support of this notion is a study of allergen avoidance in children of atopic parents.[15] Compared with a control group, reduction of allergen load in the first 2 years of life led to a lower than expected incidence of allergic rhinitis at the age of 2 years. Indeed, the prevalence in the former East Germany increased sharply during the years after reunification with the former West Germany, suggesting that factors associated with a Western affluent lifestyle may play a role in the development of allergic disease.[8] Although untested, proposed risk factors included

exposure to automobile exhaust, indoor heating and ventilation systems, indoor allergens, smaller family size, and a relatively higher socioeconomic status. Several other determinants have been identified, including breast-feeding, month of birth, and age of entry to day nursery.

Another hypothesis is that a decrease in infections in early childhood may affect the development of atopic diseases. This hypothesis is based on observations that show a higher prevalence of allergic diseases in individuals brought up in smaller and more affluent families, particularly in children and young adults in developed nations. The increase of allergic diseases is associated with certain socioeconomic conditions found in these countries. One common effect of these conditions is the decreasing prevalence of childhood infections that are assumed to have a protective effect for the development of allergy. This decreasing prevalence of early childhood infections can be explained by a lowered risk of exposure to infectious agents early in life as a result of increased immunization, improved sanitary conditions, and smaller family size.[16–25]

A skew in the balance between T-cell subpopulations from (normal) T helper 1 (TH1)-cell responses toward (atopic) TH2-cell responses, including increased production of IgE, may be caused by decreasing exposure to infectious agents. In support of the hygiene theory are studies demonstrating an inverse relationship between the frequency of allergic disorders and childhood infection.[16–18,25] Additional evidence comes from studies in which children attending day care have been shown to develop more infections than those who do not. A study of 1,035 children in Tucson, Arizona, showed that exposure to other children at day care was protective against the development of asthma and frequent wheezing in childhood.[25] Children with more exposure to other children were less likely to have wheezing after age 6 through age 13 years, although they wheezed more in the first 2 years of life. The mechanism of these phenomena is not known. However, signals to the developing immune system provided by infections may cause alterations in the functional responses of lymphocytes.[26] For example, the absence of a driving signal toward a TH1 response may allow the development of TH2 responses, resulting in a more atopic phenotype later in life (Figure 32–1).[26]

Symptoms of allergic rhinitis can begin at any age but are most frequently first reported in adolescence or young adulthood.[27] Rates of prevalence are similar for males and females, and no racial or ethnic variations are reported.[28] The incidence of developing an allergic diathesis is higher in children whose parents suffer from allergic rhinitis. If one parent has allergies, the chances of the child's having rhinitis are 29% and increase to 47% when both parents have the disease.[29] The rate of loss of allergic rhinitis symptoms has been estimated to be about 10%, which occurs in patients with the mildest form of the disease.[30]

A genetic component to allergic diseases, including asthma, allergic rhinitis, and others, has long been suggested. Several investigators have conducted genome-wide screens in a variety of populations in the search for susceptibility genes involved in allergic diseases, primarily asthma.[31–35] These atopic diseases may be linked by common mechanisms.[36] A number of chromosomal regions show evidence for linkage in multiple populations for asthma- and atopy-related phenotypes.[37] Positional cloning efforts are under way in these regions. Heterogeneity of phenotype, gene environment interaction, complexity of molecular mechanisms, an unknown mode of inheritance, and the possibility of small effects of a number of genes have complicated these approaches. To date, no loci have been conclusively identified for association with allergic rhinitis. However, positive associations of certain HLA (human leukocyte–associated antigens, also known as major histocompatibility complex [MHC]; for more information, see Chapter 28) alleles with dust and mite allergens have been reported.[38] This remains an area of intense interest and will likely accelerate with the recent completion of the first draft sequence of the human genome.

Part of the clinical importance of allergic rhinitis lies in its association with complications related to chronic nasal obstruction. Total nasal obstruction, regardless of cause, can result in sleep disturbances. Whether interference with sleep or the effects of systemically released mediators contribute to the fatigue associated with the disease is unknown. Evidence for a systemic role has been outlined.[39] Allergic rhinitis is the third most common cause of hyposmia.[40] Sinusitis often occurs in conjunction with allergic rhinitis. For example, the prevalence of abnormal sinus radiograph studies is greater in

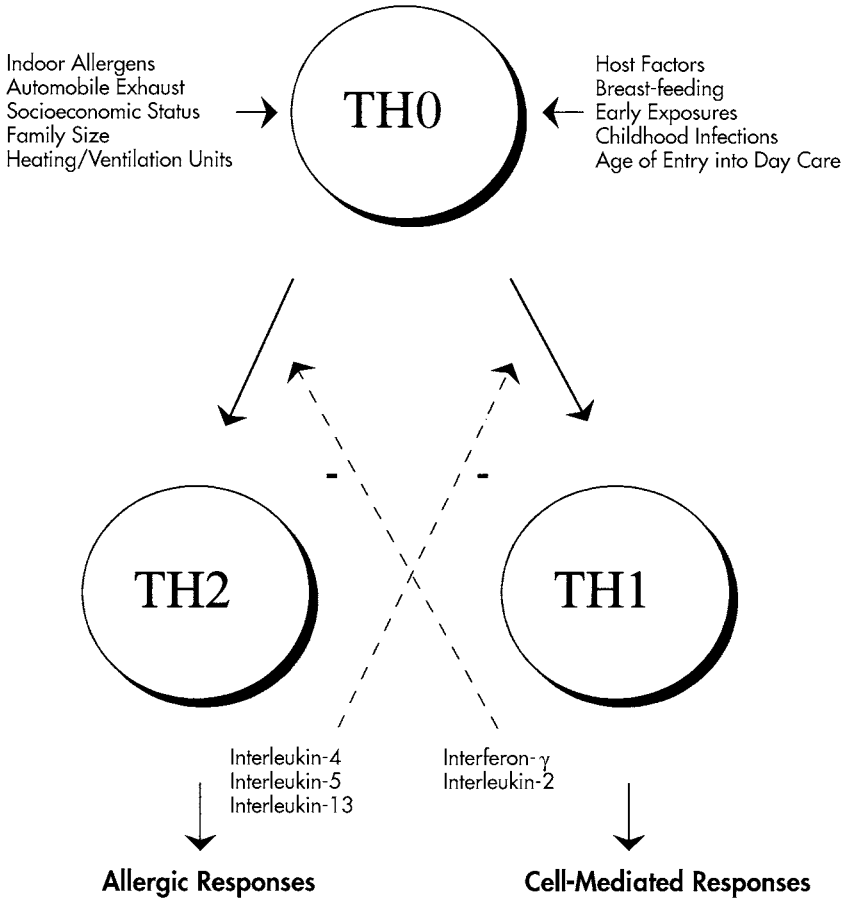


FIGURE 32–1. Current theory on the development of allergy disease. A variety of lines of evidence suggest that fewer respiratory illnesses during early childhood may predispose to later allergic disease. This may be attributable to signals to the developing immune system, driving it toward allergic responses via the TH2 pathway. Other environmental stimuli have been suggested to affect the development of allergic disease; these may similarly alter immune responses with resultant allergic disease, although they remain unproved. Cytokines produced in TH1 and TH2 responses inhibit the expression of the counterpart response. The TH1 pathway is primarily responsible for cell-mediated immunity, whereas the TH2 pathway is important in the production of allergic responses, including the production of immunoglobulin E.

patients with perennial allergic rhinitis compared with those with seasonal disease and approaches 50%.[41] Whether nasal polyps result from allergic rhinitis remains an open question. Settipane and Chafee, in an allergy practice setting, found that 2.1% of patients with allergic rhinitis had nasal polyps.[42] Although this is higher than the 0.3% incidence of nasal polyposis reported for the general population in the 1980 American Health Survey, it is less than the 12.5% incidence noted by Settipane and Chafee in nonallergic asthmatic patients.[42]

The contribution of allergic rhinitis to middle ear disease is debated. Studies by Friedman and colleagues suggest that an induced allergic reaction can cause eustachian tube dysfunction.[43] Moreover, allergic mediators have been found in the middle ear fluid of patients with otitis media with effusion.[44] Expression of TH2 cytokines and inflammatory cells has been documented in the middle ear mucosa of allergic patients.[45] Other data provide less support for a significant role of allergic rhinitis in the pathophysiology of otitis media.[46] Allergic rhinitis occurs in 80% of asthmatic subjects, and 40% of allergic rhinitic patients have asthma, suggesting a close link between the two diseases.

PATHOPHYSIOLOGY

IMMUNOGLOBULIN E

The distinguishing characteristic of allergic rhinitis is the involvement of IgE. In an individual with a susceptibility for developing allergic disease, an initial contact with allergen leads to the production of specific IgE molecules, a process termed sensitization. This begins when macrophages and other antigen-presenting cells process the allergen before presenting it to T-helper cells, which then interact with B lymphocytes, leading to their differentiation into IgE-producing plasma cells, a process involving interleukin (IL)-4 and IL-13 and accessory molecules such as CD40. These newly formed IgE molecules bind to high-affinity receptors principally located on mast cells and basophils and to low-affin-

ity receptors located on eosinophils, monocytes, and platelets. The high-affinity IgE receptors on mast cells mediate the initial allergic response. In the nasal mucosa of allergic subjects, mast cells are usually found within the mucosal connective tissue stroma, in the vicinity of blood vessels and glandular structures.[47] Seasonal exposure increases the number of mast cells located close to the surface in the nasal epithelium without affecting their overall number, suggesting that these cells migrate from the deeper to the more superficial layers of the nasal mucosa during the pollen season.[47] Recent studies demonstrating epsilon germline gene transcripts in biopsy specimens obtained from asymptomatic patients with seasonal allergic rhinitis and exposed to allergen ex vivo suggest that production of IgE may occur, at least in part, locally within the nasal mucosa.[48] This elevates the importance of local targeting of therapeutic interventions.[49]

Early Response

In a sensitized individual, another encounter with the same allergen provokes an immediate (early) allergic reaction. Within seconds of entering the nasal cavity, antigens interact with specific IgE molecules on the surface of mast cells. Cross-linking of two adjacent IgE molecules by antigen initiates a sequence of intracellular events resulting in mast cell degranulation and the subsequent release of preformed and newly generated mediators into the nasal milieu. The preformed mediators are stored within granules in the cytoplasm and include histamine, tryptase, heparin, and chondroitin sulfate. The newly generated mediators are synthesized from the phospholipids of the cell membrane subsequent to mast cell activation. These mediators include platelet activating factor (PAF), prostaglandins (PGs), and leukotrienes. Evidence suggests that mast cells produce ILs by IgE mechanisms, and these, unlike the lipid and granule mediators, require hours to be synthesized and released.[50,51] Released mediators can also initiate extracellular events that generate additional mediators; for example, the generation of kinins depends, in part, on the transudation of precursor proteins, kininogens, from the plasma that increase as a result of the action of histamine on vascular H_1 receptors. Other non–mast cell mediators are also associated with this early response to antigen.

For ascertaining the importance of any mediator in the pathophysiology of allergic rhinitis, three criteria should be satisfied: (1) the mediator should be present during the allergic reaction, (2) instilling the mediator in the nasal cavity should mimic part of the pathophysiology and symptoms of the disease, and (3) a mediator antagonist must partially or totally attenuate disease expression. To establish the importance of different mediators in allergic rhinitis, several investigators have attempted to satisfy one or more of these criteria.

Multiple mediators have been recovered at increased levels in nasal secretions after allergen challenge, and these include histamine,[52] kinins,[53] plasma and glandular kallikrein,[54,55] mast cell tryptase,[56] PGD_2,[53] leukotriene C_4 (LTC_4),[57] leukotriene B_4 (LTB_4),[58] and major basic protein (MBP)[59] (Figure 32–2). Furthermore, several of these mediators have been used to challenge the nasal mucosa and observe resultant responses. Histamine provocation, for example, produces rhinorrhea, congestion, pruritus, and sneezing by stimulating receptors on sensory nerves and blood vessels.[60] Instillation of serotonin produces sneezing,[61] whereas PGD_2[62] and kinins[63] all produce nasal congestion.[64] The other approach to establishing the importance of a mediator involves the use of antagonists; the best example is H_1 antihistamines: the clinical utility of H_1 receptor antagonists serves to demonstrate the importance of this inflammatory mediator.

Whereas mast cell degranulation and the released inflammatory mediators mimic most of the symptoms of allergic disease, the short duration of these events, as opposed to the prolonged symptoms of clinical disease, suggests that additional inflammatory processes are probably important in clinical disease. Furthermore, increased responsiveness and mucosal cellular changes that accompany chronic allergen exposure cannot be explained by this mechanism alone. Many of these phenomena relate to the cellular infiltration that is a hallmark of the inflammatory response.[65]

Late Response

Various techniques of experimental nasal allergen provocation, in addition to being useful for an understanding of the early response, assist investigators in understanding subsequent inflammatory events. Following challenge, there is an increase in

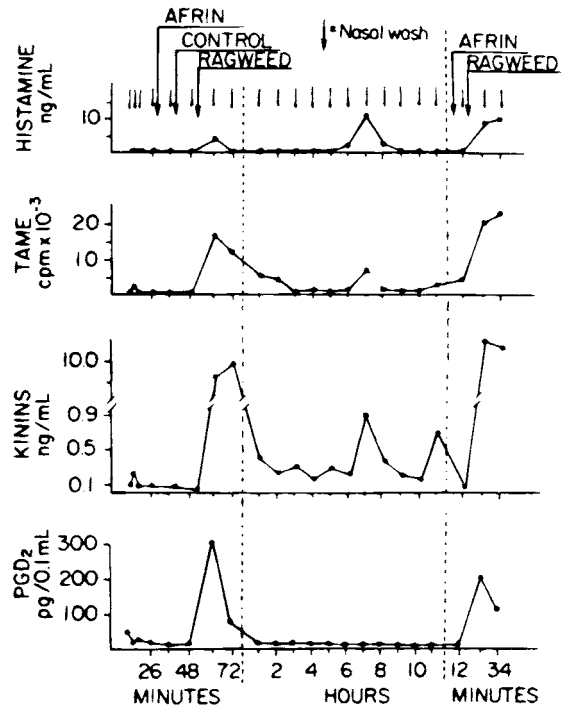


FIGURE 32–2. The nasal response to allergen provocation in allergic subjects challenged with ragweed extract. The abscissa depicts the time after allergen provocation: minutes = time points of the early response, hours = time points of the late response, and minutes (to the right) = time points of the rechallenge response. *Short arrows* at the top of the graph represent nasal lavages, and *long arrows* represent times of administration of different substances into the nasal cavities. There was a rise in all measured mediators during the early reaction. After a quiescent phase, there was a recurrence in elevations in the levels of histamine, tosyl-L-arginine methyl ester (TAME)-esterase, and kinins, but not prostaglandin D_2 (PGD_2) in nasal lavages. During rechallenge, there was an exaggeration of the nasal response to challenge with a dose identical to that used during the early reaction for histamine, TAME-esterase, and kinins. This is an illustration of the priming response, whereby subjects respond more vigorously to a dose of allergen after previous stimulation. Reproduced with permission from Naclerio RM et al.[67]

symptoms and mediator levels characteristic of the early allergic reaction, which soon returns toward baseline. If one continues to monitor the response for several hours, symptoms recur, associated with an elevation of levels of inflammatory mediators in approximately 50% of the patients, the late-phase response (Figure 32–3).[66,67] During the late-phase response, subjects have a recurrence of sneezing, rhinorrhea, and congestion, with congestion predominating.[68] The late reaction is also accompanied by increases of some, but not all, of the inflammatory mediators associated with the early reaction and the presence of mediators not seen in the early response.[66]

As in clinical disease, a striking characteristic of the late-phase response is cellular inflammation. One sees an increase in basophils, eosinophils, neutrophils, and mononuclear inflammatory cells in nasal secretions, as recovered by nasal lavage.[69] Nasal mucosal biopsies, performed 24 hours after topical allergen provocation, also show increases in the number of inflammatory cells, but, in contrast to the nasal secretions, in which eosinophils and neutrophils account for the majority of recovered cells, mononuclear cells predominate.[70] These are mostly of the T-helper category, as evidenced by positive staining with antibodies against CD4, and a portion of them are activated, as documented by IL-2 receptor expression (CD25+).[71,72] These T lymphocytes have been shown to express messenger ribonucleic acid (mRNA) for TH2-type cytokines.

The cellular changes observed after allergen provocation are similar to observations of nasal cytology during seasonal disease. In separately conducted studies, Bryan and Bryan and Okuda and Otsuka observed basophilic cells in nasal secretions during seasonal pollen exposure of allergic individuals.[73–76] In contrast to nasal secretions, which represent the most superficial compartment of the nasal mucosa, examination of nasal mucosal scrapings,[77,78] or biopsies[47,77,78] that sample deeper layers, showed that the majority of metachromatic cells in these compartments were mast cells. Enerbäck and colleagues, using biopsies and cytologic imprints to examine the cellular content of the nasal mucosa during the birch pollen season in Sweden, showed a seasonal increase in mast cell number on the surface of the nasal epithelium after 4 or 5 days of pollen exposure.[47] Mast cells are found in high concentrations beneath epithelial surfaces, whereas their counterparts, basophils, circulate in the bloodstream. The consensus in most studies is that basophils predominate in nasal secretions, whereas mast cells are more abundant in the epithelium and lamina propria of allergic subjects exposed to antigen either experimentally or naturally. Bentley and colleagues also

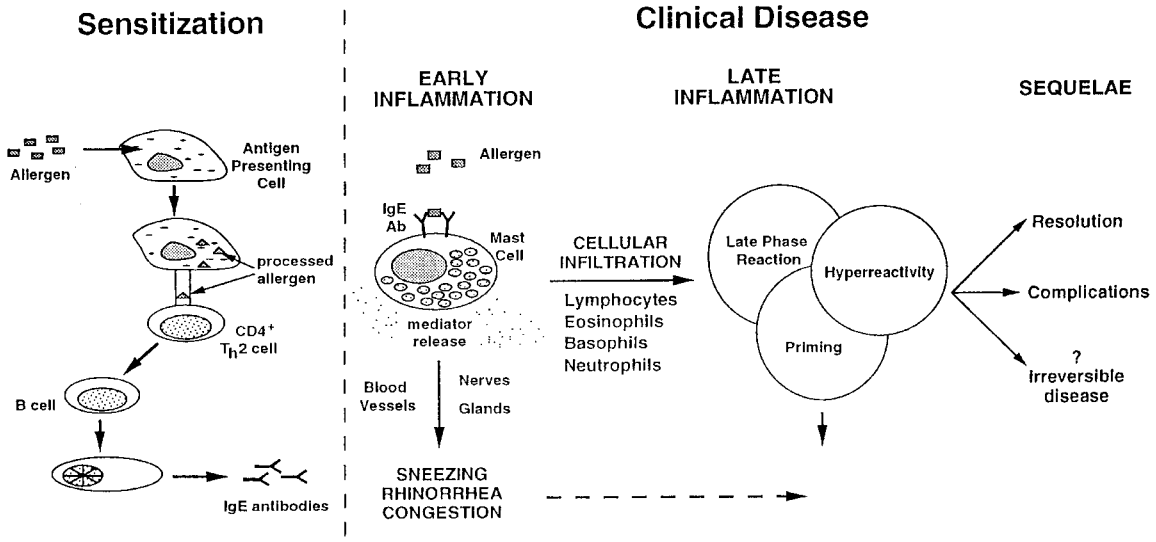


FIGURE 32–3. Pathophysiology of allergic rhinitis. On the left side of the schematic, sensitization to allergen is depicted and involves the production of specific immunoglobulin (Ig) E antibodies to certain allergens. On subsequent exposure to the allergen, depicted on the right, there is cross-linking of specific IgE receptors with resultant degranulation of mast cells and the release of inflammatory mediators. Subsequent to the early reaction, inflammatory cells infiltrate the nasal mucosa, and the late-phase response develops. This is accompanied by a state of increased reactivity that renders the nasal mucosa more responsive to subsequent exposure to allergen (priming) or other stimuli (hyperreactivity). Finally, natural allergic disease might resolve or lead to complications such as sinus infections or otitis media. Reproduced with permission from Naclerio RM. Allergic rhinitis. N Engl J Med 1991;325:860–9.

reported a significant seasonal increase in total MBP^+ and activated eosinophils in the submucosa of allergic patients when compared with preseasonal biopsies or biopsies from nonallergic control subjects.[79]

Adhesion Molecules

Studies have also demonstrated increases in endothelial adhesion molecules, namely, vascular cell adhesion molecule 1 (VCAM-1) in nasal biopsies obtained from allergic subjects 24 hours after nasal allergen challenge.[80] This molecule is expressed on the surface of vascular endothelial cells and interacts with a counterligand, very late activation antigen 4 (VLA-4), which is present on the surface of several leukocytes including lymphocytes, monocytes, eosinophils, and basophils but not neutrophils.[81,82] The VLA-4/VCAM-1 adhesion pathway has been suggested as a mechanism for specific eosinophil, as opposed to neutrophil, migration from the circulation into allergic inflammatory sites.[81,83,84] Other adhesion molecules thought to be important in the recruitment of inflammatory cells from the intravascular compartment into tissue sites of allergic inflammation are intercellular adhesion molecule 1 (ICAM-1), which is constitutively expressed in the nasal mucosa, and E-selectin, which is modestly up-regulated 24 hours after allergen provocation.[80] Intercellular adhesion molecule 1, a ligand for the $ß_2$ integrin molecules leukocyte function-associated antigen-1 and membrane attach complex-1, which are present on the surface of leukocytes, mediates the attachment of all classes of leukocytes to endothelial cells,[82] and E-selectin binds to sialyl Lewis X expressed on the surface of leukocytes.[85] The expression of adhesion molecules has also been studied by Montefort and colleagues, who compared the expression of endothelial cell adhesion molecules in nasal biopsies from subjects with perennial allergic rhinitis and normal control subjects. These investigators found enhanced expression of ICAM-1 and VCAM-1, but not E-selectin, in the mucosa of allergic subjects.[86] Intercellular adhesion molecule expression has been shown to be suppressed by H_1 antihistamines.

Cytokines and Chemokines

In addition to the role of inflammatory mediators in allergic disease, cytokines are increasingly being recognized as important mediators of allergic inflammation and have been shown to have proinflammatory effects in vitro. Such effects include increasing the production of cells in the bone marrow, aiding in the survival of cells in the nasal mucosa, synergizing with other cellular activators, and acting as chemoattractants. Although initially thought to originate exclusively from lymphocytes, other inflammatory cells such as mast cells, basophils, and eosinophils can also produce cytokines.[87,88] Inflammatory cell recruitment enhances the interactions among cells and leads to additional cytokine production, thereby amplifying the reaction, causing further cellular recruitment, and increasing the survival of already recruited inflammatory cells.[89] Some of these inflammatory cells elaborate substances that can cause local damage and mucosal modifications characteristic of chronic rhinitis. The presence of cytokines in nasal secretions and tissues after allergen provocation and natural allergen exposure is beginning to be recognized. Durham and colleagues have demonstrated an increase in the number of cells bearing mRNA for IL-3, IL-4, IL-5, and granulocyte-macrophage colony-stimulating factor (GM-CSF) in nasal mucosal biopsy specimens after allergen challenge.[90] Furthermore, increased levels of IL-1ß and GM-CSF have been measured in nasal secretions during the late-phase reaction, and these elevations peaked at the fifth hour after allergen provocation.[91]

The chemokines are chemoattractant cytokines that play an important role in the function and expression of leukocyte and endothelial adhesion molecules. One such mediator, IL-8, is released by epithelial cells and appears to be necessary for transendothelial migration of neutrophils. Another chemokine, RANTES (*r*egulated on *a*ctivation, *n*ormal *T e*xpressed and *s*ecreted), a member of the CC chemokine subfamily, selectively promotes chemotaxis of memory T lymphocytes, monocytes, and eosinophils. Monocyte chemotactic proteins (MCPs), also members of the CC chemokine subfamily, include MCP-1, MCP-2, MCP-3, and MCP-4. These exert chemotactic activity on monocytes, basophils, lymphocytes, and eosinophils.[92] Another CC chemokine, eotaxin, appears to contribute to the specific recruitment of eosinophils through the eotaxin receptor CCR-3, which is expressed on eosinophils and other leukocytes.[93] Corroborating the importance of these chemokines in allergic reactions are studies demonstrating RANTES, macrophage inflammatory protein 1-α, and IL-8 in nasal lavages after allergen provocation.[94] Interleukin-8 protein has also been localized to the nasal mucosa of both patients with perennial allergic rhinitis and normal controls.[95] Additional studies have shown increased numbers of RANTES, eotaxin, and MCP-3 and MCP-4 mRNA-positive cells in nasal mucosal biopsy specimens 24 hours after allergen challenge.[92] In another study, Kuna and colleagues showed that challenging the nasal cavity with RANTES resulted in the influx of eosinophils, basophils, and lymphocytes into nasal secretions, which were increased significantly compared with those after diluent challenge and more marked in the allergen-challenged nose.[96] Hanazawa and colleagues administered eotaxin intranasally and showed a significant eosinophil influx into nasal secretions, coupled with an increase in nasal nitric oxide 8 hours after challenge.[93] Thus, in vitro, in vivo, and challenge studies all support the importance of chemokines in the pathophysiology of allergic rhinitis, specifically in enhancing cellular recruitment.

Nasal Reactivity

Besides the development of recurrent symptoms hours after a nasal challenge, the inflammatory cellular influx is accompanied by increased nasal reactivity. This has been demonstrated by responsiveness to specific stimuli, such as the antigen that initiated the reaction, and to nonspecific stimuli, such as the secretagogues histamine and methacholine. In the 1960s, John T. Connell[97] coined the term "priming" to refer to increased specific responsiveness. Priming may explain why patients become more symptomatic to the same amounts of pollen in the environment later during the allergy season. Nonspecific reactivity implies increased responsiveness to nonantigenic substances and has been studied by provoking of individuals with methacholine; histamine; cold, dry air; and bradykinin. Like priming, nonspecific reactivity is not obligatorily linked to the appearance of a late reaction. Walden and colleagues found increased sensitivity to histamine 24 hours after antigen challenge.[98] The response to histamine challenge returned to baseline 10 days later, suggesting that increased respon-

siveness to histamine was reversible. Other studies showed a positive correlation between the number of eosinophils 24 hours after antigen challenge and the magnitude of the responsiveness to histamine, and this hyperresponsiveness was inhibited by pretreatment with topical corticosteriods.[99,100]

Majchel and colleagues examined the effect of seasonal exposure on nasal responsiveness to histamine by challenging allergic subjects with histamine before, at the peak of, near the end of, and 2 weeks after the ragweed pollen season.[101] These investigators observed a significant increase in symptoms at the peak of the pollen season that returned to baseline with the disappearance of pollen. The increase above baseline at the peak of the season was not significant for any of the parameters measured. However, they suggested that increased reactivity to histamine with seasonal exposure appears to represent a change in baseline rather than an increased sensitivity to histamine itself. This change in baseline reactivity was inhibited in subjects who were receiving immunotherapy.

Similar studies demonstrate nasal hyperresponsiveness to the cholinomimetic agonist methacholine.[102] Klementsson and colleagues measured the volume of nasal secretions after intranasal administration of 6 mg of methacholine before and after antigen challenge of allergic subjects out of allergy season and observed significant increases in methacholine-induced secretions at 2, 4, 6, 8, 10, and 24 hours after antigen challenge, compared with baseline.[103] Although eosinophils in nasal secretions increased significantly after antigen challenge, no correlation was seen between their numbers and the increase in nonspecific hyperresponsiveness to methacholine. These workers also showed that premedication with two different H_1 antihistamines (terfenadine and cetirizine) resulted in inhibition of both the acute allergic response and the allergen-induced increase in responsiveness to methacholine without affecting eosinophil influx after antigen.[104] This observation was confirmed by Baroody et al for the H_1 antihistamine terfenadine and extended to the antihistamine loratidine.[105] In contrast, corticosteroids, given topically or orally, have dramatic inhibitory effects on cellular infiltration and both specific and nonspecific hyperresponsiveness.[99,100,106] An important consequence of nonspecific hyperactivity is the increased symptoms on exposure to irritants, such as gasoline odors, reported by patients during their allergy season.

Neural Reflexes

Neural reflexes are thought to be involved in the pathophysiology of allergic rhinitis. Sneezing and itching seen during the early response to allergen provocation involve the nervous system. Konno and Togawa and others demonstrated the importance of neural reflexes in patients with allergic rhinitis by showing that stimulating one nasal cavity with histamine led to bilateral nasal secretions.[107] Unilateral intranasal challenge with antigen in subjects with allergic rhinitis led to an increase in sneezes, rhinorrhea, nasal secretions, histamine, nasal airway resistance,[108] and PGD_2[109,110] on the side of challenge. Contralateral to the challenge, rhinorrhea and secretion weights increased significantly, as did PGD_2.[110] The contralateral secretory response was rich in the glandular markers lactoferrin and lysozyme[109] and was inhibited by atropine, an anticholinergic agent, suggesting that the efferent limb was cholinergically mediated.[108] The muscarinic receptors that mediate the actions of acetylcholine in the human nasal mucosa are of both the M_1 and M_3 receptor subtypes and coexist at high densities in submucosal glands.[111]

Immunohistochemical studies have established the presence of several neuropeptides in addition to sympathetic and parasympathetic nerves and their transmitters in the nasal mucosa. These neuropeptides are secreted by unmyelinated nociceptive C fibers (tachykinins, calcitonin gene-related protein [CGRP], neurokinin A [NKA], and gastrin-releasing peptide), parasympathetic nerve endings (vasoactive intestinal protein [VIP], peptide histidine methionine), and sympathetic nerve endings (neuropeptide Y). Substance P (SP), a member of the tachykinin family, is often found as a cotransmitter with NKA and CGRP and has been found in high density in arterial vessels and, to some extent, in veins, gland acini, and the epithelium of the nasal mucosa.[112] In addition to the identification of these neuropeptides in the nasal mucosa, Okamoto et al showed that incubation of nasal biopsy specimens of perennial rhinitics and nonallergic rhinitics with SP or mite allergen resulted in significant increases in mRNA for IL-1β, IL-2, IL-3, IL-4, IL-5, IL-6, tumor necrosis factor-α, and interferon-γ in specimens from allergic subjects but not nonallergic con-

trols.[113] In nasal challenge studies in allergic subjects, Mosimann and colleagues showed that levels of SP, CGRP, and VIP all increased significantly immediately after antigen challenge, with significant but modest increases in SP during the late response.[114] These experiments suggest that neuropeptides are released in vivo in man after allergen challenge and might be partly responsible for symptoms of the allergic reaction. Repetitive application of capsaicin, the essence of chili peppers, depletes sensory nerves of their content of SP and CGRP and initiates both central and axonal reflexes. Unlike its effects in rodents, the capsaicin-induced nasal secretory response in humans is primarily glandular and not attributable to increased vascular permeability.[115] In support of its proinflammatory effects, capsaicin nasal challenge caused significant increases in neutrophils, eosinophils, and mononuclear cells from the prechallenge baseline, with no difference between rhinitic and normal subjects.[116] Furthermore, capsaicin desensitization reduced nasal symptoms recorded for 24 hours after allergen challenge, and this reduction persisted for up to 2 months.[117]

Therefore, several experimental findings highlight the importance of neurogenic control of the allergic response. These include the presence of nasonasal reflexes after nasal antigen provocation, the presence of neuropeptides in nasal tissues and their recovery in nasal secretions after antigen challenge, the ability of these peptides to produce symptoms and inflammatory responses similar to those obtained after exposure to antigen, and the clinical efficacy of capsaicin, which depletes the stores of these substances.

Relationships with Related Organs

Multiple lines of evidence support the observation that allergic rhinitis and sinusitis are closely associated disease entities, although controlled studies of the incidence of rhinosinusitis in patients with allergic rhinitis have not been conducted. Benninger, in a study of outpatients in an otolaryngology practice, found that 54% of those with chronic sinusitis also had allergic rhinitis.[118] Grove and Farrior described a 50% incidence of positive skin tests in patients undergoing sinus surgery, whereas Friedman reported an incidence of atopy in 94% of patients undergoing sphenoethmoidectomies.[119,120] Van Dishoeck and Franssen described allergy as an underlying factor in 40 to 67% of patients with chronic sinusitis.[121] Complementing the above data that allergic rhinitis occurs frequently in patients with chronic sinusitis are data showing a high prevalence of sinus disease in patients with allergic rhinitis. Sinus radiographs are abnormal in more than 50% of adults and children with perennial allergic rhinitis.[122,123] In contrast, Enberg reported a low incidence of sinus abnormalities in patients with perennial nonallergic rhinitis.[124] Magnetic resonance imaging scans demonstrate increased evidence of sinus mucosal abnormalities during major pollinating seasons.[125] A study of subjects with ragweed rhinitis during the pollen season showed that 60% of the subjects had sinus mucosal abnormalities on computed tomography (CT).[126] Multiple mechanisms are postulated to explain these relationships. In one recent study of subjects with seasonal allergic rhinitis outside their season, the nose was challenged with allergen, and nasal and ipsilateral maxillary sinus responses were monitored by use of lavage. There was a late increase in total cell count (the percentage of eosinophils and total eosinophils within the sinus cavity in the allergen-challenged subjects) but not in the control experiments. There were significantly more total cells after allergen challenge compared with the control group as well as a higher number of total eosinophils.[127] This study highlights the fact that important relationships exist between the nose and sinus in the response to allergen, possibly through regional or central neural reflexes.

There are important relationships between the nose and other organs in the pathophysiology of allergic rhinitis. A relationship between pulmonary disease such as asthma and allergic rhinitis has long been observed. Allergic rhinitis and asthma often coexist. Allergic rhinitis can lessen the severity of asthma,[128] and treatment of allergic rhinitis by intranasal corticosteroids can improve the symptoms of asthma and reduce bronchial hyperresponsiveness.[129] Treatment with orally inhaled corticosteroids not only prevented development of increased bronchial hyperresponsiveness to methacholine but also reduced nasal symptoms, eosinophils from nasal brushings, markers of eosinophil activation in nasal lavage fluid, and peripheral blood eosinophilia.[130]

A recent study by Braunstahl and colleagues demonstrated that segmental provocation with allergen in the lungs produced allergic inflammation in the nose in atopic patients.[131] Another study

examined the ability of the nose to condition cold, dry air. Assanasen et al demonstrated that asthmatics have a reduced ability to warm and humidify cold, dry air as compared with normal subjects; subjects with seasonal allergic rhinitis also had a reduced ability to condition air, with possibly adverse implications for the lower airways.[132] Pathophysiologic connections between the nose and lungs are still not entirely understood and remain an active topic of investigation. The above data, however, support the presence of important physiologic and pathophysiologic connections between contiguous areas of the respiratory tract.

Nasal inflammation with an allergic or infectious cause may be a factor involved in the development of otitis media. Studies of the pathogenesis of otitis media have identified interactions among infection, allergic reactions, and eustachian tube dysfunction.[133] In the presence of allergic rhinitis, treatment may improve symptom resolution and therapeutic response. The inflammation found in allergic rhinitis is thought to promote eustachian tube dysfunction and support the development of otitis through mediators and cytokines. This remains an area of ongoing investigation.

Summary of Pathophysiology

The pathophysiology of allergic rhinitis can be summarized as follows: After sensitization of the nasal mucosa to a certain allergen, subsequent exposure leads to cross-linking of specific IgE receptors on mast cells and their resultant degranulation, with the release of a host of inflammatory mediators that are responsible for allergic nasal symptoms (see Figure 32–3). Proinflammatory substances produced by other inflammatory cells are also generated after antigen exposure, most prominent among which are eosinophil products and cytokines. Cytokines are thought to be generated, in part, by lymphocytes, which are found in abundance in both resting and stimulated nasal mucosa. Recent evidence also points to an important role for mast cells in the storage and probable production and secretion of cytokines. Cytokines can up-regulate adhesion molecules on the vascular endothelium and possibly marginating leukocytes, leading to the migration of these cells into tissues. Other cytokines also promote chemotaxis and survival of recruited inflammatory cells. Another important player is the nervous system, which amplifies the allergic reaction by both central and peripheral reflexes that result in changes at sites distant from those of antigen deposition. All of these changes lower the threshold of mucosal responsiveness and amplify it to a variety of specific and nonspecific stimuli, making allergic individuals more responsive than nonallergic individuals to stimuli to which they are exposed in daily life (see Figure 32–3).

ALLERGENS

Allergens are foreign substances capable of provoking an IgE-mediated response. Most allergens are between 5 and 20 μm in diameter, a size that permits their complete removal by the nose. They are proteins with molecular weights between 10 and 40 kD. No distinguishing surface characteristics appear to differentiate allergens from nonantigenic substances.

Allergens are often categorized into indoor and outdoor types. In general, outdoor allergens are responsible for seasonal allergic rhinitis, whereas indoor allergens usually cause perennial rhinitis. Pollens causing allergy in temperate climates are released into the air from plants, trees, weeds, and grasses and are carried over great distances. Thus, cutting down trees around a suburban home in an effort to reduce the amount of pollen has little effect. About 75% of patients with seasonal allergens have symptoms caused by ragweed, 40% by grasses, and 5% by trees alone. Approximately 25% have allergies to both grass and ragweed, and 5% have allergies to all three pollens.[27] Trees clearly have geographic variations; for example, Western red cedar is limited to the Northwest. Grasses are diverse and include timothy grass, often used for feeding horses; Kentucky bluegrass, widely used in lawn grass mixtures; orchard grass; rye grass; and English plantain. Common short ragweed is found throughout North America, with the exception of Newfoundland, and is conspicuously absent from the European continent.

Pollination, and hence the allergy season, occurs in a predictable annual pattern for different regions of the country.[134] The pattern, however, varies throughout the country. In the Northeast, trees pollinate in mid-March to late April, grasses follow in May and June, and ragweed flowers from mid-August until the first frost. In the South, tree blooming begins in early February. In contrast to the sharply demarcated grass season that occurs in the North, in the South, grasses may pollinate from

March through September, and in some areas, pollination may be a year-round process. The pattern in the central United States resembles the patterns seen on the east coast. In the California lowlands, grass pollen is present from early March through November, and trees and short ragweed are present as in other regions. In the Northwest coastal region, trees and grass pollen are present, but the region is ragweed free. In the traditionally arid Southwest, previously a haven for allergy sufferers, increased urbanization and irrigation have contributed to increasing the pollen load. Humans tend to bring their allergens with them.

The most frequent perennial allergens are animal danders, dust mites, cockroaches, and molds. Dust mites are microscopic, eight-legged organisms of the genus *Dermatophagoides*, including *D. pteronyssinus*, *D. farinae,* and *Euroglyphus maynei*. They are the major allergens in "house dust." Dust mites are found throughout the world, with the exception of extremely dry climates such as northern Sweden, central Canada, and areas at elevations above 10,000 feet. These mites feed on human epithelial scales and thrive in warm, humid environments (60 to 70% relative humidity, temperature 65 to 80°F). Bedding provides an ideal environment for proliferation of dust mites. Other sites for mite accumulation are upholstered furniture, carpets, and stuffed toys. Dust mite feces, the source of the allergen, are relatively large particles that remain airborne for short periods, unlike outdoor pollen. When an individual sits on a bed, the particles become airborne and are inhaled. Because these particles are large, they settle from the air rapidly, and air filtration systems cannot effectively remove them. Lowering the indoor relative humidity to less than 50% during the summer months has a profound effect on the mite population and the antigen load throughout the year,[135] suggesting a role for dehumidifiers even in homes with central air conditioning.

Animal danders are an important source of indoor allergens. Cat and dog danders are the most frequent, but mice, guinea pigs, and horses can all be responsible for allergic symptoms. Laboratory workers can become allergic to animals at work. In most cases, the allergen is found in secretions. In cats, *Fel d I* is the principal allergen secreted in cat saliva. It dries on fur and is spread to furniture, bedding, and carpets. When these reservoirs are disturbed, the allergen becomes airborne and can provoke symptoms. Cat dander is "sticky," and children with cats can carry enough to school to cause symptoms in cat-allergic children who have no cats in their homes.[136]

The cockroach is an important source of allergen in inner-city populations. Both the American (*Periplaneta americana*) and German cockroach (*Blattella germanica*) have been identified as important allergens in asthma. Allergenicity occurs to body parts and to feces. Molds, although less well studied, are sources of allergens, particularly in warm, humid environments. They tend to be found inside older homes in areas of decreased ventilation or increased dampness. Although a phenomenal variety of molds exists, *Alternaria* and *Cladosporium* are principally responsible for symptoms owing to outdoor exposure, and *Aspergillus* and *Penicillium* are most prevalent indoors.

The patient's work environment may also be a source of allergens. Symptoms occurring only at work and subsiding on weekends may reflect an occupational disorder. At risk are flour handlers, workers in paint and plastic industries, woodworkers, fish and shellfish processors, and animal handlers. Unfortunately, few specific tests exist for the diagnosis of these disorders.

Allergic reactions to natural rubber latex have increased, especially in health care workers who have high exposure by direct skin contact and inhalation of latex particles from powdered gloves. This problem is being addressed by studies on improving diagnostic methods for latex allergy and evaluating strategies for prevention.

CLINICAL PRESENTATION

History

Antigen exposure causes itching within seconds, which is soon followed by sneezing. Rhinorrhea ensues, and within about 15 minutes, nasal congestion peaks. Besides nasal symptoms, patients often complain about ocular pruritus, tearing, pharyngeal itching, throat clearing, cough, and ear popping. Other commonly reported symptoms include postnasal drip, increased lacrimation, dry cough, red eyes, headaches (pressure) over the paranasal sinus areas, and loss of smell or taste. These symptoms are nonspecific and have significant clinical overlap with other disorders. Itching of mucous membranes and repetitive sneezing, however, are the symptoms most suggestive of

allergic disease. The relative importance of each symptom may vary among individuals, but nasal congestion tends to be the most bothersome, although each symptom is usually present, at least to some degree.

When obtaining the history, the physician should attempt to link exposure to allergens temporally with the occurrence of symptoms. This temporal correlation is the hallmark of allergic rhinitis. Patients with seasonal allergies complain of recurrent symptoms only at specific times of each year that coincide with pollination periods. In contrast, a history of year-round symptoms may indicate sensitivity to a perennial allergen or multiple seasonal allergens. Symptoms immediately following exposure to a potential source of allergen, such as a cat, strongly suggest an allergy to that source. Exposures to perennial allergens tend to be accentuated in winter in colder climates, when ventilation is reduced. Symptoms occurring only at work or during the workweek and subsiding on weekends may reflect an occupationally related disorder. The presence of domestic pets (including birds) and whether these sleep in the patient's room must be determined.

Additional considerations in history taking include the response to prior therapy and evidence of complications. Related effects of the pathophysiology must be addressed. For example, nasal obstruction may lead to mouth breathing. In children, this may be manifested as adenoid facies, with a high palatal arch and abnormal dental development. In adults, nasal obstruction may contribute to snoring and sleep-disordered breathing. Obstruction of sinus ostia may predispose to sinusitis. Eustachian tube dysfunction may occur,[137] but this has not been shown to lead to an increased incidence of otitis media with effusion.[138,139]

A general medical history remains important. The medical history may document systemic disorders that affect the nose, such as hypothyroidism. Pregnancy can produce nasal congestion and may require modification of treatment strategies. The presence of pulmonary disease such as asthma should be sought. Indeed, a significant percentage of patients with allergic rhinitis have asthma. Between 5 and 10% of asthmatic subjects may have intolerance to aspirin and nonsteroidal anti-inflammatory drugs. A family history of allergic rhinitis increases the chances of the patient's having an allergic disorder. Nasal symptoms might also be caused by intake of medications such as beta blockers, which may contribute to nasal congestion through interference with the adrenergic mechanism. Tricyclic antidepressants may produce dryness of the nasal mucosa by virtue of their anticholinergic effects. Angiotensin-converting enzyme inhibitors can produce a chronic cough. Birth control pills can cause nasal congestion, and topical eyedrops can also induce nasal symptoms.

Examination

Attentive history taking and physical examination, combined with appropriate diagnostic tests, are required for establishing the correct diagnosis because allergic rhinitis shares features of other nasal disease entities. The classic description of allergic facies includes mouth breathing, allergic "shiners" (resulting from periorbital venous stasis from chronic nasal obstruction), and a transverse supratip nasal crease from long-term rubbing of the nose upward to relieve itching. These classic presentations occur especially often in children, but absence of these signs does not exclude the disease.

Physical examination must be complete. Ocular examination may demonstrate injection of the conjunctiva or swelling of the eyelids. Examination of the nose begins with observing the external appearance for gross deformities such as a deviation suggesting previous trauma or expansion of the nasal bridge suggestive of nasal polyps. A nasal speculum permits evaluation of the anterior third of the internal nasal architecture and the character of the nasal mucosa. Structural anomalies providing an anatomic basis for obstruction or recurrent infections such as septal deviations or spurs should be sought. The character and consistency of nasal secretions should be noted. These can vary from thin and clear to thick and whitish. The nasal mucosa may be swollen and pale-bluish, although these signs are not pathognomonic of the disease, as previously thought. The examination of allergic individuals often appears normal, and the primary importance of the physical examination is to rule out other causes of or contributors to the symptoms.

Decongestion of swollen nasal mucosa with a topical decongestant improves visualization and allows the differentiation of reversible from irreversible changes. Combining the vasoconstrictor with a topical anesthetic allows complete examination with an endoscope. The choanae and the nasopharynx can be visualized in this manner. The

region of the middle meatus should also be examined carefully because secretions there might be suggestive of acute or chronic sinusitis. Nasal polyps that were not visualized by anterior rhinoscopy may be seen during a careful endoscopic examination. Nasal polyps are infrequent in allergic rhinitis (< 2%) but are found in up to 20% of patients with cystic fibrosis.[42] The presence of nasal polyps in children suggests the diagnosis of cystic fibrosis because polyps are rarely found in this age group. These children should undergo sweat or genetic testing.

DIAGNOSIS

The identification of allergen(s) responsible for the patient's symptoms is important both for establishing the diagnosis and for the institution of avoidance measures. Symptoms occurring in temporal relation to allergen exposure suggest sensitization but are not diagnostic. Sensitization implies the presence of elevated levels of IgE directed against a specific allergen and can be demonstrated by a wheal and flare response to skin testing with allergen extracts or by measuring the level of antigen-specific IgE antibodies in the serum. However, individuals can show evidence of sensitization by a positive skin test or elevated specific antibody levels in the serum without having evidence of clinical disease. This emphasizes the importance of obtaining a good history in the evaluation of patients with suspected allergic disorders. In patients with a positive history, the magnitude of skin responses often corresponds to the severity of symptoms.

Skin Testing

Skin testing furnishes an excellent in vivo method to demonstrate sensitivity to a given allergen. This test evaluates the presence of specific IgE antibodies on skin mast cells, the reactivity of these cells, and the reaction of the end-organ to released mediators. Its advantages include greater sensitivity, the rapidity with which results can be obtained, and low cost. Like all diagnostic tests, skin testing also has disadvantages, which include the inability to perform the test in patients with dermatologic problems such as dermatographism and extensive eczema, poor tolerance of many children for multiple needle pricks, the inhibitory effect of certain ingested drugs such as antihistamines on skin test reactivity,[140,141] the need to maintain the potency of the allergen extracts, and the possibility of systemic reactions.

Skin testing often begins with puncture testing, which provides low-dose allergen exposure. A small drop of concentrated allergen is placed on the skin (usually on the volar surface of the forearm or the back), and a minute quantity is introduced into the dermis with a sharp object. Positive responses occur within 10 to 15 minutes and produce a characteristic raised central area of induration (wheal), with a surrounding zone of erythema (flare). The response is graded in comparison with a positive histamine or codeine response, and a negative control with the diluent for the allergen extracts is also included to control for nonspecific reactivity to the vehicle. The positive control ensures that the patient can mount a cutaneous reaction to histamine, and the absence of a reaction can unmask interference by medications, decreased skin reactivity, or technical problems with the procedure. Skin testing is valid in infants and young children, but the criteria for a positive reaction need to be adjusted because the reactions are smaller.[142] Measurement of serum-specific IgE levels is also valid in this younger age group.[143] It is important, however, to test with only relevant antigens. Infants are more likely to be allergic to foods, and children are more likely to be allergic to perennial rather than seasonal allergens.

Negative puncture tests are usually confirmed by intradermal tests, which are more sensitive. In an intradermal test, a small (0.01 to 0.05 mL) quantity of dilute allergen is injected into the superficial dermis, and the same wheal and flare responses are observed and graded in comparison with a positive histamine or codeine control. Because antihistamines can interfere with the results of skin testing, most H_1 receptor antagonists are withheld for 2 days before skin testing. Tricyclic antidepressants suppress responses for several weeks, as can tranquilizers and antiemetics of the phenothiazine class through intrinsic anti-H_1 activity. Short-term oral corticosteroid treatment has no effect on skin test reactivity but may have an inhibitory effect if the agent is taken for long periods. Testing with extracts that are standardized (ragweed, grass pollens) is more reliable than for nonstandardized antigens such as foods and chemicals.

In Vitro Immunoglobulin E Measurements

Drawing blood for the measurement of specific IgE can circumvent some of the disadvantages of skin

testing. False-positive results may occur if patients have elevated IgE levels in their sera because of nonspecific binding. Therefore, although IgE levels alone are of limited usefulness in the diagnosis of allergic rhinitis, and because elevated levels can also exist in patients with nonallergic conditions, these levels must be obtained in conjunction with a determination of specific IgE levels. False-negative results may also occur from inhibition by IgG antibodies with similar affinities as in patients receiving immunotherapy. Data from clinical studies comparing the results of skin testing and in vitro tests for specific IgE determination (the radioallergosorbent assay or RAST) in allergic subjects suggest a good correlation between the two,[144] with a higher sensitivity for skin testing.[145] Therefore, both determinations of specific IgE levels and skin testing are useful in the diagnosis of allergic disorders, but their results should always be interpreted in the context of clinical symptoms. The disadvantages of this form of allergy testing include cost, slightly lower sensitivity, and the time delay between drawing the blood and obtaining the results.

Other Diagnostic Tests

Peripheral eosinophilia, although nonspecific, may indicate the presence of other atopic diseases. Nasal cytologic examination allows the identification of eosinophils and other inflammatory cells in nasal secretions and may be helpful in differentiating an infectious from an allergic cause during a clinical exacerbation of symptoms. In normal individuals, smears show the presence of epithelial cells, including some ciliated and goblet cells, with few eosinophils, neutrophils, basophils, or bacteria. In subjects with an infection, neutrophils increase in nasal secretions, and in symptomatic allergic subjects, the percentage of eosinophils increases. A value greater than 10% is suggestive of allergic disease.[146–150] Eosinophils, however, may also be present in the absence of IgE-mediated disease. Approximately 25% of patients with chronic rhinitis and negative skin tests demonstrate eosinophilia on nasal cytologic study, and this entity is known as the nonallergic rhinitis with eosinophilia syndrome (NARES). Regardless of the cause, the presence of eosinophilia usually implies a favorable clinical response to corticosteroid therapy.[151]

Soft tissue radiography of the neck can evaluate adenoid size, a major consideration in the differential diagnosis of rhinitis in children, especially when the predominant symptom is nasal obstruction. Sinus disease often complicates perennial allergic rhinitis and may be a consideration in the differential diagnosis. In the past, sinus plain films were used, but CT is now standard for evaluation of the presence and extent of sinus abnormalities. The common association of upper and lower airway disease makes tests of pulmonary functions useful adjuvants. This statement applies to such diverse disorders as cystic fibrosis, asthma, and bronchopulmonary aspergillosis.

THERAPY

Environmental Modifications

The most potent treatments cannot eliminate symptoms in the face of an overwhelming allergen load. Complete avoidance of the allergen(s) to which the patient is sensitive eliminates symptoms of the disease. Thus, measures aimed at reducing the allergen load from the patient's environment are effective in reducing symptoms. Although many methods of environmental control (removal of a pet, restriction of activities, home renovations) are difficult to execute, particularly with children, simple measures can be very effective. In seasonal allergic rhinitis, patients can reduce exposure by keeping windows closed on days when pollen counts are high and limiting physical activity outdoors in the early morning and evening when pollen counts peak. Air conditioning can help, and the addition of special filters can prevent pollen grains and mold spores from entering the home. Pollen counts are often given during daily weather reports, and the previous day's pollen counts are the best predictor of the next day's counts. Rain, however, profoundly reduces the levels of outdoor pollens.

Measures to reduce exposure to dust mites concentrate on bedding.[152] These include replacing feather pillows and bedspreads with synthetic ones that can be washed in hot water (hotter than 130°F) and covering mattresses with commercially available plastic covers ($< 10\ \mu m$). Removing carpets and frequent vacuuming also help. Where carpets cannot be removed, acaricidal products can directly kill mites. Stuffed toys can be placed in the freezer for 2 days to reduce the number of dust mites.

In subjects with allergies to animal dander, removal of a domestic pet is the best step. Reduction in the allergen load after pet removal may take up to

6 months, however; thus, symptoms may persist. Furthermore, complete removal of a pet may be difficult, but eliminating the animal from the bedroom, where we spend an average of 8 hours per day, is a helpful alternative. In the case of cats, regular washing of the animal may help decrease the allergen load.[153]

Humidifiers must be cleaned regularly to avoid becoming a source of mold allergens, which thrive in moist environments. The humidity should not exceed 40 to 45% because higher levels encourage the growth of dust mites and molds. The use of high-efficiency particulate air (HEPA) filters and of electrostatic filters effectively removes particulates larger than 1 μm in diameter. Particles, however, must be airborne for removal, such as pollens, whereas heavy particles that settle from the air rapidly, such as dust mites, are not eliminated by these measures.[154] High-efficiency particulate air filters are preferred to electrostatic ones as they filter out particulate matter without generating ozone, an irritant.

General household changes may also be helpful. Repair of any plumbing or drainage problem areas in the house must be undertaken to eliminate mold. Decontamination of surfaces contaminated with mold by use of diluted bleach and elimination of soiled organic materials is useful. Appropriate packaging of garbage and pesticide application are necessary to eliminate cockroaches. Use of well-fitting, semirigid masks when vacuuming offers useful protection from airborne allergens such as dust mites.[155]

PHARMACOLOGIC THERAPY

Antihistamines

Three receptors exist for histamine. H_1 receptors are found on blood vessels, on sensory nerves, on smooth muscles of the respiratory and digestive tracts, and in the central nervous system. Stimulation leads to vasodilatation, increased vascular permeability, sneezing, pruritus, glandular secretion, and increased intestinal motility. H_2 receptors have a distribution similar to that of H_1 receptors but are principally involved in the regulation of gastric acid secretion. H_3 receptors are located principally in the brain and seem to be involved in the regulation of histamine synthesis and release. The contribution of histamine to the early allergic response, largely mediated by the H_1 receptor, has long been recognized and is the rationale for the large number of H_1 antagonists in clinical use.

H_1 antihistamines have been classified as first-generation, or sedating, and second-generation, or nonsedating, antihistamines. The first generation of antihistamines is effective, but they have some undesirable side effects because of their lack of selectivity and subsequent nonspecific stimulation of other receptors. Among these side effects are sedation, anticholinergic effects, functional and performance impairment, and gastrointestinal distress. When applied topically, some act as local anesthetics. The most important of the side effects is sedation, which is reported in approximately 20% of patients.[156] Objective studies of performance have documented the problem. The correlation with the subjective reporting of sedation was weak, however, and the problem appeared more significant than previously believed; that is, a larger number of subjects developed impaired performance compared with the number reporting sedation. Furthermore, impaired performance can persist after the subjective feeling of somnolence dissipates—hence, the importance of warning patients who are taking these medications not to perform some tasks such as operating heavy machinery or driving. Although some side effects may be useful, such as the use of diphenhydramine for treating insomnia, they are usually a nuisance and limit compliance. In one study, diphenhydramine caused worse impairment to driving than consuming alcohol to the level of being legally drunk.[157] In this study, performance after a second-generation agent was similar to that after a placebo.

Second-generation antihistamines are less lipophilic than first-generation H_1 antihistamines and do not penetrate the blood-brain barrier. Therefore, they produce no more somnolence than does placebo. Their greater receptor selectivity also reduces the incidence of anticholinergic side effects. In addition to antagonizing histamine at the H_1 receptor, some antihistamines, such as azatadine, of the piperidine class, and terfenadine, a nonsedating antihistamine, inhibit histamine release after intranasal antigen challenge.[158,159] Treatment with some antihistamines also reduces the production of leukotrienes and kinins, which are mediators with proinflammatory effects,[159,160] as well as the allergen-induced increased responsiveness to methacholine.[104] Another anti-inflammatory property of antihistamines is a reduction of soluble ICAM-1 lev-

els in nasal secretions, a property demonstrated by both loratidine and cetirizine.[161]

Oral antihistamines are readily absorbed. Their onset of action is rapid, usually within 60 minutes, and maximum benefit occurs within hours.[160] Metabolism of most antihistamines occurs primarily through the hepatic cytochrome P-450 system. Drugs that interfere with this system, such as antifungal agents, can lead to the accumulation of antihistamines to toxic levels. One exception is cetirizine, which is primarily excreted in the urine and does not depend on the cytochrome P-450 system. The clinical effectiveness of antihistamines exceeds the duration of measurable serum levels. This phenomenon may be attributable to the presence of active metabolites. Another explanation for the prolonged efficacy of H_1 receptor antagonists beyond their measurable serum levels relates to extended tissue levels.

In the United States, four common second-generation antihistamines are in clinical use: azelastine, cetirizine, fexofenadine, and loratidine.[162] Azelastine, a phthalazinone derivative, is an intranasal preparation with efficacy comparable to that of other antihistamines. Although it does not cause somnolence, azelastine may cause a sensation of altered taste immediately after use.[163] Cardiac arrhythmias, fatal ventricular tachycardia, or prolongation of the QT interval associated with concomitant administration of agents that interfere with the cytochrome P-450 system have been reported with two older second-generation antihistamines, terfenadine and astemizole. These agents are no longer available in the United States. Fexofenadine, a metabolite of terfenadine, does not possess these cardiac risks. Second-generation agents have been shown to be effective. Both cetirizine and fexofenadine have been demonstrated to improve quality of life, as measured by generic and disease-specific tools.[164,165] They cause little or no somnolence, do not affect performance, and have no anticholinergic effects.[166] Azelastine and fexofenadine are approved for children older than 12 years, loratidine may be used in children age 6 or older, and cetirizine may be used in children as young as 2 years.[167]

All antihistamines are effective in the treatment of allergic rhinitis and differ principally in their side effects, duration of action, and cost. In equipotent doses, they are equally effective in suppressing histamine-induced skin wheals.[168] H_1 receptor antagonists are most effective in treating sneezing, nasal and ocular pruritus, and rhinorrhea associated with allergic rhinitis but have little or no effect on nasal congestion. Thus, they are often combined with an oral decongestant. Generic first-generation H_1 blockers are considerably less expensive than their nonsedating counterparts. Some clinicians have tried to circumvent the sedation caused by first-generation antihistamines by directing that they be taken before bedtime, when somnolence is not a problem. The next day, prolonged tissue levels provide continued efficacy without the undesirable side effect of drowsiness.[169] Some studies, however, suggest that performance is impaired despite the lack of drowsiness. It is therefore important to warn patients receiving these drugs about their effect on daily activities, such as driving or operating heavy machinery. Patients who lack medical insurance or formulary coverage often use first-generation agents despite their significant side effects. This poses considerable public health concerns, given recent studies demonstrating the marked performance impairment associated with their use. Unfortunately, the nonsedating antihistamines are more expensive than the traditional antihistamines.

Decongestants

Decongestants exert their effect through stimulation of α_1- or α_2-adrenergic receptors. These receptors are present on resistance vessels, where they control blood flow, and on capacitance vessels, where they control blood volume. In capacitance vessels, α_2 receptors outnumber α_1 receptors.[170] In resting conditions, sympathetic nervous activity regulates nasal patency by maintaining the sinusoids contracted to approximately half maximal capacity. The resting state is affected by the nasal cycle, a periodic, reciprocal alteration of nasal cavity congestion and decongestion that affects about 80% of normal individuals. Increased sympathetic stimulation, such as occurs during exercise, reduces nasal congestion.

Oral decongestants exert their effects directly and by stimulating norepinephrine release. The two major decongestants are pseudoephedrine and phenylpropanolamine, which can be prescribed separately or in combination with antihistamines. Use of phenylpropanolamine in diet medication has been associated with rare hemorrhagic stroke and has been withdrawn from the US market. Because oral decongestants also stimulate adrenergic recep-

tors other than those in the nasal vasculature, overdosage has been associated with hypertensive crisis. When given in prescribed doses, however, they do not induce hypertension in normotensive patients, nor do they alter the pharmacologic control of stable hypertensive patients. Current recommendations suggest that decongestants should not be used in patients with uncontrolled hypertension, in those with severe coronary artery disease, or in patients receiving monoamine oxidase inhibitors. Decongestants should be prescribed with caution in patients with diabetes, hyperthyroidism, closed-angle glaucoma, coronary artery disease, cardiac insufficiency, prostatic hypertrophy, or urinary retention.[171] Their major side effect is insomnia, which occurs in approximately 25% of patients.

Topical decongestants are effective in reducing nasal congestion, regardless of the cause, and these include catecholamines (such as phenylephrine) and imidazoline derivatives (such as xylometazoline or oxymetazoline). Prolonged use can bring about rhinitis medicamentosa, which is characterized by reduced duration of action and rebound nasal congestion after cessation of therapy. Because this phenomenon can appear even after a short period, use of these agents should be limited to a few days. These agents are best reserved for patients in whom nasal congestion is so severe that it precludes the use of other topical preparations such as intranasal corticosteroids or more restful sleep is required during acute exacerbations of disease. Seizures have occurred in children given these medications intranasally.

Topical Intranasal Corticosteroids

Topical intranasal glucocorticosteroids are potent medications for the treatment of allergic rhinitis. These agents profoundly reduce multiple aspects of the inflammatory response to allergen. Corticosteroids penetrate the interior of the cell, where they are bound by a glucocorticoid receptor in the cytoplasm. The glucocorticoid-receptor complex then penetrates the nucleus, where it inhibits the synthesis of the proinflammatory cytokines IL-1, -2, -3, -5, and -6; interferon-γ; tumor necrosis factor-α; and GM-CSF[172] and induces the synthesis of other anti-inflammatory substances such as vasocortin and lipocortin. These agents reduce eosinophil survival and function induced by IL-1, -3, and -5.[173,174] Treatment with intranasal flunisolide resulted in significant inhibition of mediator release during both early- and late-phase reactions after antigen challenge, along with a significant inhibition of the influx of basophils, eosinophils, neutrophils, and mononuclear cells in nasal secretions.[174,175] Treatment with intranasal corticosteroids also reduced the antigen-induced hyperresponsiveness of the nasal mucosa to subsequent antigen[174] and histamine provocation.[100] Intranasal corticosteroid administration also leads to a reduction in inflammatory cells and TH2-type cytokines within the nasal mucosa.[92] For example, treatment with fluticasone inhibits the allergen-induced increase in IL-4, IL-13, eotaxin, and MCP-4 mRNA. In addition, intranasal corticosteroids increase the level of TH1-type cytokines such as interferon-γ and IL-12, which can suppress the transcription of IL-4.[92] These beneficial anti-inflammatory effects of intranasal corticosteroids support the findings in several placebo-controlled clinical trials of a variety of these agents that demonstrated their efficacy in the reduction of nasal symptoms in both seasonal[176] and perennial[177] allergic rhinitis. Corticosteroid-induced cellular modifications require several hours to appear, and this might explain why their onset of action was gauged in days, although more recent studies show that fluticasone affects symptoms within 12 hours.

Topical corticosteroids effectively suppress the response to allergen provocation. In contrast to systemic corticosteroids, pretreatment with topical corticosteroids reduces the acute nasal response to allergen challenge, as shown by a reduction in symptoms and in levels of recovered inflammatory mediators in nasal secretions (Figure 32–4).[174] Treatment with topical corticosteroids also reduces symptoms, the levels of mediators, and cellular infiltration during the late-phase reaction to allergen challenge and the priming response to antigen (see Figure 32–4).[70,174] These agents also inhibit hyperresponsiveness to nonantigenic stimuli such as histamine.[99] Topical corticosteroids also prevent the increase in mast cells and inflammatory cells seen during seasonal exposure to allergen.[147] Furthermore, they also resulted in suppression of the seasonal increase in specific IgE antibodies during the ragweed season.[178] A direct vasoconstrictor effect of topical glucocorticosteroids, found in the skin, does not occur on the nasal mucosa.[179]

Currently, topical forms of corticosteroids include flunisolide, beclomethasone dipropionate, triamcinolone, budesonide, mometasone, and fluti-

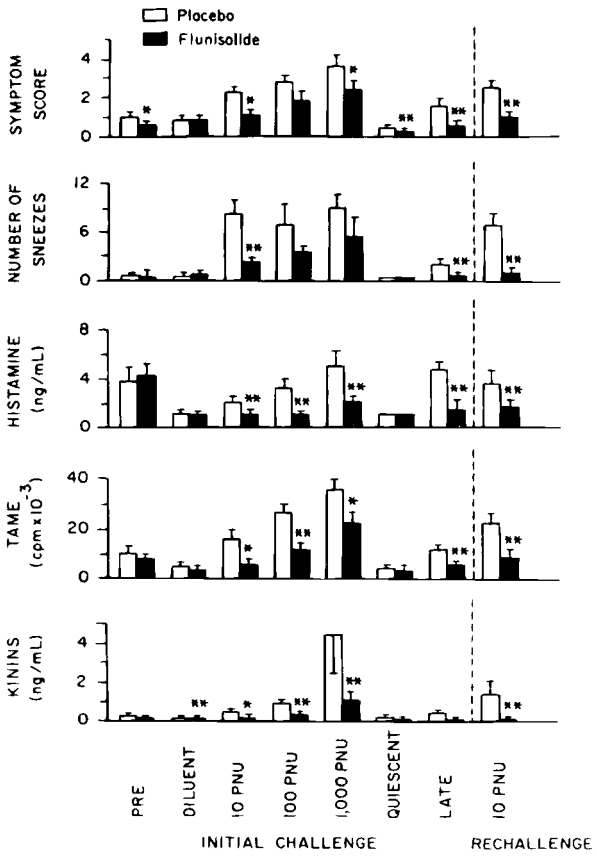


FIGURE 32–4. Effect of intranasal corticosteroids (flunisolide) on the nasal response to allergen. The protocol of challenge is seen on the abscissa: Pre = prewash at the initiation of the allergen challenge protocol; Diluent = challenge with the diluent for the allergen extract to control for nonspecific reactivity; 10, 100, and 1,000 protein nitrogen units (PNU) = increasing doses of allergen used for challenge; Quiescent = the initial hours after allergen provocation; Late = the late-phase response; 10 PNU = rechallenge with the lowest dose of allergen 10 hours after the initial challenge. As can be seen from the response with subjects on placebo (*open bars*), there was a significant increase over diluent in all of the parameters measured during both the early and late responses. There was also an increase in responsiveness to the lowest dose of allergen (10 PNU) as assessed by levels of histamine, tosyl-L-arginine methyl ester (TAME)-esterase, and kinins in nasal lavages, indicative of the priming response. Pretreatment with flunisolide (*closed bars*) resulted in inhibition of the early, late, and rechallenge responses to allergen. Adapted from Pipkorn U et al.[174]

casone. Although differences in the potency and strength of receptor binding among these molecules can be demonstrated in vitro and in certain in vivo models, none of these variations has been shown to translate into major clinical differences. These drugs are principally distinguished by the form of administration (pressurized aerosol, metered-dose inhaler, powder), by the frequency of administration (either once or twice daily), and by the potential to cause systemic toxicity. Their onset of action has been reported to be as short as 1 day, with most preparations having a noticeable clinical effect by 3 days and a peak effect by 2 weeks. Previous studies suggested that these medications work best with continued use as opposed to intermittent, as needed use.[180] However, there is evidence that they may be effective when used intermittently.[181] A study compared the efficacy of intranasal fluticasone and placebo nasal sprays used as needed in patients with seasonal allergic rhinitis and showed that as needed use of the intranasal corticosteroid led to significant improvement in symptoms and quality-of-life scores compared to placebo.[182] Additionally, data indicate that the as needed use of intranasal corticosteroids is more effective for perennial allergic rhinitis than is a second-generation antihistamine.[183]

Side effects are relatively rare.[184] The most frequent is nasal irritation, which occurs in approximately 10% of patients. This is manifested as a nasal burning sensation or sneezing. Two percent of patients have blood-tinged secretions either because of the medication or because of the delivery system.[183] Although septal perforations have been reported, they are extremely rare.[185] Nasal biopsies after prolonged use of these agents have not shown thinning of the nasal epithelium or abnormalities in the nasal mucosa.[186–188] Mucosal superinfection with *Candida albicans*, occasionally found with the use of topical, orally inhaled corticosteroids in the treatment of asthma, has not been a significant problem in the nose.[189]

Systemic side effects have long been a matter of concern. Dexamethasone, the first available topical corticosteroid in the United States, had measurable systemic absorption, leading to adrenal suppression after prolonged use.[190] The newer preparations now available on the market have lower systemic absorption, and at the standard doses used for the treatment of allergic rhinitis, no detectable effects on the hypothalamic-pituitary-adrenal axis have been found. One report of the development of retrocapsular cataract during beclomethasone therapy appeared,

but some of these patients had also received systemic corticosteroids.[191] An Australian study suggested a significant increase in cataract formation associated with prolonged use of these medications, but the effect was small and not much greater than not wearing sunglasses.[192] In a more recent study in the United Kingdom, the use of intranasal corticosteroids was not associated with an increased risk of cataracts.[193] A reduction in bone growth in children has concerned pediatricians.[194] This problem has been studied best in asthmatic children, and the majority of studies suggest no effect and are confounded by the effect of asthma itself on growth in these patients.[195] Therefore, for long-term use, a topical corticosteroid with low bioavailability, administered in the lowest dose necessary to provide relief of symptoms, seems advisable and safe. A recent placebo-controlled clinical study has shown that treatment of children with perennial allergic rhinitis with beclomethasone dipropionate caused a reduction in growth velocity compared with the effect of placebo.[196] This effect was observed in the absence of laboratory evidence of hypothalamic-pituitary-adrenal axis suppression, suggesting that growth velocity may be a more sensitive indicator of systemic corticosteroid exposure than assessment of the hypothalmic-pituitary axis in pediatric patients. On the other hand, treatment of children with perennial allergic rhinitis with mometasone furoate for a year did not result in any growth retardation or hypothalamic-pituitary-adrenal axis suppression compared to placebo.[197] The long-term effects of the reduction in growth velocity, including the final impact on adult height and the ability to catch up in growth after discontinuation of treatment, have not been adequately studied. Recent recommendations state that the growth of pediatric patients receiving intranasal corticosteroids should be monitored regularly (every 3 to 6 months) with an accurate instrument (stadiometer) by trained staff in a consistent way. Furthermore, it seems wise to use the newer agents with the lower systemic bioavailability such as mometasone and fluticasone for the treatment of children. These agents have been approved by the US Food and Drug Administration (FDA) starting at the ages of 3 (mometasone) and 4 years (fluticasone), with the recommended dose being half that used for adults. Mometasone and fluticasone are poorly absorbed from the gastrointestinal tract, with the remaining fraction of absorbed drug rapidly metabolized by the liver.[198,199]

In comparative trials, topical corticosteroids were more potent in relieving nasal symptoms than cromolyn[200] or antihistamines.[201,202] Previously, the addition of antihistamines to intranasal corticosteroids was thought to benefit ocular symptoms and add more rapid onset of effectiveness at the initiation of treatment, although a meta-analysis showed antihistamines and topical corticosteroids to have equivalent effects on ocular symptoms.[203] In a study comparing intranasal corticosteroids with immunotherapy, budesonide was found to be superior to Pollinex-R (Bencard Allergy Service, Weston, Ontario, Canada). Pollinex-R is not a standard, accepted form of immunotherapy, however, and further investigation is warranted in that regard.[204]

Systemic Corticosteroids

Clinical practice confirms the impression that oral corticosteroids reduce symptoms during seasonal allergies, but this has not been documented in placebo-controlled trials.[205] These agents are usually administered to patients during severe exacerbations of allergic symptoms when total nasal obstruction prevents the introduction of a topical intranasal corticosteroid. Furthermore, these agents are used successfully in combination with antibiotics for treatment of sinus infections complicating allergic exacerbations. Depot injections of corticosteroids have efficacy comparable with short-term oral prednisone therapy and enjoy some popularity in Europe, including Scandinavia.[206] Corticosteroid injections into the turbinates are also perceived as clinically effective but have rarely been used in North America since the advent of intranasal corticosteroids and because of a small associated risk of blindness.[207] Although this complication can be minimized by use of a corticosteroid with small particle size, the need to take this infrequent risk in lieu of alternative successful therapies has led to a decline in the popularity of this form of treatment.

Cromolyn Sodium

Cromolyn is available over the counter as a 4% solution for intranasal use and has been shown to be clinically effective in the treatment of allergic rhinitis. It exerts a protective effect on the allergic response when given four to six times daily beginning before the development of symptoms.[208] Although it was

initially thought to prevent mast cell degranulation, the exact mechanism of action of this agent is unknown. Its effectiveness approximates that of antihistamines,[209] but the need for frequent dosing limits compliance. Like antihistamines, cromolyn is more helpful for sneezing, rhinorrhea, and nasal itching than for nasal congestion.[210] Its safety profile, however, makes it an attractive treatment, especially in children and pregnant women. When effective, the potency of this agent parallels that of antihistamines but is less than that of intranasal corticosteroids.

Anticholinergic Agents

Ipratropium bromide is the only anticholinergic agent available for topical use in the United States. Anticholinergic agents inhibit parasympathetic stimulation of glandular secretion by competing for muscarinic receptors on glands. They are highly effective in reducing rhinorrhea but have no effect on the other symptoms of allergic rhinitis.[211,212] The clinical benefit of anticholinergic agents is primarily limited to the treatment of patients with rhinitis in whom rhinorrhea is the predominant complaint. This could occur in a variety of nasal conditions such as allergic and nonallergic rhinitis, as well as the rhinorrhea precipitated by exposure to cold, windy environments, often referred to as "skiers' nose."[213] The dosage should be titrated to avoid excessive drying of the nasal mucosa and epistaxis, which are the most frequent side effects. This agent serves as a useful adjuvant therapy with topical corticosteroids and antihistamines for control of rhinorrhea.

Leukotriene Modifiers

Among the numerous mediators released into the nose on antigen challenge, leukotrienes have been shown to contribute to nasal congestion in allergic rhinitis.[57,58] Symptoms of allergic rhinitis have been shown to be reduced by two clinically available leukotriene modifiers, zafirlukast and montelukast.[214,215] These drugs, although not currently indicated in the primary treatment of allergic rhinitis, may be useful in the treatment of patients with allergic rhinitis and concomitant asthma. The combination of a leukotriene modifier with an antihistamine increases the efficacy of both medications.[216] This combination can be considered as an alternative in patients who do not tolerate intranasal corticosteroids.

Ophthalmic Preparations

Treatment of ocular symptoms of allergy employs a regimen of medications similar to those used for nasal manifestations. After avoidance, pharmacotherapy includes the use of topical decongestants, antihistamines, mast cell–stabilizing agents, and anti-inflammatory preparations.[217] Topical decongestants such as phenylephrine and tetrahydrozoline decrease vascular congestion and eyelid edema through α-adrenergic receptors. Several over-the-counter topical antihistamines are available, some in combination with a decongestant. Recently, second-generation topical antihistamines have become available. Levocabastine is a long-lasting topical antihistamine with a rapid onset of action that has been shown to be effective.[218] Azelastine has recently been approved for the treatment of allergic conjunctivitis.[219] Emedastine and olopatadine are also new topical agents that have also been shown to be effective.[220,221] Ophthalmic formulations of ketotifen fumarate, pemirolast potassium, and nedocromil sodium, agents with mast cell–stabilizing agents and other properties, have recently been approved by the FDA for use in adults and children with itching of the eyes owing to allergic conjunctivitis. Ketorolac, a nonsteroidal anti-inflammatory agent, diminishes ocular itching and hyperemia and is the only agent in this class approved by the FDA for treatment of allergic conjunctivitis.[222] Mild topical corticosteroids can be used in patients who continue to have symptoms despite treatment with the aforementioned agents, but these can cause local complications. Rapidly inactivated topical corticosteroids are under investigation for use in this disease. Although a number of new drugs are now in use, there are few randomized, placebo-controlled trials comparing the efficacy of these agents.

Immunotherapy

The exact mode by which immunotherapy achieves efficacy remains to be fully elucidated.[223] Studies show that this treatment induces a state of specific T-cell tolerance with a subsequent reduction in mediator release and tissue inflammation.[224] Immunotherapy produces several immune modifications in the peripheral blood and nasal mucosa that probably

contribute to its efficacy.[225,226] Treatment causes a rise in serum-specific IgG antibodies, a suppression in the usual seasonal rise in specific IgE antibodies with a decline over years, and an increase in IgA and IgG antibodies in nasal secretions.[227,228] It reduces in vitro lymphocyte responsiveness to antigen[229] and decreases IL-2 release by inflammatory cells.[230] In experimental models of nasal provocation, immunotherapy reduces both early and late responses to ragweed antigen challenge, cellular influx into nasal secretions, and priming.[231] During the allergy season, immunotherapy leads to an inhibition of eosinophil migration into nasal secretions[232] and a reduction in nonspecific hyperreactivity to histamine.[233] A randomized, double-blind, placebo-controlled trial of the discontinuation of immunotherapy for grass-pollen allergy showed a sustained reduction in the late skin response and associated CD3+ T-cell infiltration and IL-4 mRNA expression.[234] These results suggest that long-term changes occur in immune function as a result of immunotherapy, possibly owing to a switch from TH2 to TH1 responses.

Indications for immunotherapy have not been established by experimental studies but rather have evolved over years of clinical experience.[233] The primary indication is symptoms not adequately controlled by avoidance measures and pharmacotherapy. Patients with perennial symptoms may prefer immunotherapy to yearlong daily medication. In making their selection, patients must be advised that immunotherapy offers control of symptoms but is slow in onset and, unlike pharmacotherapy, is effective only for the allergens for which the patient is treated. Whether years of successful therapy cure the disease after discontinuation of immunotherapy is an important but unanswered question. However, one study suggested that improvement in rhinitis symptoms persists for several years after the end of treatment.[234]

Immunotherapy begins with low-dose injections of allergen extracts and builds to a maintenance dose. Injections usually begin at weekly intervals and are then reduced in frequency when maintenance doses are reached. The choice of allergens for treatment must be made after a careful diagnostic workup so that the probability of treatment being started with extracts from all of the allergens responsible for symptoms is increased. Treatment should not be given if skin or serum testing cannot confirm evidence of an IgE-mediated mechanism. Whereas excellent relief can be expected with pollen allergens, dust mites, and some animal danders, treatment with molds is less reliable. Immunotherapy with too many allergens is impractical, and usually treatment with no more than 6 to 10 allergens is attempted. Enzymes from some allergens can destroy other extracts, reducing the expected potency of the treatment.[233] Furthermore, combining allergens of different clinical sensitivities may interfere with reaching adequate maintenance doses for all allergens. Patients who are taking beta blockers should not receive immunotherapy because, if anaphylaxis occurs, they cannot be resuscitated. Immunotherapy of symptomatic asthmatic patients should be administered with extreme caution because these patients have the greatest morbidity. Immunotherapy should not be started in pregnant women because of the risk of anaphylaxis and the resultant effect of hypotension on the fetus.

Improvement of symptoms may begin as soon as 12 weeks, but an optimal effect usually takes 1 year to attain.[233] Effectiveness of immunotherapy requires administration of sufficient amounts of allergen. For ragweed, this has been shown to be between 6 and 12 μg of Amb a 1 (the major antigen in ragweed) per injection. The effect is antigen specific; thus, selection of relevant antigens for treatment is paramount. Patients who do not achieve symptomatic improvement after 1 year of immunotherapy should have it discontinued. No information is available on the length of time over which effective immunotherapy should be pursued, but most physicians treat for 2 to 5 years.

Allergen injections should be administered under the supervision of a qualified medical practitioner, and patients should be observed for at least 30 minutes after every injection. Local reactions at the site of injection are frequent with effective dosages and require no therapy. Repeated strong reactions (ie, greater than 4 cm in diameter) persisting for 24 hours should lead to consideration of dose reduction. Proper resuscitative equipment should be present because anaphylactic reactions can occur at any time during treatment, even in patients receiving maintenance dosages. Factors that seem to contribute to increased complications, including fatalities, seem to be labile, corticosteroid-dependent asthma that has required prior hospitalizations, high sensitivity to allergen (as demonstrated by serum-specific IgE tests or skin test), and a history of prior systemic reactions.[235] Therefore, although death from immunotherapy is uncommon (the risk is estimated at one

fatality for every 2 million injections), special precautions need to be taken with asthma, and a waiting period of at least 20 minutes after administration of the injection is recommended for all patients, with longer intervals (30 minutes) appropriate for high-risk patients.[236] Concern over mortality has led to decreased use in the United Kingdom.

Future Therapies

As the molecular dissection of the complex immunologic pathophysiology of allergic rhinitis has developed, new targets are surfacing for pharmacologic development. Perhaps the most obvious target is IgE, given its central role in allergic rhinitis. An antihuman IgE antibody (rhuMAb-E25) showed promise in early trials[237,238]; a phase 3 trial suggested that this therapy may have potential.[239] A third generation of antihistamines is in phase 2 and phase 3 clinical trials (desloratadine and norastemizole). Development of drugs combining antihistamines and antileukotriene agents is under way. Perhaps most exciting are advances in cytokine biology. Drugs targeting several cytokines known to play a role in the inflammatory process are being developed and tested. These include monoclonal antibodies to IL-4, IL-13, and IL-5; monoclonal antibodies to the receptors of IL-4 and IL-13; and soluble IL-4 and IL-13 receptors.[240] Classes of investigational agents that target other molecules thought to be important in the inflammatory processes of allergic rhinitis include tryptase inhibitors, phosphodiesterase-4 inhibitors, chemokine inhibitors, and adhesion receptor antagonists.[240]

A perhaps more elegant therapeutic strategy would be to effect long-lasting changes in immune responses away from an allergic phenotype. Such a strategy might alter the natural course of disease and allow discontinuation of medication. This type of immunotherapy may involve immunization with allergen, modified allergen, peptides of allergen, or complementary deoxyribonucleic acid (cDNA) of allergen, with adjuvants, including immunostimulatory DNA sequences, cytokines, and bacterial products.[241] These therapies will require further study but ultimately may prove useful in expanding the armamentarium against allergic rhinitis.

Special Considerations

Several patient populations require careful attention in the treatment of allergic rhinitis. In the case of elite athletes, preventive therapy must be initiated so that effects on peak performance can be avoided.[242] The Olympic Committee bans pseudoephedrine-containing agents. Similarly, care must be taken in the treatment of elderly patients with respect to side effects, clearance, and drug interactions. Treatment of rhinitis during pregnancy poses special problems. Rhinitis and nasal congestion frequently occur during pregnancy (30%) and are related to hormonal changes.[243,244] Pregnant patients require careful consideration in the choice of therapy. Ideally, no medication should be used, particularly during the first trimester. Avoidance measures should be implemented first. If symptoms of rhinitis interfere with maternal well-being, pharmacologic management is considered.[245] The patient must be advised that no drug can be considered absolutely safe because most drugs cross the placenta and can be measured in fetal blood.

The American College of Obstetricians and Gynecologists (ACOG) and the American College of Allergy, Asthma and Immunology (ACAAI) make the following recommendations. Pseudoephedrine has been recommended as the decongestant to be used in pregnancy, even though studies show an association with a rare birth defect, gastroschisis.[246] Therefore, the recommendation remains to avoid oral decongestants during the first trimester.[247] Chlorpheniramine and tripelennamine have been recommended as the antihistamines of choice for use in pregnancy, based on animal and human studies. Animal studies have been reassuring for cetirizine and loratidine but not for azelastine or fexofenadine. In patients who do not tolerate chlorpheniramine and tripelennamine and who fail other therapies, use of cetirizine and loratidine could be considered.[247] Sodium cromoglycate is a safe drug during pregnancy and should be considered in the treatment of allergic rhinitis before corticosteroids. No data exist on the teratogenicity of intranasal corticosteroids in humans,[170] although they are thought to be safe and effective.[248] They are widely used during gestation, given their effectiveness in allergic disease and the lack of systemic effects. Studies on inhaled corticosteroids and pregnancy would suggest that beclomethasone and budesonide would be the agents to be initiated. Because the available intranasal corticosteroids do not vary in efficacy or safety and lack systemic side effects, one may consider continuing treatment with these agents in a patient who has well-controlled disease prior to pregnancy and who continues to require

treatment.[247] Newer forms of intranasal corticosteroids with lower bioavailability may be a more prudent choice. Leukotriene modifiers show adverse effects in animals and cannot be recommended in pregnancy. Maintenance immunotherapy may be safely continued during pregnancy in patients who are not prone to systemic reactions. Because of the increased risk of systemic reactions, immunotherapy should not be initiated during pregnancy.

TOWARD A RATIONAL CHOICE OF THERAPY

Prevention remains the mainstay of treatment of allergic rhinitis (Figure 32–5). If allergen exposure can be reduced, this should be part of long-term management. Short-term avoidance does not result in an instant resolution of symptoms and is rarely completely achievable.

Pharmacotherapy provides the fastest relief. Antihistamines begin to take effect within 1 hour and traditionally constitute the first line of intervention. They are excellent for the treatment of sneezing and watery rhinorrhea. When cost is not an issue, nonsedating antihistamines should always be prescribed. To minimize cost, one can administer a long-acting, sedating antihistamine around bedtime and allow its efficacy to persist into the next day without its sedative side effect, although adverse central nervous system effects may still occur. Decongestants can be added to antihistamines in fixed combinations or as separate agents for relief of nasal congestion. Cromolyn sodium is an alternative to antihistamines as an initial treatment, but the need for frequent dosing, with the resultant reduction in compliance, should be kept in mind. Leukotriene modifiers may be used for effects against congestion in asthmatic or refractory cases and may be used in combination with antihistamines in patients who do not tolerate other therapies.

When antihistamines and/or decongestants are insufficient in relieving symptoms, topical corticosteroids should be recommended. These have been shown to be equally effective as antihistamines even when used on an as needed basis. They are highly effective in reducing all of the nasal symptoms of allergic rhinitis, including congestion. They are nonsedating, have few side effects, and are well tolerated by patients. The daily cost of treatment with nasal corticosteroids is less than that of the daily use of nonsedating antihistamines. Topical corticosteroids are initially given at the dose recommended in the *Physicians' Desk Reference*. Patients should be seen 2 weeks after initiation of therapy for monitoring for the development of local side effects. Superficial septal erosions can occur secondary to trauma from the nozzle, and the application technique should be carefully reviewed with these patients. The dosage is adjusted depending on the response: if a patient is doing better but continues to have breakthrough symptoms, the frequency of administration is increased; if excellent control is achieved, the frequency or the dose should be reduced.

Furthermore, periods of exacerbations can be predicted, based on a patient's pattern of allergies; therefore, the medication dosage can be varied accordingly. Ocular symptoms may be minimally controlled by intranasal corticosteroids; thus, adding an ophthalmic preparation or an oral antihistamine may be necessary.

In children, the long-term use of certain topical corticosteroids is approved in patients as young as age 3 years or older. For most topical corticosteroids, it is recommended to reduce the dose by half in young children. Because of the slight possibility that intranasal corticosteroids could interfere with growth, these medications should always be given in the lowest effective dose. Growth should be carefully and regularly monitored. Treatment with topical corticosteroids requires patient education. The physician must often reassure patients as to the safety of intranasal corticosteroids compared with oral preparations.

Initiation of immunotherapy depends on patient preference and the response to pharmacotherapy. Immunotherapy has not been shown to be more effective than intranasal corticosteroids.

Surgical management of nasal obstruction is presented in Chapter 33. Major septal deviations should be corrected because they may interfere with the delivery of intranasal medications. Less severe obstructions should be corrected only if symptoms persist after adequate control of the allergic symptoms. Similarly, turbinate reduction by means of submucous resection, vaporization with the laser, electrocautery, or radiofrequency ablation may be adjuvant procedures in patients with allergic rhinitis but more often find use in the management of nonallergic rhinitis (see Chapter 33). Functional endoscopic sinus surgery, presented in Chapter 34, can be used for the treatment of chronic rhinosinusitis, which frequently complicates perennial rhinitis.[249]

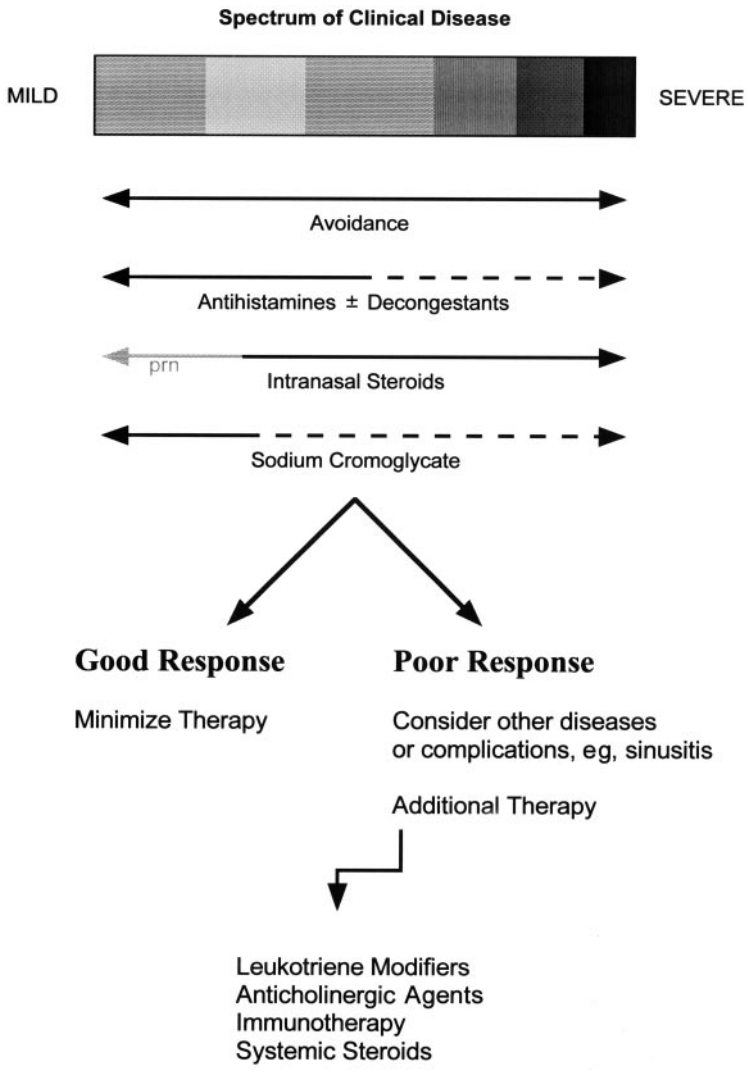


FIGURE 32–5. Approach to the treatment of allergic rhinitis. This schematic depicts the different treatments available for allergic rhinitis and establishes some guidelines for the use of these therapies in different stages of severity of the disease. Avoidance should be used in all stages. Antihistamines and/or decongestants as well as sodium cromoglycate are commonly used for allergic rhinitis of mild to moderate severity. Topical intranasal corticosteroids are also effective against a range of symptoms, including on an as needed basis. Severe disease usually warrants more aggressive treatment. This may require combination or adjuvant therapy. Anticholinergic agents play a limited role in the control of rhinorrhea. Leukotriene modifiers are useful in the treatment of congestion, especially in asthmatic patients. Systemic corticosteroids are reserved for severe allergic rhinitis leading to complete nasal obstruction. Immunotherapy can be offered to patients who fail medical therapy. Clearly, these guidelines need to be varied in the context of different clinical settings as some patients with mild disease might benefit from intranasal corticosteroids, whereas some other patients with more severe disease might respond well to antihistamine/decongestant combinations. prn = as needed.

REFERENCES

1. Sly RM. Changing prevalence of allergic rhinitis and asthma. Ann Allergy Asthma Immunol 1999;82: 233–48.
2. Collins J, editor. Prevalence of selected chronic conditions, United States, 1983–1985. Hyattsville (MD): US Public Health Service. Advance data from Vital and Health Statistics, No 155. Vol DHHS Pub No (PHS): 88.
3. Juniper EF, Guyatt GH. Developing and testing a new measure of health status for clinical trials in rhinoconjunctivitis. Clin Exp Allergy 1991;21:77–83.
4. Juniper EF, Guyatt GH, Anderson B, et al. Comparison of power and aerosolized budesonide in perennial rhinitis: validation of rhinitis quality of life questionnaire. Ann Allergy 1993;70:225–30.
5. Task Force on Allergic Disorders. Promoting best practice: raising the standard of care for patients with allergic disorders. Executive summary report. Milwaukee (WI): American Academy of Allergy, Asthma, and Immunology; 1998. p. 1–15.
6. Aberg N. Asthma and allergic rhinitis in Swedish conscripts. Clin Exp Allerg 1989;19:59–63.
7. Linneberg A, Nielsen NH, Madesen F, et al. Increasing prevalence of specific IgE to aeroallergens in an adult population: two cross sectional studies 8 years apart. The Copenhagen Allergy Study. J Allergy Clin Immunol 2000;106:247–52.
8. von Mutius E, Weiland SJ, Fritzsch C, et al. Increasing prevalence of hay fever and atopy among children in Leipzig, East Germany. Lancet 1998;308:1591–6.
9. Bascom R, Naclerio RM, Fitzgerald TK, et al. Effect of ozone inhalation on the response to nasal chal-

lenge with antigen of allergic subjects. Am Rev Respir Dis 1990;142:594–601.
10. Molfino NA, Slutsky AS, Zamel N. The effect of air pollution on allergic bronchial responsiveness. Clin Exp Allergy 1992;22:667–72.
11. Matsumura Y. The effects of ozone, nitrogen dioxide, and sulfur dioxide on the experimentally induced allergic respiratory disorders in guinea pigs 1. The effect of sensitization with albumin through the airway. Am Rev Respir Dis 1970;102:430–7.
12. Nel AE, Diaz Sanchez D, Ng D, et al. Enhancement of allergic inflammation by the interaction between diesel exhaust particles and the immune system. J Allergy Clin Immunol 1998;102(Pt 1):539–54.
13. von Mutius E, Martinez FD, Fritzsch C, et al. Prevalence of asthma and atopy in two areas of West and East Germany. Am J Respir Crit Care Med 1994;149:358–64.
14. Magnussen H, Jörres R, Nowak D. Effect of air pollution on the prevalence of asthma and allergy: lessons from the German reunification. Thorax 1993;48:879–81.
15. Hide DW, Matthews S, Matthews L, et al. Effect of allergen avoidance in infancy on allergic manifestations at age two years. J Allergy Clin Immunol 1994;93:842–6.
16. Shaheen SO, Aaby P, Hall AJ, et al. Measles and atopy in Guinea-Bissau. Lancet 1996;347:1792–6.
17. Shirakawa T, Enomoto T, Shimazu S, Hopkin JM. The inverse association between tuberculin responses and atopic disorders. Science 1997;275:77–9.
18. Matricardi PM, Rosmini F, Ferrigno L, et al. Cross sectional, retrospective study of prevalence of atopy among Italian military students with antibodies against hepatitis A virus. BMJ 1997;314:999–1003.
19. von Mutius E, Martinez FD, Fritzsch C, et al. Skin test reactivity and number of siblings. BMJ 1994; 308:692–5.
20. Strachan DP, Harkins LS, Johnston IDA, Anderson HR. Childhood antecedents of allergic sensitization in young British adults. J Allergy Clin Immunol 1997;99:6–12.
21. Strachan DP. Hay fever, hygiene, and household size. BMJ 1989;299:1259–60.
22. Strachan, DP, Taylor EM, Carpenter RG. Family structure, neonatal infection, and hay fever in adolescence. Arch Dis Child 1996;74:422–6.
23. Rona RJ, Duran-Tauleria E, Chinn S. Family size, atopic disorders in parents, asthma in children, and ethnicity. J Allergy Clin Immunol 1997;99:454–60.
24. Ponsonby AL, Couper D, Dwyer T, Carmichael A. Cross sectional study of the relation between sibling number, asthma, hay fever, and eczema. Arch Dis Child 1998;79:328–33.
25. Ball TM, Castro-Rodriguez JA, Griffith KA, et al. Siblings, day-care attendance, and the risk of asthma and wheezing during childhood. N Engl J Med 2000;343:538–43.
26. Martinez FD. Maturation of immune responses at the beginning of asthma. J Allergy Clin Immunol 1999;103:355–61.
27. Evans RI. Epidemiology and natural history of asthma, allergic rhinitis, and atopic dermatitis. In: Middleton EJ, Reed CE, Ellis EF, et al, editors. Allergy: principles and practice. 4th ed. St. Louis: Mosby-Year Book; 1993.
28. Malmberg H. Symptoms of chronic and allergic rhinitis and occurrence of nasal secretion granulocytes in university students, children and infants. Allergy 1979;34:389–94.
29. Zeiger RS. Development and prevention of allergic disease in childhood. In: Middleton EJ, Reed CE, Ellis EF, et al, editors. Allergy: principles and practice. 4th ed. St. Louis: Mosby-Year Book; 1993. p. 1137–72.
30. Aberg N, Engstrom I. Natural history of allergic diseases in children. Acta Paediatr Scand 1990;79: 206–11.
31. The Collaborative Study on the Genetics of Asthma (CSGA). A genome-wide search for asthma susceptibility loci in ethnically diverse populations. Nat Genet 1997;15:389–92.
32. Daniels SE, Bhattacharrya S, James A, et al. A genome-wide search for quantitative trait loci underlying asthma. Nature 1996;383:247–50.
33. Dizier MH. Genome screen for asthma and related phenotypes in the French EGEA study. Am J Respir Crit Care Med 2000;162:1812–8.
34. Ober C, Tsalenko A, Parry R, Cox NJ. A second-generation genome wide screen for asthma-susceptibility alleles in a founder population. Am J Hum Genet 2000;67:1154–6.
35. Wjst M. Specific IgE—one gene fits all? German Asthma Genetics Group. Clin Exp Allergy 1999;29 Suppl 4:5–10.
36. Barnes KC. Evidence for common genetic elements in allergic disease. J Allergy Clin Immunol 2000; 106(5 Suppl):S192–200.
37. Ober C, Moffatt MF. Contributing factors to the pathobiology. The genetics of asthma. Clin Chest Med 2000;21:245–61.
38. Blumenthal JB, Blumenthal MN. Immunogenetics of allergy and asthma. Immunol Allergy Clin North Am 1996;16:517–35.

39. Togias AG. Systemic immunologic and inflammatory aspects of allergic rhinitis. J Allergy Clin Immunol 2000;106(5 Suppl):S247–50.
40. Church JA, Bauer H, Bellanti JA, et al. Hyposmia associated with atopy. Ann Allergy 1978;40:105–9.
41. Binder E, Holopainen E, Malmberg K, Salo OP. Clinical findings in patients with allergic rhinitis. Rhinology 1984;22:255–60.
42. Settipane GA, Chafee FH. Nasal polyps in asthma and rhinitis. A review of 6,037 patients. J Allergy Clin Immunol 1977;59:17–21.
43. Friedman RA, Doyle WJ, Casselbrant ML, et al. Immunologic-mediated eustachian tube obstruction: a double-blind crossover study. J Allergy Clin Immunol 1983;71:442–7.
44. Hurst DS, Amin K, Seveus L, Venge P. Mast cells and tryptase in the middle ear of children with otitis media with effusion. Int J Pediatr Otorhinolaryngol 1999;49 Suppl 1:S315–9.
45. Wright ED, Hurst D, Miotto D, et al. Increased expression of major basic protein (MBP) and interleukin-5 (IL-5) in middle ear biopsy specimens from atopic patients with persistent otitis media with effusion. Otolaryngol Head Neck Surg 2000; 123:533–8.
46. Bernstein JM, Lee J, Conboy K, et al. Role of IgE mediated hypersensitivity in recurrent otitis media with effusion. Am J Otol 1983;5:66–9.
47. Enerbäck L, Pipkorn U, Olofsson A. Intraepithelial migration of mucosal mast cells in hay fever: ultrastructural observations. Int Arch Allergy Appl Immunol 1986;81:289–97.
48. Cameron L, Hamid Q, Wright E, et al. Local synthesis of epsilon germline gene transcripts, IL-4, and IL-13 in allergic nasal mucosa after ex vivo allergen exposure. J Allergy Clin Immunol 2000;106(1 Pt 1): 46–52.
49. Durham SR, Smurthwaite L, Gould HJ. Local IgE production. Am J Rhinol 2000;14:305–7.
50. Plaut M, Pierce JH, Watson CJ, et al. Mast cell lines produce lymphokines in response to cross-linkage of Fc epsilon RI or to calcium ionophores. Nature 1989;339:64–7.
51. Bradding P, Feather IH, Howarth PH, et al. Interleukin 4 is localized to and released by human mast cells. J Exp Med 1992;176:1381–6.
52. Naclerio RM, Meier HL, Kagey-Sobotka A, et al. Mediator release after airway challenge with antigen. Am Rev Respir Dis 1983;128:597–602.
53. Baumgarten CR, Togias AG, Naclerio RM, et al. Influx of kininogens into nasal secretions after antigen challenge of allergic individuals. J Clin Invest 1985;76:191–7.
54. Baumgarten CR, Nichols RC, Naclerio RM, et al. Plasma kallikrein during experimentally induced allergic rhinitis: role in kinin formation and contribution of TAME-esterase activity. J Immunol 1986;137:977–82.
55. Baumgarten CR, Nichols RC, Naclerio RM, Proud D. Concentration of glandular kallikrein in human nasal secretions increase during experimentally induced allergic rhinitis. J Immunol 1986;137:1323–8.
56. Castells M, Schwartz LB. Tryptase levels in nasal-lavage fluid is an indicator of the immediate allergic reaction. J Allergy Clin Immunol 1988;82:348–55.
57. Creticos PS, Peters SP, Adkinson NF, et al. Peptide leukotriene release after antigen challenge in patients sensitive to ragweed. N Engl J Med 1984; 310:1626–30.
58. Peters SP, Freeland HS, Kelly SJ, et al. Is leukotriene B4 an important mediator in human IgE-mediated allergic reactions? Am Rev Respir Dis 1987;135(6 Pt 2):S42–5.
59. Bascom R, Pipkorn U, Proud D, et al. Major basic protein and eosinophil-derived neurotoxin concentrations in nasal-lavage fluid after antigen challenge: effect of systemic corticosteroids and relationship to eosinophil influx. J Allergy Clin Immunol 1989;84: 338–46.
60. Togias AG, Lykens K, Kagey-Sabotka A, et al. Studies on the relationship between sensitivity to cold dry air, hyperosmolar solutions and histamine in the adult nose. Am Rev Respir Dis 1990;141:1428.
61. Tønnesen P, Mygind N. Nasal challenge with serotonin and histamine in normal persons. Allergy 1985;40:350–3.
62. Karim SM, Adaikian PG, Kunaratnam N. Effect of topical prostaglandins on nasal patency in man. Prostaglandins 1978;15:457–62.
63. Bisgaard H, Olsson P, Bende M. Effect of leukotriene D4 on nasal mucosal blood flow, nasal airway resistance and nasal secretions in humans. Clin Allergy 1986;16:289–97.
64. Proud D, Reynolds CJ, Lacapra S, et al. Nasal provocation with bradykinin induces symptoms of rhinitis and a sore throat. Am Rev Respir Dis 1988;137:613–6.
65. Doyle WJ, Boehm S, Skoner DP. Physiologic responses to intranasal dose-response challenges with histamine, metacholine, bradykinin, and prostaglandin in adult volunteers with and without nasal allergy. J Allergy Clin Immunol 1990;86: 924–35.

66. Gleich GJ. The late phase of the immunoglobulin E-mediated reaction: a link between anaphylaxis and common allergic disease? J Allergy Clin Immunol 1982;70:160–9.
67. Naclerio RM, Proud D, Togias AG, et al. Inflammatory mediators in late antigen-induced rhinitis. N Engl J Med 1985;313:65–70.
68. Iliopoulos O, Proud D, Adkinson NF, et al. Relationship between the early, late, and rechallenge reaction to nasal challenge with antigen: observations on the role of inflammatory mediators and cells. J Allergy Clin Immunol 1990;86:851–61.
69. Taylor G, Shivalkar PR. 'Arthus-type' reactivity in the nasal airways and skin in pollen sensitive subjects. Clin Allergy 1971;1:407–14.
70. Bascom R, Wachs M, Naclerio RM, et al. Basophil influx occurs after nasal antigen challenge: effects of topical corticosteroid pretreatment. J Allergy Clin Immunol 1988;81:580–9.
71. Lim-Mombay M, Baroody FM, Taylor R, Naclerio RM. Mucosal cellular changes after nasal antigen challenge [abstract]. J Allergy Clin Immunol 1993;89:205.
72. Varney VA, Jacobson MR, Sudderick RM, et al. Immunohistology of the nasal mucosa following allergen-induced rhinitis. Identification of activated T lymphocytes, eosinophils, and neutrophils. Am Rev Respir Dis 1992;146:170–6.
73. Bryan WTK, Bryan MP. Significance of mast cells in nasal secretions. Trans Am Acad Ophthalmol Otolaryngol 1959;63:613.
74. Okuda M, Otsuka H. Basophilic cells in allergic nasal secretions. Arch Otorhinolaryngol 1977;214:283–9.
75. Okuda M, Kawabori S, Otsuka H. Electron microscope study of basophilic cells in allergic nasal secretions. Arch Otorhinolaryngol 1978;221:215–20.
76. Hastie R, Heroy JH, Levy DA. Basophil leukocytes and mast cells in human nasal secretions and scrapings studied by light microscopy. Lab Invest 1979;40:554–61.
77. Okuda M, Otsuka H, Kawabori S. Basophil leukocytes and mast cells in the nose. Eur J Respir Dis 1983;64 Suppl 128:7–15.
78. Otsuka H, Denburg J, Dolovich J, et al. Heterogeneity of metachromatic cells in the human nose: significance of mucosal mast cells. J Allergy Clin Immunol 1985;76:695–702.
79. Bentley AM, Jacobson MR, Cumberworth V, et al. Immunohistology of the nasal mucosa in seasonal allergic rhinitis: increases in activated eosinophils and epithelial mast cells. J Allergy Clin Immunol 1992;89:877–83.
80. Lee BJ, Naclerio RM, Bochner BS, et al. Nasal challenge with allergen upregulates the local expression of vascular endothelial adhesion molecules. J Allergy Clin Immunol 1994;94(6 Pt 1):1006–16.
81. Walsh GM, Mermod JJ, Hartnell A, et al. Human eosinophil, but not neutrophil, adherence to IL-1 stimulated human umbilical vascular endothelial cells is alpha 4 beta 1 (very late antigen-4) dependent. J Immunol 1991;146:3419–23.
82. Lobb RR. Integrin-immunoglobulin superfamily interactions in endothelial-leukocyte adhesion. In: Harlan JM, Liu DY, editors. Adhesion: its role in inflammatory disease. New York: WH Freeman; 1992. p. 1–15.
83. Schleimer RP, Sterbinsky SA, Kaiser J, et al. IL-4 induces adherence of human eosinophils and basophils but not neutrophils to endothelium. Association with expression of VCAM-1. J Immunol 1992;148:1086–92.
84. Bochner BS, Luscinskas FW, Gimbrone MA, et al. Adhesion of human basophils, eosinophils, and neutrophils to interleukin1-activated human vascular endothelial cells: contributions of endothelial cell adhesion molecules. J Exp Med 1991;173:1553–7.
85. Phillips ML, Nudelman E, Gaeta FC, et al. ELAM-1 mediates cell adhesion by recognition of carbohydrate ligand, sialyl-Lex. Science 1990;250:1130–2.
86. Montefort S, Feather IH, Wilson SJ, et al. The expression of leukocyte-endothelial adhesion molecules is increased in perennial allergic rhinitis. Am J Respir Cell Mol Biol 1992;7:393–8.
87. Robinson DS. Interleukin-5, eosinophils and bronchial hyperreactivity. Clin Exp Allergy 1993;23:1–3.
88. Altman LC, Ayars GH, Baker C, Luchtel DL. Cytokines and eosinophil-derived cationic proteins upregulate intercellular adhesion molecule-1 on human nasal epithelial cells. J Allergy Clin Immunol 1993;92: 527–36.
89. Holgate S. Mediator and cytokine mechanisms in asthma. Thorax 1993;48:103–9.
90. Durham SR, Ying S, Varney VA, et al. Cytokine messenger RNA expression for IL-3, IL-4, IL-5, and granulocyte/macrophage-colony-stimulating factor in the nasal mucosa after local allergen provocation: relationship to tissue eosinophilia. J Immunol 1992;148:2390–4.
91. Alam R, Sim TC, Hilsmeier K, Grant JA. Development of a new technique for recovery of cytokines from inflammatory sites in situ. J Immunol Methods 1992;155:25–9.
92. Christodoulopoulos P, Cameron L, Durham S, Hamid Q. Molecular pathology of allergic disease.

II: upper airway disease. J Allergy Clin Immunol 2000;105:211–23.
93. Hanazawa T, Antuni JD, Kharitonov SA, Barnes PJ. Intranasal administration of eotaxin increases nasal eosinophils and nitric oxide in patients with allergic rhinitis. J Allergy Clin Immunol 2000;105:58–64.
94. Sim TC, Reece LM, Hilsmeier KA, et al. Secretion of chemokines and other cytokines in allergen-induced nasal responses: inhibition by topical steroid treatment. Am J Respir Crit Care Med 1995;152:927–33.
95. Bradding P, Feather IH, Wilson S, et al. Immunolocalization of cytokines in the nasal mucosa of normal and perennial rhinitic subjects. The mast cell as a source of IL-4, IL-5, and IL-6 in human allergic mucosal inflammation. J Immunol 1993;151: 3853–65.
96. Kuna P, Alam R, Ruta U, Gorski P. RANTES induces nasal mucosal inflammation rich in eosinophils, basophils and lymphocytes in vivo. Am J Respir Crit Care Med 1998;157:873–9.
97. Connell JT. Quantitative intranasal pollen challenges. 3. The priming effect in allergic rhinitis. J Allergy 1969;43:33–44.
98. Walden SM, Proud D, Lichtenstein LM, et al. Antigen-provoked increase in histamine reactivity. Observations on mechanisms. Am Rev Respir Dis 1991;144:642–8.
99. Andersson M, Andersson P, Pipkorn U. Allergen-induced specific and non-specific nasal reactions: reciprocal relationship and inhibition by topical glucocorticosteroids. Acta Otolaryngol (Stockh) 1989;107:270–7.
100. Baroody FM, Cruz AA, Lichtenstein LM, et al. Intranasal beclomethasone inhibits antigen-induced nasal hyperresponsiveness to histamine. J Allergy Clin Immunol 1992;90:373–6.
101. Majchel AM, Proud D, Freidhoff L, et al. The nasal response to histamine challenge: effect of the pollen season and immunotherapy. J Allergy Clin Immunol 1992;90:85–91.
102. Druce HM, Wright RH, Kossoff D, Kaliner MA. Cholinergic nasal hyperreactivity in atopic subjects. J Allergy Clin Immunol 1985;76:445–52.
103. Klementsson H, Andersson M, Baumgarten CR, et al. Changes in nonspecific nasal reactivity and eosinophil influx and activation after allergen challenge. Clin Exp Allergy 1990;20:539–47.
104. Klementsson H, Andersson M, Pipkorn U. Allergen-induced increase in nonspecific nasal reactivity is blocked by antihistamines without a clear-cut relationship to eosinophil influx. J Allergy Clin Immunol 1990;86:466–72.
105. Baroody FM, Lim MC, Proud D, et al. Effects of loratadine and terfenadine on the induced nasal allergic reaction. Arch Otolaryngol Head Neck Surg 1996;122:309–16.
106. Klementsson H, Svensson C, Andersson M, et al. Eosinophils, secretory responsiveness and glucocorticoid-induced effects on the nasal mucosa during a weak pollen season. Clin Exp Allergy 1991;21:705–10.
107. Konno A, Togawa K. Role of the vidian nerve in nasal allergy. Ann Otol Rhinol Laryngol 1979;88:258–66.
108. Baroody FM, Ford S, Lichtenstein LM, et al. Physiologic responses and histamine release after nasal antigen challenge. Effect of atropine. Am J Respir Crit Care Med 1994;149:1457–65.
109. Raphael GD, Igarashi Y, White MV, Kaliner MA. The pathophysiology of rhinitis. V. Sources of protein in allergen-induced nasal secretions. J Allergy Clin Immunol 1991;88:33–42.
110. Wagenmann M, Baroody FM, Desrosiers M, et al. Unilateral nasal allergen challenge leads to bilateral release of prostaglandin D_2. Clin Exp Allergy 1996;26:371–8.
111. Okayama M, Mullol J, Baraniuk JN, et al. Muscarinic receptor subtypes in human nasal mucosa: characterization, autoradiographic localization, and function in vitro. Am J Respir Cell Mol Biol 1993;8:176–87.
112. Baraniuk JN, Lundgren JD, Okayama M, et al. Substance P and neurokinin A in human nasal mucosa. Am J Respir Cell Mol Biol 1991;4:228–36.
113. Okamoto Y, Shirotori K, Kudo K, et al. Cytokine expression after the topical administration of substance P to human nasal mucosa. The role of substance P in nasal allergy. J Immunol 1993;151:4391–8.
114. Mosimann BL, White MV, Hohman RJ, et al. Substance P, calcitonin gene-related peptide, and vasoactive intestinal peptide increase in nasal secretions after allergen challenge in atopic patients. J Allergy Clin Immunol 1993;92:95–104.
115. Philip G, Baroody FM, Proud D, et al. The human nasal response to capsaicin. J Allergy Clin Immunol 1994;94:1035–45.
116. Philip G, Sanico AM, Togias A. Inflammatory cellular influx follows capsaicin nasal challenge. Am J Respir Crit Care Med 1996;153:1222–9.
117. Stjarne P, Rinder J, Heden-Blomquist E, et al. Capsaicin desensitization of the nasal mucosa reduces symptoms upon allergen challenge in patients with allergic rhinitis. Acta Otolaryngol (Stockh) 1998; 118:235–9.

118. Benninger M. Rhinitis, sinusitis and their relationships to allergies. Am J Rhinol 1992;6:37–43.
119. Grove R, Farrior J. Chronic hyperplastic sinusitis in allergic patients. A bacteriologic study of two hundred operative cases. J Allergy 1990;11:271–6.
120. Friedman W. Surgery for chronic hyperplastic rhinosinusitis. Laryngoscope 1975;85:199–201.
121. van Dishoeck H, Franssen M. The incidence and correlation of allergy and chronic maxillary sinusitis. Pract Otolaryngol 1957;19:502–6.
122. Rachelefsky G, Goldberg M, Kutz R, et al. Sinus disease in children with respiratory allergy. J Allergy Clin Immunol 1978;61:310–4.
123. Shapiro G. Role of allergy in sinusitis. Pediatr Infect Dis 1985;4:555–8.
124. Enberg R. Perennial nonallergic rhinitis: a retrospective review. Ann Allergy 1989;63:513–6.
125. Conner BL, Roach ES, Laster WS, Georgitis JW. Magnetic resonance imaging of the paranasal sinuses: frequency and type of abnormalities. Ann Allergy 1989;62:457–60.
126. Naclerio RM, DeTineo ML, Baroody FM. Ragweed allergic rhinitis and the paranasal sinuses: a computed tomographic study. Arch Otolaryngol Head Neck Surg 1997;123:193–6.
127. Baroody FM, DeTineo M, Haney L, et al. Influx of eosinophils into the maxillary sinus after nasal challenge with allergen. J Allergy Clin Immunol 2000;105 Suppl:S70.
128. Vignola AM, Chanez P, Godard P, Bousquet J. Relationships between rhinitis and asthma. Allergy 1998;53:833–9.
129. Corren JAD, Adinoff AD, Buchmeier AD, Irvin CG. Nasal beclomethasone prevents the seasonal increase in bronchial hyperresponsiveness in patients with allergic rhinitis and asthma. J Allergy Clin Immunol 1992;90:250–6.
130. Greiff L, Andersson M, Svensson C, et al. Effects of orally inhaled budesonide in seasonal allergic rhinitis. Eur Respir J 1998;11:1268–73.
131. Braunstahl GJ, Kleinjan A, Overbeek SE, et al. Segmental bronchial provocation induces nasal inflammation in allergic rhinitis patients. Am J Respir Crit Care Med 2000;161:2051–7.
132. Assanasen PA, Baroody FM, Naureckas E, et al. The nasal passages of asthmatic subjects have a decreased ability to warm and humidify inspired air. Am J Respir Crit Care Med 2001;164:1640–46.
133. Fireman P. Otitis media and eustachian tube dysfunction: connection to allergic rhinitis. J Allergy Clin Immunol 1997;99 Suppl:S787–97.
134. Solomon WR. Aerobiology and inhalant allergens: pollens and fungi. In: Middleton EJ, Reed CE, Ellis EF, et al, editors. Allergy: principles and practice. 4th ed. St. Louis: Mosby-Year Book; 1993. p. 469–528.
135. Arlian LG, Neal JS, Morgan MS, et al. Reducing relative humidity is a practical way to control dust mites and their allergens in homes in temperate climates. J Allergy Clin Immunol 2001;107:99–104.
136. Almqvist C, Larsson PH, Egmar AC, et al. School as a risk environment for children allergic to cats and a site for transfer of cat allergen to homes. J Allergy Clin Immunol 1999;103:1012–7.
137. Skoner DP, Doyle WJ, Boehm S, Fireman P. Priming of the nose and eustachian tube during natural pollen season. Am J Rhinol 1989;3:53.
138. Reisman RE, Bernstein JM. Allergy and secretory otitis media: clinical and immunologic studies. Pediatr Clin North Am 1975;22:251–7.
139. Van Cauwenberge P. Secretoire otitis media: een epidemiologische, klinische en experimentale stuie. Ghent, Belgium: Facultent Geneeskunde 1988.
140. Long WF, Taylor RJ, Wagner CJ, et al. Skin test suppression by antihistamines and the development of subsensitivity. J Allergy Clin Immunol 1985;76: 113–7.
141. Bousquet J, Michel FB. In vivo methods for the study of allergy: skin tests. In: Middleton EJ, Reed CE, Ellis EF, et al, editors. Allergy: principles and practice. 4th ed. St. Louis: Mosby-Year Book; 1993. p. 573–94.
142. Menardo JL, Bousquet J, Rodiere M, et al. Skin test reactivity in infancy. J Allergy Clin Immunol 1985;75:646–51.
143. Ownby DR. Allergy testing: in vivo versus in vitro. Pediatr Clin North Am 1988;35:995–1008.
144. Norman P. Correlations of RAST with in vivo and in vitro assays. In: Evans R, editor. Advances in diagnosis of allergy, RAST. Miami (FL): Symposia Specialists; 1975. p. 45–56.
145. Berg TLO, Johansson SGO. Allergy diagnosis with radioallergosorbent test: a comparison with the results of skin and provocation tests in an unselected group of children with asthma and hay fever. J Allergy Clin Immunol 1974;54:209.
146. Bickmore J. Nasal cytology in allergy and infection. ORL Allergy 1978;40:39.
147. Meltzer E, Orgel H, Bronsky E, et al. A dose ranging study of fluticasone propionate aqueous nasal spray for seasonal allergic rhinitis assessed by symptoms, rhinomanometry and nasal cytology. J Allergy Clin Immunol 1990;86:221–30.
148. Lans D, Alfano N, Rocklin R. Nasal eosinophilia in allergic and non-allergic rhinitis: usefulness of the

nasal smear in the diagnosis of allergic rhinitis. Allergy Proc 1989;10:275–80.
149. Miller RE, Paradise JL, Friday GA, et al. The nasal smear for eosinophils. Its value in children with seasonal allergic rhinitis. Am J Dis Child 1982; 136:1009–11.
150. Meltzer E, Orgel HA, Jalowayski A. Cytology. In: Naclerio RM, Mygind NM, editors. Allergic and non-allergic rhinitis: clinical aspects. Copenhagen: Munksgaard; 1993. p. 66–81.
151. Mullarkey MF, Hill JS, Webb DR. Allergic and non-allergic rhinitis: their characterization with attention to the meaning of nasal eosinophilia. J Allergy Clin Immunol 1980;65:122–6.
152. Kniest FM, Young E, Van Praag MC, et al. Clinical evaluation of a double-blind dust-mite avoidance trial with mite-allergic rhinitic patients. Clin Exp Allergy 1991;21:39–47.
153. De Blay F, Chapman MD, Platts-Mills TA. Airborne cat allergen (Fel d I). Environmental control with cat in situ. Am Rev Respir Dis 1991;143:1334–9.
154. Reisman RE, Mauriello PM, Davis GB, et al. A double blind study of the effectiveness of a high-efficiency particulate air (HEPA) filter in the treatment of patients with perennial allergic rhinitis and asthma. J Allergy Clin Immunol 1990;85: 1050–7.
155. Platts-Mills TAE, Chapman MD. Dust mites: immunology, allergic disease, and environmental control. J Allergy Clin Immunol 1987;80:755–75.
156. Drovin MA. H_1 antihistamines: perspective on the use of conventional and new agents. Ann Allergy 1985;55:797.
157. Weiler JM, Bloomfield JR, Woodworth GG, et al. Effects of fexofenadine, diphenhydramine, and alcohol on driving performance. A randomized, placebo-controlled trial in the Iowa driving simulator. Ann Intern Med 2000;132:354–63.
158. Togias AG, Naclerio RM, Warner J, et al. Demonstration of inhibition of mediator release from human mast cells by azatadine base. In vivo and in vitro evaluation. JAMA 1986;255:225–9.
159. Naclerio RM, Kagey-Sobotka A, Lichtenstein LM, et al. Terfenadine, an H_1 antihistamine, inhibits histamine release in vivo in the human. Am Rev Respir Dis 1990;142:167–71.
160. Naclerio RM, Proud D, Kagey-Sobotka A, et al. The effect of cetirizine on early allergic response. Laryngoscope 1989;99:596–9.
161. Campbell A, Chanal I, Czarlewski W, et al. Reduction of soluble ICAM-1 levels in nasal secretion by H1-blockers in seasonal allergic rhinitis. Allergy 1997;52:1022–5.
162. Virant FS. Allergic rhinitis. Immunol Allergy Clin North Am 2000;20:265–83.
163. LaForce C, Dockhorn RJ, Prenner BM, et al. Safety and efficacy of azelastine nasal spray (Astelin NS) for seasonal allergic rhinitis: a 4-week comparative multicenter trial. Ann Allergy Asthma Immunol 1996; 76:181–8.
164. Bousquet J, Duchateau J, Pignat JC, et al. Improvement of quality of life by treatment with cetirizine in patients with perennial allergic rhinitis as determined by a French version of the SF-36 questionnaire. J Allergy Clin Immunol 1996;98:309–16.
165. Meltzer EO, Casale TB, Nathan RA, Thompson AK. Once-daily fexofenadine HCl improves quality of life and reduces work and activity impairment in patients with seasonal allergic rhinitis. Ann Allergy Asthma Immunol 1999;83:311–7.
166. Corren J. Allergic rhinitis: treating the adult. J Allergy Clin Immunol 2000;105:S610–5.
167. Fireman P. Therapeutic approaches to treating allergic rhinitis: treating the child. J Allergy Clin Immunol 2000;105:S616–21.
168. Simons FE, McMillan JL, Simons KJ. A double-blind, single dose, cross over comparison of cetirizine, terfenadine, loratadine, astemizole, and chlorpheniramine versus placebo: suppressive effects on histamine-induced wheal and flares during 24 hours in normal subjects. J Allergy Clin Immunol 1990;86:540–7.
169. Majchel AM, Proud D, Kagey-Sobotka A, et al. Evaluation of a bedtime dose of a combination antihistamine/analgesic/decongestant product on antigen challenge the next morning. Laryngoscope 1992;102:330–4.
170. Lacroix J. Adrenergic and non-adrenergic mechanisms in the sympathetic vascular control of the nasal mucosa. Acta Physiol Scand 1989; Suppl 581:1–63.
171. Compendium of pharmaceuticals and specialties. 29th ed. Toronto: CK Productions; 1994.
172. Schleimer R. Glucocorticosteroids: their mechanisms of activation and use in allergic diseases. In: Middleton EJ, Reed CE, Ellis EF, et al, editors. Allergy: principles and practice. 4th ed. St. Louis: Mosby-Year Book; 1993. p. 893–925.
173. Cox G, Ohtoshi T, Vancheri C, et al. Promotion of eosinophil survival by human bronchial epithelial cells and its modulation by steroids. Am J Respir Cell Mol Biol 1991;4:525–31.
174. Pipkorn U, Proud D, Lichtenstein LM, et al. Inhibition of mediator release in allergic rhinitis by pre-

treatment with topical glucocorticosteroids. N Engl J Med 1987;316:1506–10.
175. Bascom R, Wachs M, Naclerio RM, et al. Basophil influx occurs after nasal antigen challenge: effects of topical corticosteroid pretreatment. J Allergy Clin Immunol 1988;81:580–9.
176. Ratner PH, Paull BR, Findlay SR, et al. Fluticasone propionate given once daily is as effective for seasonal allergic rhinitis as beclomethasone dipropionate given twice daily. J Allergy Clin Immunol 1992;90:285–91.
177. van As A, Bronsky EA, Dockhorn RJ, et al. Once daily fluticasone propionate is as effective for perennial allergic rhinitis as twice daily beclomethasone dipropionate. J Allergy Clin Immunol 1993;91:1146–54.
178. Naclerio RM, Adkinson NF, Creticos PS, et al. Intranasal steroids inhibit seasonal increases in ragweed-specific immunoglobulin E antibodies. J Allergy Clin Immunol 1993;92:717–21.
179. Bende M, Lindqvist N, Pipkorn U. Effect of topical glucocorticoid, budesonide, on nasal mucosal blood flow as measured with ^{133}Xe wash-out technique. Allergy 1983;38:461–4.
180. Juniper EF, Guyatt GH, O'Byrne PM, Viveiros M. Aqueous beclomethasone dipropionate nasal spray: regular versus "as required" use in the treatment of seasonal allergic rhinitis. J Allergy Clin Immunol 1990;86:380–6.
181. Juniper EF, Guyatt GH, Archer B, Ferrie PJ. Aqueous beclomethasone diproprionate in the treatment of ragweed pollen-induced rhinitis: further exploration of "as needed" use. J Allergy Clin Immunol 1993;92:66–72.
182. Jen A, Baroody F, deTineo M, et al. PRN use of Flonase reduces symptoms of ragweed rhinitis. J Allergy Clin Immunol 2000;105:732–8.
183. Baroody FM, deTineo M, Kaszuba SM, et al. Superiority of an intranasal corticosteroid compared to an oral antihistamine in the as needed treatment of seasonal allergic rhinitis. J Allergy Clin Immunol 2001;107 Suppl:S312.
184. Naclerio RM, Mygind N. Intranasal steroids. In: Naclerio RM, Mygind N, editors. Allergic and nonallergic rhinitis: clinical aspects. Copenhagen: Munksgaard; 1993. p. 114–22.
185. Soderberg-Warner M. Nasal septal perforation associated with topical corticosteroid therapy. J Pediatr 1984;105:840–1.
186. Sahay JN, Ibrahim NB, Chatterjee SS, et al. Long-term study of flunisolide treatment in perennial rhinitis with special reference to nasal mucosal histology and morphology. Clin Allergy 1980;10:451–7.
187. Orgel HA, Meltzer EO, Bierman CW, et al. Intranasal fluorocortin butyl in patients with perennial rhinitis: a 12-month efficacy and safety study including nasal biopsy. J Allergy Clin Immunol 1991;88:257–64.
188. Klemi PJ, Virolainen E, Puhakka H. The effect of intranasal beclomethasone dipropionate on the nasal mucosa. Rhinology 1980;18:19–24.
189. Sorensen H, Mygind N, Pedersen C, Prytz S. Long term treatment of nasal polyps with beclomethasone dipropionate aerosol. III. Morphological studies and conclusions. Acta Otolaryngol (Stockh) 1976;82:260–2.
190. Norman PS, Winkerwerder WL, Agbayon BF, Migeon CJ. Adrenal function during the use of dexamethasone aerosols in the treatment of ragweed hay fever. J Allergy 1967;40:57–61.
191. Frauenfelder F, Myer S. Posterior subcapsular cataracts associated with nasal or inhalational steroids. Am J Ophthalmol 1990;109:489.
192. Cumming RG, Mitchell P, Leeder SR. Use of inhaled corticosteroids and the risk of cataracts. N Engl J Med 1997;337:8–14.
193. Derby L, Maier WC. Risk of cataract among users of intranasal corticosteroids. J Allergy Clin Immunol 2000;105:912–6.
194. Wolthers OD, Pedersen S. Growth of asthmatic children during treatment with budesonide: a double blind trial. BMJ 1991;303:163–5.
195. Agertoft L, Pedersen S. Effect of long-term treatment with inhaled budesonide on adult height in children with asthma. N Engl J Med 2000;343:1064–9.
196. Skoner DP, Rachelefsky GS, Meltzer EO, et al. Detection of growth suppression in children during treatment with intranasal beclomethasone dipropionate. Pediatrics 2000;105:E23.
197. Schenkel EJ, Skoner DP, Bronsky EA, et al. Absence of growth retardation in children with perennial allergic rhinitis after one year of treatment with mometasone furoate aqueous nasal spray. Pediatrics 2000;105:E22.
198. Wiseman LR, Benfield P. Intranasal fluticasone proprionate: a reappraisal of its pharmacology and clinical efficacy in the treatment of rhinitis. Drugs 1997;53:885–907.
199. Onrust SV, Lamb HM. Mometasone furoate: a review of its intranasal use in allergic rhinitis. Drugs 1998;56:725–45.
200. Rhuno J, Denburg J, Dolovich J. Intranasal nedocromil sodium in the treatment of ragweed allergic rhinitis. J Allergy Clin Immunol 1998;81:570.
201. Beswick KB, Kenyon GS, Cherry JR. A comparative study of beclomethasone dipropionate aqueous

nasal spray with terfenadine tablets in seasonal allergic rhinitis. Curr Med Res Opin 1985;9:560–7.
202. Juniper EF, Kline PA, Hargreave FE, Dolovich J. Comparison of beclomethasone dipropionate aqueous nasal spray, astemizole, and the combination in the prophylactic treatment of allergen-induced rhinoconjunctivitis. J Allergy Clin Immunol 1989;83:627–33.
203. Weiner JM, Abramson MJ, Puy RM. Intranasal corticosteroids versus oral H_1 receptor antagonists in allergic rhinitis: systematic review of randomized controlled trials. BMJ 1998;317:1624–9.
204. Juniper EF, Kline PA, Ramsdale EH, Hargreave FE. Comparison of efficacy and side effects of aqueous steroid nasal spray (budesonide) and allergen-injection therapy (Pollinex-R) in the treatment of seasonal allergic rhinoconjunctivitis. J Allergy Clin Immunol 1990;85:606–11.
205. Mygind N. Glucocorticosteroids and rhinitis. Allergy 1993;48:476–90.
206. Borum P, Gronberg H, Mygind N. Seasonal allergic rhinitis and depot injection of corticosteroid. Evaluation of the efficacy of medication early and late in the season based on detailed symptom recording. Allergy 1987;42:26–32.
207. Mabry RL. Intranasal corticosteroid injection: indications, technique and complications. Otolaryngol Head Neck Surg 1979;87:207–11.
208. Orie N, Booij-Nord H, Pelikan Z. Protective effect of disodium cromoglycate on nasal and bronchial reactions after allergen challenge. In: Pepys J, Frankland A, editors. Disodium cromoglycate in airways disease. London: Butterworth; 1970. p. 33–41.
209. Lindsay-Miller A, Chambers A. Group comparative trial of cromolyn sodium and terfenadine in the treatment of seasonal allergic rhinitis. Ann Allergy 1987;58:28–32.
210. Coffman DA. A controlled trial of disodium cromoglycate in seasonal allergic rhinitis. Br J Clin Pract 1971;25:403–6.
211. Borum P, Larsen FS, Mygind N. Nasal methacholine provocation and ipratropium therapy of perennial rhinitis. Acta Otolaryngol Suppl (Stockh) 1979;360:35–9.
212. Mygind N, Borum P. Anticholinergic treatment of watery rhinorrhea. Am J Rhinol 1990;4:1.
213. Østberg B, Winther B, Mygind N. Cold air-induced rhinorrhea and high-dose ipratropium. Arch Otolaryngol Head Neck Surg 1987;113:160–2.
214. Donnelly AL, Glass M, Minkwitz MC, Casale TB. ICI 204,219 relieves the symptoms of acute seasonal allergic rhinitis. Am J Respir Crit Care Med 1995;151:1734–9.
215. Malmstrom K, Meltzer E, Prenner B, et al. Effects of montelukast + loratadine and placebo in seasonal allergic rhinitis and conjunctivitis [abstract]. J Allergy Clin Immunol 1999;103:A406.
216. Meltzer EO, Malmstrom K, Lu S, et al. Concomitant montelukast and loratadine as treatment for seasonal allergic rhinitis: a randomized, placebo-controlled clinical trial. J Allergy Clin Immunol 2000; 105:917–22.
217. Bielory L. Allergic and immunologic disorders of the eye. Part II: ocular allergy. J Allergy Clin Immunol 2000;106:1019–32.
218. Azevedo M, Castel-Branco MG, Oliveira JF, et al. Double-blind comparison of levocabastine eye drops with sodium cromoglycolate and placebo in the treatment of seasonal allergic conjunctivitis. Clin Exp Allergy 1991;21:689–94.
219. Friedlander MH, Harris J, LaVallee N, et al. Evaluation of the onset and duration of effect of azelastine (0.05%) versus placebo in patients with allergic conjunctivitis using an allergen challenge model. Ophthalmology 2000;107:2152–7.
220. Netland PA, Leahy C, Krenzer KL. Emedastine ophthalmic solution 0.05% versus levocabastine ophthalmic suspension 0.05% in the treatment of allergic conjunctivitis using the conjunctival allergen challenge model. Am J Ophthalmol 2000;130: 717–23.
221. Abelson MB. Evaluation of olopatadine, a new ophthalmic antiallergic agent with dual activity, using the conjunctival allergen challenge model. Ann Allergy Asthma Immunol 1998;81:211–8.
222. Tinkelman DG, Rupp G, Jaufman H, et al. Double-masked, paired comparison clinical study of ketorolac tromethamine 0.5% ophthalmic solution compared with placebo eye drops in the treatment of seasonal allergic conjunctivitis. Surv Ophthalmol 1993;38:133–40.
223. Creticos PS, Marsh DG, Proud D, et al. Responses to ragweed-pollen nasal challenge before and after immunotherapy. J Allergy Clin Immunol 1989;84: 197–205.
224. van Neerven RJ, Wikborg T, Lund G, et al. Blocking antibodies induced by specific allergy vaccination prevent the activation of CD4+ T cells by inhibiting serum-IgE-facilitated allergen presentation. J Immunol 1999;163:2944–52.
225. Bousquet J, Becker WM, Hejjaoui A, et al. Differences in clinical and immunologic reactivity of patients allergic to grass pollens and to multiple pollen species. II. Efficacy of a double-blind,

CHAPTER 33

Acute and Chronic Nasal Disorders

Valerie J. Lund, MS, FRCS, FRCS(Ed)

A wide range of acute and chronic conditions may affect the nose, either arising locally or occurring as part of systemic disease. These may be broadly divided into allergic and nonallergic, with the latter divided into infectious and noninfectious (Table 33–1). This chapter considers some of the infectious conditions, the nonallergic, noninfectious diseases, and a number of other conditions that enter into the differential diagnosis of patients presenting with nasal symptoms. From an anatomic and physiologic perspective, the nose is rarely affected in isolation, and the majority of conditions impact on the adjacent paranasal sinuses to a greater or lesser extent. Thus, there has been agreement on the term "rhinosinusitis" in a number of international consensus documents[1,2] as primarily relating to infection although, in reality, probably applicable to most disease processes in this region. Similarly, the upper and lower respiratory tract should be considered as one organ, with many conditions manifesting themselves in both areas, sometimes in subtly different ways (eg, asthma and nasal polyps).

TABLE 33–1. Classification of Nose and Sinus Disorders

Allergic
Nonallergic
Infectious
Viral
Bacterial
Fungal
Parasitic
Protozoal
Noninfectious
Idiopathic
Occupational/environmental
Hormonal
Drug induced
Food
Emotional
Atrophic
Nonallergic rhinitis with eosinophilia
Gastroesophageal reflux
Differential diagnosis
Polyps
Mechanical factors
Tumors
Granulomas
Cerebrospinal rhinorrhea

Adapted from Lund VJ et al.[1]

INVESTIGATION

A considerable number of tests and procedures are now available for the investigation of sinonasal disease (Table 33–2). However, not all are routinely available or indeed applicable to each individual patient, and their selection will depend on clinical indicators and local factors.

HISTORY AND EXAMINATION

In practical terms, the nose has a limited repertoire of responses in any given condition, that is, nasal congestion, discharge, sneezing, itching, and epistaxis with associated symptoms of headache and/or facial pain. However, a careful history will usually suggest the diagnosis. A thorough general medical history should be followed by questions specific to rhinologic symptoms including information on environmental and occupational factors and family history. The frequency, duration, and severity of symptoms should also be discussed, and a visual analog score can be used to semiquantify severity by simply asking the patients to mark on a 10 cm line where they feel they lie for a particular symptom during the last week.[3] This method can be repeated on each subsequent visit, with the patients acting as

TABLE 33–2. Diagnostic Techniques

History
General otolaryngologic examination
Allergy tests
Skin tests
Total serum IgE
Serum-specific IgE
Endoscopy
Rigid
Flexible
Nasal smear
Cytology
Nasal swab
Bacteriology
Radiology
Plain sinus radiography
Computed tomography
Magnetic resonance imaging
Chest radiography
Mucociliary function
Nasomucociliary clearance
Ciliary beat frequency
Electron microscopy
Nitric oxide measurement
Nasal airway assessment
Nasal inspiratory peak flow
Rhinomanometry (anterior and posterior)
Acoustic rhinometry
Olfaction
Threshold testing
"Scratch and sniff" tests
Blood tests
Full blood count and white cell differential
Erythrocyte sedimentation rate
Thyroid function tests
Antineutrophil cytoplasmic antibody
Immunoglobulins and IgG subclasses
Antibody response to immunization with protein and carbohydrate antigens

Adapted from Lund VJ et al.[1]

their own controls. Consistent obstruction on the same side suggests a polyp, structural problem, or, more rarely, a tumor or even a congenital unilateral choanal atresia. Hyposmia and anosmia are most often associated with nasal polyps or more severe disease such as Wegener's granulomatosis or sarcoidosis. Secondary symptoms related to blockage of the airways include frequent sore throats, dryness of the mouth and oropharynx, a nasal quality to the voice, and snoring. A full drug history may also give a clue to the source of the problem. Symptoms specific to particular conditions are discussed in the relevant sections.

Examination of the nose should be performed in all cases of rhinologic complaint and should include anterior rhinoscopy using a nasal speculum and head light and rigid endoscopy combined with flexible nasoendoscopy in selected cases to ensure a full evaluation of the nasal cavity and nasopharynx.

The quantity and quality of secretion should be noted, both the viscosity and color, and specimens may be taken for a nasal smear. Similarly, under endoscopic control, a swab may be taken from the middle meatus, which provides a reasonable correlation with the bacteriology of the dependent sinuses.[4]

The appearance of the mucosa is rarely altered in a way that is pathognomonic for a particular disease. However, it is usually reddened in acute infections and overuse of topical medications, whereas a typical allergic mucosa appears pale and swollen. In granulomatous conditions such as Wegener's and sarcoidosis, the mucosa is generally reddened, friable, and associated with crusting and evidence of previous epistaxis. A careful examination of the nasal cavity should reveal polyps, tumor, foreign bodies, or septal deflections, although it is important that the unwary examine beyond the deflection wherever possible as secondary pathology may be present and go unnoticed. The presence of a septal perforation should raise the possibility of cocaine abuse, previous surgery, or one of the systemic granulomatous diseases, although most often it is "idiopathic," resulting from minor or repeated trauma.

Physical examination should not be confined to the nose alone and should be accompanied by a full otolaryngoloic examination including posterior rhinoscopy, indirect laryngoscopy, and palpation of the neck in selected cases. It should also be remembered that many conditions affecting the nose and sinuses also affect the lower respiratory tract, and this may be revealed by appropriate examination.

(For allergy testing, see Chapter 32, and for imaging, see Chapter 31.)

MUCOCILIARY FUNCTION

Nasomucociliary Clearance

A simple test of mucociliary function can be performed by placing a 0.5 mm piece of saccharin on the anterior end of the inferior turbinate approximately 1 cm from the end to avoid areas of squamous metaplasia. The time taken to taste something sweet in the mouth is measured,[5] which normally takes 30 minutes or less. If longer than an hour has elapsed, it is worth repeating the test in case the particle has fallen out and checking that the patient is capable of tasting saccharin.

Ciliary Beat Frequency

When the saccharin test is prolonged or if specific ciliary abnormalities are suspected, it is possible to examine the cilia directly by taking a sample with a small disposable, cupped spatula (Rhinoprobe) and observing cilia activity under a phase-contrast microscope with a photometric cell.[6] The frequency can be measured with a real-time analyzer and expressed in hertz, the normal range from the inferior turbinate being 12 to 15 Hz. However, this technique is not available in every center.

Electron Microscopy

If the nasomucociliary clearance time and ciliary beat frequency are abnormal, samples may be obtained with the spatula or via direct biopsy for electron microscopy studies to diagnose conditions such as primary ciliary dyskinesia (PCD).

Nitric Oxide Measurement

The level in parts per million of nitric oxide present in the expired nasal and pulmonary air can be helpful in establishing normal mucociliary function.[7] In PCD, the nitric oxide level that is an indirect marker of ciliary metabolism is much reduced to double figures or less in the nose and single figures from the chest. However, as nitric oxide is predominantly produced in the sinuses, conditions such as nasal polyposis that obstruct gas exchange from these areas can also result in a very low reading from the nasal gases. The level is commensurately elevated in the presence of inflammation.

NASAL AIRWAY ASSESSMENT

Nasal Expiratory or Inspiratory Peak Flow

Nasal expiratory or inspiratory peak flow, which uses a peak-flow meter, has the advantage of being inexpensive, quick, and easy to perform. It is useful for repeated examinations and compares well with active anterior rhinomanometry.[8] Of the two methods, forced inspiration is preferred, although it can produce significant vestibular collapse in some individuals. Forced expiration inflates the eustachian tube, which may be uncomfortable and may also produce an unpleasant quantity of mucus in the mask.

Rhinomanometry

Rhinomanometry attempts to measure nasal airway resistance by making quantitative measurement of nasal flow and pressure.[9] It employs the principle that air will flow through a tube only when there is a pressure differential passing from areas of high to low pressure. When the nasal mucosa is decongested, the reproducibility of rhinomanometry is good, but it requires some expertise to produce consistent results and has therefore remained primarily a research tool. Active anterior rhinomanometry is more commonly used as the posterior technique cannot be used in 20 to 25% of individuals owing to an inability to relax the soft palate.

Acoustic Rhinometry

In acoustic rhinometry, an audible sound pulse is electronically generated and is passed into the nose, where it is altered by variations in the cross-sectional area. The reflected signal is picked up by a microphone and analyzed, allowing determination of area within the nasal cavity as a function of distance. From this, volumes may be derived, thus providing topographic information rather than a measure of airflow.[10] It appears to be more reproducible and to have greater applicability than rhinomanometry but is still being evaluated as a clinical tool.

OLFACTION

Olfactory Thresholds

Estimation of olfactory thresholds may be established by presentation of serial dilutions of pure odorants such as carbinol.[11] The patient is presented with two bottles, one containing only the diluent solution as the control and the other the odorant in progressively increasing or decreasing concentrations. Each is sniffed in turn until a point is reached when the patient cannot distinguish between the control and test bottle. This indicates the minimum detectable odor.

Scratch and Sniff Tests

The University of Pennsylvania Smell Identification Test (UPSIT) uses patches impregnated with microencapsulated odorants.[12] The patient is forced to choose among a number of options after scratching the patch to release the odor. The results are well validated for age and sex and take into account answers guessed correctly as well as give an indication of malingering.

Alternative identification tests include the Zurich Smell Test, in which eight odorants must be correctly identified. The test choices are offered pictorially and in English.[13] The Sniffin' Sticks combine both a qualitative and a quantitative assessment of olfaction.[14]

Blood Tests

A wide range of hematologic investigations may be appropriate given the suspected differential diagnosis (eg, thyroid dysfunction, Wegener's granulomatosis, sarcoidosis, and immune deficiency).

INFECTIOUS RHINITIS

The separation of infections in the nose from those in the paranasal sinuses is now recognized as incorrect, and it is generally accepted that the term "rhinosinusitis" better encompasses infections in this region, although one area may be affected to a greater or lesser extent. This has been supported by the work of Gwaltney et al,[15] who have demonstrated with computed tomographic scanning that during a simple uncomplicated viral cold, the majority of subjects have some involvement of the paranasal sinuses. A long list of causative agents can produce infection in this region (Table 33–3).

Etiologic Factors

A number of congenital conditions may predispose patients to infection of the respiratory tract.

Primary Ciliary Dyskinesia Some individuals are born with a congenital abnormality of the cilia that affects their motility. Electron microscopic examination reveals disorganized microtubules and the absence of dynein arms. In addition to nasal problems, patients usually exhibit lower respiratory tract infection, serous otitis media, and infertility problems. In full-blown Kartagener's syndrome, patients have bronchiectasis, sinus infection, and situs inversus in approximately 50% of cases.

Cystic Fibrosis Cystic fibrosis is an autosomal recessive disease, the most common inherited fatal disease of white children, caused by a defective mucosal chloride transport gene. In its most severe form, it presents with malabsorption and progressive obstructive pulmonary disease, but approximately one-third of children and one half of adults suffer from multiple bilateral nasal polyps, and the majority of patients have a chronic pansinusitis (Figure 33–1). The polyps have been attributed to an abnormality of sodium and water transport across membranes, and the histology of the polyps differs significantly from that in other conditions in that they have less eosinophils and more lymphocytes and plasma cells. *Pseudomonas aeruginosa* is commonly found in association with the sinus disease and may be targeted specifically with appropriate antibiotics such as the quinolones. The majority of patients are managed by a combination of medical therapies with intranasal corticosteroids and surgery. The results of surgery are compromised by the condition although are still regarded as worthwhile by patients. Other medical therapies have included intranasal furosemide, amiloride, and messenger ribonuclease.

Viral Infections

The most obvious and frequently occurring example of infection is the common cold. A number of viruses are associated with upper respiratory tract infections of varying degrees of severity including over 100 different types of rhinoviruses. This, combined with the

TABLE 33–3. Infective Rhinitis: Causative Agents

Viruses	
Common cold	Picornavirus
	Rhinovirus
	Coxsackievirus
	Reovirus
	Echovirus
	Parainfluenza
	Adenovirus
	Respiratory syncytial virus
Influenza	Influenza virus
Mucocutaneous herpes	Herpes simplex
Chickenpox, herpes zoster	Varicella zoster
Bacteria	
Furuncle	*Staphylococcus aureus*
Secondary infection	Various aerobes, anaerobes, bacteroides
Actinomycosis	*Actinomyces israelii*
Diphtheria	*Corynebacterium diphtheriae*
Rhinoscleroma	*Klebsiella rhinoscleromatis*
Leprosy	*Mycobacterium leprae*
Glanders	*Pseudomonas mallei*
Tuberculosis	*Mycobacterium tuberculosis*
Syphilis	*Treponema pallidum*
Fungi	
Aspergillosis	*Aspergillus* sp
Zygomycosis	
Subcutaneous	*Conidiobolus coronatus*
Rhinocerebral	*Mucor* or *Rhizopus oryzae*
Blastomycosis	*Blastomyces dermatidis*
Cryptococcosis	*Cryptococcus neoformans*
Histoplasmosis	*Histoplasma capsulatum*
Sporotrichosis	*Sporothrix schenkii*
Candidiasis	*Candida*
Protozoa	
Leishmaniasis	*Leishmania* sp
Parasites	
Myiasis	*Chrysomyia*
Rhinosporidiosis	*Rhinosporidium seeberi*

ability of viruses for mutation, has so far confounded attempts at prophylactic immunization. Most young adults suffer from two to three colds each year, and it is estimated that 0.5 to 2% of these become bacterially infected. Predisposing factors such as climatic and environmental changes, fatigue, and stress have all been implicated, but the huge variety of viruses makes it unlikely that there is preexisting immunity, and, as a consequence, most individuals are susceptible most of the time. The virus is normally transported into the nose by direct contact with the fingers rather than by airborne contamination, and once the virus gains access to the respiratory epithelium, it will produce ciliary stasis and damage to the cilia. The

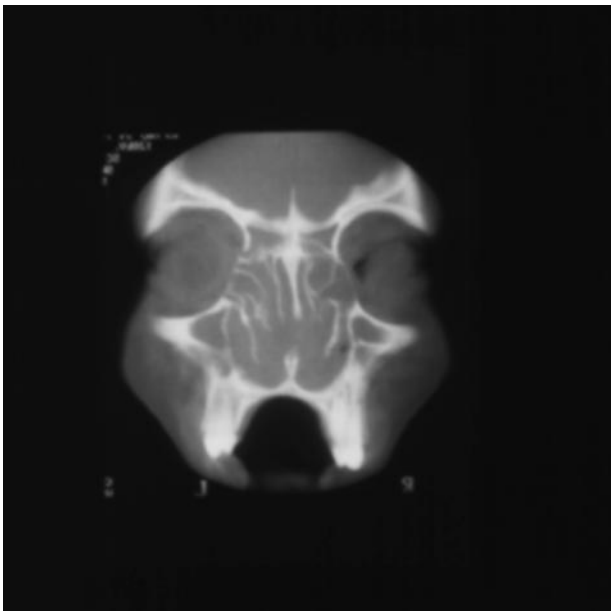


FIGURE 33–1. Coronal computed tomographic scan showing massive nasal polyposis with expansion of the nasal cavity and ethmoid complexes in a patient with cystic fibrosis.

condition is characterized by edema, increased secretion, and desquamation of respiratory epithelium. Clinically, after 1 to 3 days incubation, a prodromal or "dry" phase follows. The subject feels unwell and shivery with a headache, muscle aches, and loss of appetite, while the nose feels irritated and sneezing occurs. The catarrhal phase follows after a few hours accompanied by profuse watery secretion, nasal obstruction, a reduced sense of smell, and an increase in constitutional symptoms. The infection then enters a mucous phase in which symptoms improve or may continue after 5 days owing to secondary bacterial infection with a progressively mucopurulent discharge. Other areas of the upper respiratory tract may be affected including the pharynx, middle ear, and tonsils.

From a practical point of view, little can be done other than symptomatic relief with decongestants and antipyretics, although recent studies have suggested that a combination of oral chlorpheniramine and intranasal ipratropium bromide may be effective in reducing symptoms if taken early in the prodromal phase. Antibiotics should be reserved for secondary bacterial infection.

Influenza, in its many forms, constitutes a more specific and potentially serious viral upper respiratory tract infection, resulting in pandemics with significant morbidity and mortality in susceptible populations. Damage to the ciliated epithelium is more profound, facilitating secondary bacterial colonization, and the olfactory epithelium may be permanently damaged, leading to anosmia. Vaccines have been developed and are increasingly used as prophylaxis in vulnerable populations, but their usefulness is again limited by the potential for viral mutation.

Bacterial Infections

Folliculitis and Vestibulitis Specific bacterial infections can occur in the nasal vestibule when *Staphylococcus aureus* invades a pilosebaceous follicle to produce an extremely painful, indurated area. This will spontaneously burst and discharge its purulent contents after 4 to 5 days. The condition is extemely painful as the vestibular skin is tightly bound to the underlying cartilages, providing little room for swelling. Local cleaning and systemic antibiotics, which should include floxacillin, should be given, taking into account problems of local bacterial resistance. If the abscess is localized and fluctuant, it may be incised, but the condition should always be treated seriously as it is associated with the potential complication of cavernous sinus thrombosis by retrograde spread of infection along the angular and superior ophthalmic veins.

The cavernous sinus lies on either side of the sphenoid bone. It is broken up by trabeculae into many venous cavernous spaces through which the third, fourth, and first two divisions of the fifth cranial nerve, sixth cranial nerve, and internal carotid artery pass. It is connected by an extensive valveless venous system to the nose, adjacent face, nasopharynx, pharynx, orbit, and paranasal sinuses, allowing retrograde spread of infection from any of these areas. The condition carries a serious morbidity, often with bilateral blindness and mortality. It is a complication predominantly of the young, with two-thirds being under 20 years.[16] The patient will complain of headache and painful paresthesia in the distribution of the trigeminal nerve, following which other cranial nerves related to the cavernous sinus become involved, affecting extraocular movement and resulting in ophthalmoplegia. The sudden development of bilateral orbital signs should alert the clinician to this complication. Prior to antibiotics, cavernous sinus thrombosis carried a 50% mortality, which nevertheless still stands at a significant 10 to 27%.[17] Computed tomography and

magnetic resonance imaging confirm the diagnosis, and high-dose broad-spectrum systemic antibiotic therapy should be given, together with anticoagulation in selected cases.

The increase of asymptomatic carriage of methicillin-resistant *S. aureus* (MRSA) in the nasal vestibule is a source of considerable concern as it is implicated in nosocomial infection, with serious consequences for debilitated patients. Despite vigorous eradication programs, the incidence of MRSA appears to be increasing worldwide.

A mild form of nonspecific infective vestibulitis is often encountered, producing mild irritation, small fissures, or pustules. The application of white petroleum jelly or antibiotic ointments such as chlortetracycline hydrochloride 3% may be used.

Erysipelas and Other Specific Infections Erysipelas is an acute beta-hemolytic streptococcus infection of the skin and soft tissue of the head and neck following a cut or surgical incision. The onset is usually abrupt, with fever, redness, tenderness, and induration of the involved skin. On the face, the infection frequently displays a butterfly pattern across the nose, adjacent cheeks, and upper lip. The infection generally responds well to penicillin or erythromycin and should be distinguished from cellulitis, which is a more generalized deep infection of the skin and subcutaneous tissues also caused by *Staphylococcus* or *Streptococcus*.

A variety of other specific infections may occur in this region including diphtheria, rhinosceleroma, leprosy, tuberculosis, and syphilis.

In diphtheria, caused by *Corynebacterium diphtheriae,* the nose may be affected primarily or secondarily to infections of the oropharynx. The false membrane associated with the condition is made of fibrin, neutrophils, necrotic epithelium, and the bacillae. Removal of the membrane leaves a bleeding surface. Treatment is with parenteral penicillin, with the antitoxin reserved for acute cases. A carrier state can occur after contact, and the individual may need to be isolated until nose swabs are negative.

In rhinoscleroma, a chronic inflammatory condition is characterized by large, deforming masses distending the nasal cavity. The condition, ascribed to *Klebsiella rhinoscleromatis*, was once found in Central and Eastern Europe but now occurs in Central and South America, Africa, the Middle East, and India, where it is associated with a poor standard of health and living conditions. The disease classically has three phases: rhinitic, infiltrative, and nodular, finally resulting in adhesions, stenosis, and atresia. The large red tumor-like masses of the nodular phase are characterized histologically by the presence of Mikulicz's cells, large cells with clear cytoplasm containing the bacilli, and by Russell's bodies, which are plasma cells with eosinophilic staining cytoplasm and prominent nuclei. Large doses of streptomycin and tetracycline over 4 to 6 weeks are needed until two consecutive biopsies are negative. Topical acriflavine solution has also been used.

In leprosy caused by the acid-fast *Mycobacterium leprae*, two forms in the nose are recognized: tuberculoid and lepromatous. The microorganism may enter via the nose, transferred by a fingernail; incubation can take up to 10 years. In tuberculoid leprosy, solitary lesions occur as small anesthetic patches, whereas in lepromatous leprosy, a diffuse infiltration of the skin, mucous membranes, and nerves is found, producing nodular thickening particularly in the region of the anterior inferior turbinate. The septum is progressively involved, leading to perforation of the cartilage and a typical nasal deformity of collapse and stenosis. Dapsone is still the treatment most generally used worldwide, although increasing resistance has led to newer agents and combinations with rifampin and clofazimine.

In tuberculosis, the nose may be involved secondary to pulmonary involvement, with nodular or ulcerative lesions usually found on the anterior part of the cartilaginous septum, on the inferior turbinate, or in the choanae. The lesions show the usual appearances of caseating epithelioid granulomas together with the presence of acid-alcohol-fast bacilli. The condition responds to standard systemic antituberculosis therapy together with topical saline douching. Lupus vulgaris is a chronic indolent form of tuberculosis, producing the classic "apple jelly" nodules at the mucocutaneous junction of the vestibule.

In syphilis caused by *Treponema pallidum*, the nose may be affected in the primary, secondary, or tertiary phases as well as in its congenital form. A primary chancre can occur on the external nose or vestibule around 3 to 4 weeks after contact and then disappears spontaneously within 6 to 10 weeks. Smears taken at this time may reveal the *Treponema*, although the serology can be negative in the early stages. Secondary syphilis appears around 6 to 10

weeks after infection and may produce no more than simple catarrhal rhinitis, although a serologic diagnosis can be made during this phase. Tertiary syphilis is associated with a gumma that appears as a firm, reddened nodule invading the mucous membrane, periosteum, and bone and often producing tenderness over the bridge of the nose, posterior septal perforations, and nasal collapse. Serology is positive in 90% of these patients. Congenital syphilis may manifest itself with "snuffles" a few weeks after birth or as a latent form occurring at puberty, with the characteristic nasal saddling, dental abnormalities, and sensorineural hearing loss. The condition must be distinguished from other granulomatous diseases and should be treated with standard antibiotics such as penicillin. For more information on infections, see Chapter 30.

NONALLERGIC, NONINFECTIOUS RHINITIS

Idiopathic (Vasomotor) Rhinitis

The term "idiopathic rhinitis" relates to any inflammation of the nose the cause of which is unknown. It is the diagnosis of exclusion as underlying causative or contributory factors can often be found if carefully sought. Patients are generally manifesting an upper respiratory hyperresponsiveness or reactivity, which is simply an exaggeration of normal defense mechanisms to nonspecific environmental triggers such as changes in temperature and humidity or exposure to irritants (eg, cigarette smoke or strong odors). An extreme example of a normal physiologic response can be seen in "skier's nose" in reaction to cold, dry air.[18] The commonly used term "vasomotor rhinitis" is less satisfactory as it suggests a known pathophysiologic mechanism for the condition that is far from proven.[1]

Cigarette smoke is known to affect mucociliary clearance[19] and has been shown to cause an eosinophilic inflammation in the nasal mucosa of nonatopic children.[20] Challenge of smoke-sensitive individuals produces rhinorrhea and obstruction, and in smokers, eye irritation and hyposmia are more common than in nonsmokers. It has also been shown that the more people smoke, the more they experience the symptoms of chronic rhinitis.[21]

In nonallergic, noninfectious rhinitis, symptoms vary in intensity but consist of nasal blockage, anterior rhinorrhea, and postnasal discharge, sometimes with sneezing. There is no characteristic seasonal variation, and ocular symptoms are almost always missing. Skin prick tests are negative or do not correlate with the symptoms (ie, may be positive to grass pollen).

Therapy may encompass avoidance of obvious triggers, but most patients require medication, the mainstays of which are topical corticosteroids that are safe for long-term use. If watery rhinorrhea is the most significant or only symptom, a topical anticholinergic agent such as ipratropium bromide may be helpful. Topical decongestants are best avoided because of their side effects in the long term (sic), and although oral decongestants are popular, their effect is variable and not without cardiovascular and central nervous system effects. Topical capsaicin has been investigated for intractable nonallergic rhinitis[22] as it produces a short-term neuronal defunctionalization, but the effects are short-lived and are accompanied by an unpleasant burning sensation, so its role remains undetermined.

Surgery has a limited role to play. Vidian neurectomy has had a vogue but has largely been abandoned as the effects are unpredictable and at best short-lived. Correction of septal deformity and/or turbinate hypertrophy is occasionally indicated if medication fails.

Occupational/Environmental Rhinitis

Airborne agents present in the workplace may affect the upper respiratory tract, producing either an allergic reaction or a nonallergic hyperresponsiveness. Chemicals such as acid anhydrides, platinum salts, glues, and solvents, as well as dusts from grain and wood, may act as irritants.[23] Considerable controversy now surrounds the adverse effects of environmental pollutants such as ozone, sulfur dioxide, nitric dioxide, particulate matter, volatile organic compounds, and formaldehyde. Some contradictory evidence needs to be clarified by carefully constructed epidemiologic studies that may ultimately support present concerns.

A careful history will usually indicate the cause, confirmed by improvement once the patient is removed from the suspected irritant. However, there may be a reluctance to complain if a worker's job is jeopardized, and chronic exposure may produce

chronic symptoms, which do not abate for a week or more after exposure ceases.

The nasal symptoms of blockage and watery discharge may be accompanied by occupational asthma, conjunctivitis, and dermatitis. Clinical findings are usually nonspecific, and confirmation may require direct nasal challenge evaluated with symptom scores and assessment of nasal patency. Carefully graded exposure to increasing doses of the suspected agent in special ventilated environmental chambers may be necessary, especially if there are medicolegal consequences of the diagnosis.

Separation of patient and irritant, either by change of job, improved ventilation, or use of masks, may suffice complemented by conventional pharmacotherapy (eg, topical corticosteroids).

Hormonal Rhinitis

A range of physiologic and pathologic endocrine conditions can affect the nose. Changes are known to occur in the nose during puberty, the menstrual cycle, and pregnancy,[24] with some women experiencing significant blockage, watery discharge, and, sometimes, severe epistaxis. The symptoms are worst during the last trimester of pregnancy, relating to the blood estrogen level, and resolve with delivery. In women with preexisting perennial rhinitis, the nasal symptoms may improve or deteriorate during pregnancy. Hormonal change may also be responsible for the mucosal atrophy that occurs after the menopause.

In specific endocrine disorders such as hypothyroidism[25] and acromegaly, nasal stuffiness and discharge may occur.

From a therapeutic perspective, if the symptoms during pregnancy are sufficiently debilitating, medication must be considered, always taking the risk/benefit ratio into consideration. In the first instance, saline douche may be of benefit, but, if ineffective, topical nasal corticosteroids have been used for many years without any evidence of a teratogenic or other adverse effects. Certainly, extensive studies in asthmatics using inhaled corticosteroids have failed to show any abnormalities.[26] Oral decongestants theoretically could cause vascular disturbance in the placenta and fetus, but oral pseudoephedrine given within the recommended dosage is approved for use in pregnancy and has been widely prescribed in the United States.

Drug-Induced Rhinitis

A number of medications are known to produce nasal symptoms. These include

- aspirin and other nonsteroidal anti-inflammatory agents (NSAIDs). Aspirin intolerance typically increases nasal congestion and secretion and is associated with nasal polyposis and nonallergic asthma. Avoidance of aspirin and NSAIDs is clearly important, although there is some controversy as to whether avoidance of salicylate-like substances in the diet is beneficial or not. Some have argued that complete avoidance may render individuals more susceptible to inadvertent challenge, whereas others believe a completely salicylate-free diet to be worse than the disease. Other strategies have included salicylate desensitization, which must be performed with great care in a hospital to manage any severe respiratory reactions.
- cardiovascular preparations such as reserpine,[27] guanethidine,[28] phentolamine, methyldopa, angiotensin-converting enzyme (ACE) inhibitors,[29] and α-adrenoceptor antagonists, which are all known to affect the nose, typically producing nasal stuffiness.
- topical ophthalmic beta blockers.
- chlorpromazine.
- oral contraceptives, again by hormonal modulation.
- cocaine, which is more often used recreationally than medically and is a significant irritant, producing frequent rhinorrhea and resultant sniffing, hyposmia, and septal perforation.[30,31] In extreme cases, a midline destruction of the nose may result, which has proved difficult to distinguish from T-cell lymphoma or Wegener's granulomatosis.

Rhinitis medicamentosa is a well-described condition classically resulting from long-term abuse of topical decongestants.[32,33] The patient initially suffers a rebound congestion, which leads to further use of the decongestant, and ultimately develops significant mucosal atrophy. Consequently, these preparations are not recommended for longer than 7 to 10 days. Once the condition is established, it can prove difficult to wean a patient off the medication. The reason for originally starting the drug should be considered, and in mild cases, it may be possible to substitute a topical nasal corticosteroid.

However, if there are no contraindications, an initial short course of oral prednisolone (30 mg/day for 5 days) may be necessary.[1]

Food-Induced Rhinitis

Food can produce nasal symptoms in a variety of ways. Although allergy to food is a rare cause of isolated rhinitis, hypersensitivity may result in reaction to the foods themselves or to colorants and preservatives.[34] Alcoholic beverages can produce physiologic vasodilatation with associated nasal blockage as well as allergic and nonallergic reactions to their many components.

Gustatory rhinorrhea may result from spicy hot foods, probably owing to the capsaicin content of red pepper. This substance is known to stimulate sensory nerves, inducing release of neuropeptides such as tachykinins.[22]

Emotion

Sexual arousal and stress are known to have an effect on the nose—hence "honeymoon rhinitis," probably owing to autonomic stimulation, although therapeutic intervention is rarely sought.

Atrophic Rhinitis

Primary atrophic rhinitis has been attributed to infection, most notably to *Klebsiella ozaenae*[35] (from the word "ozaena" meaning "stench"), but may also relate to environmental factors and general health. The condition is characterized by progressive atrophy of the mucosa with loss of the turbinate bone and resulting in a capacious cavity full of foul-smelling crust. This should be distinguished from secondary atrophy following excessive surgery, trauma, radiotherapy, and chronic granulomatous conditions.

Primary atrophic rhinitis has become less common in countries where social conditions and health have generally improved. It affects both sides of the nose, occurs after puberty, and is more common in women. Because of this, an endocrine imbalance has been postulated as a cause, whereas others believe it has an autoimmune basis, possibly initiated by a virus or owing to vitamin or iron deficiencies.

The most unpleasant symptom is the foul smell, of which the patient is often unaware. The copious crusts often bleed when they detach and may extend into the nasopharynx, producing an unpleasant choking sensation and snorting. The nose paradoxically feels blocked, owing to the drying effect in the abnormally patent airway.

Having eliminated other possible underlying conditions by appropriate hematology and imaging, local treatment with saline or alkaline douching, emollients and lubricants such as 25% glucose and glycerine drops, and regular decrusting may suffice. Antibiotics directed at *Klebsiella* are rarely successful in the long term, although the fluoroquinolones and metronidazole have been used. A wide range of surgical procedures have been suggested, largely aimed at reduction of the size of the cavity, all with limited success. Closure of the nostrils with small skin flaps, as advocated by Young,[36] can be helpful, but relapse often occurs on reopening. It was suggested that reduction of airflow inhibited growth of the microorganisms, but this is unproven.

Nonallergic Rhinitis with Eosinophilia/ Eosinophilic Rhinitis Syndrome

Nonallergic rhinitis with eosinophilia syndrome (NARES) was described 20 years ago, although its existence as a separate entity is disputed. It should probably be regarded as a subgroup of idiopathic rhinitis characterized by nasal eosinophilia and perennial symptoms of sneezing, itching, rhinorrhea, blockage, and sometimes hyposmia in the absence of demonstrable allergy. It is possible that it represents an early stage of aspirin sensitivity and may be associated with asthma and nasal polyposis. Intranasal corticosteroids are the mainstay of medical treatment combined with surgery for the polyps.[37]

Gastroesophageal Reflux

Reflux is increasingly thought to play a role in nonallergic rhinitis, particularly in children.[38]

MECHANICAL NASAL OBSTRUCTION

A common nasal complaint is one of nasal blockage, congestion, or obstruction. This may arise for a variety of reasons, be it a genuine mechanical obstruction to airflow, a response to autonomic or sensory changes in the nose, inflammation within the osteomeatal complex, allergic changes, or even a per-

ception paradoxically in the presence of a patent or overpatent airway. It is important to investigate this complaint carefully to elucidate the cause since in the past there has been a tendency to offer surgery for the septum and/or turbinates without consideration or treatment of other contributory factors, inevitably compromising a successful symptomatic outcome.

Septal Deviation

Causes Septal deformity can be congenital or acquired, although it should be recognized that a completely straight septum is the exception rather than the rule. This may be attributable to prenatal or perinatal factors, in utero or during delivery. A true dislocation of the septum is rare; more often, a horizontal fracture above the maxillary crest occurs, which can be relocated immediately. Small microfractures may go unnoticed, presenting in later life, or be exacerbated by falls during the "toddler" stage that may go unnoticed. In adolescents and young adults, sports injuries, assault, and traffic accidents may all have their impact.

Effects Any part of the nasal internal and external structure may be damaged, leading to deflections and deformities of the cartilages and bones. The effects may not be immediately obvious to the individual until infection or allergy intervenes. With age, loss of tensile strength in collagen may lead to collapse of the nasal valve region, uncovering preexisting nasal asymmetry. The sensation of nasal obstruction may alternate with the nasal cycle or be apparent when lying on that side. Disordered airflow may result in crusting and epistaxis on the deflected area. "As goes the septum, so goes the nose" correctly indicates that septal deviation can be associated with cosmetic deformity of the bridge and columella. (For nasal reconstruction, see Chapter 38.)

Assessment of the external and internal nose, including anterior rhinoscopy and rigid endoscopy, is important to assess the septal position, spurs, and turbinates and exclude other pathology, particularly in the middle meatus and nasopharynx. Instillation of a topical decongestant may help determine the degree of mucosal swelling, particularly of the anterior part of the nasal septum and inferior turbinates. Where available, objective measures of airway can be helpful.

Treatment Septal manipulation is only of value in neonates when trauma has occurred during delivery. Septoplasty or submucous resection have been described to correct septal deviation. The distinction between these two techniques relates primarily to the approach, and in practice, most surgeons perform a combination of the two. The anterior approach used for septoplasty via a transfixion incision gives excellent access to the whole septum, whereas the lateral approach used for submucous resection denies access to the caudal septum. Theoretically, there is also a difference in the amount of cartilage resected, but in practice, this will vary from individual to individual.

Complications

Septal Hematoma The septal cartilage relies on the perichondrium and overlying mucosa for its blood supply. A hematoma may form between the cartilage and the perichondrium or between the perichondrium and the septal mucosa (Figure 33–2). In the former case, the blood supply to the cartilage will be damaged, and unless the hematoma is drained, cartilage death may ensue. This situation is further exacerbated by secondary infection leading

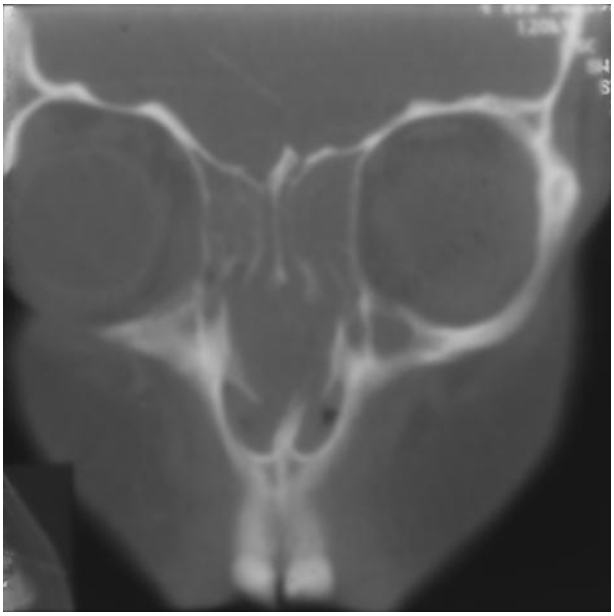


FIGURE 33–2. Coronal computed tomographic scan showing expansion of the anterior part of the nasal septum in a septal hematoma with opacification of the anterior ethmoid cells.

TABLE 33–4. Causes of Septal Perforation

Traumatic
Septal surgery
Cauterization
Packing
Nasotracheal intubation
Foreign bodies, rhinolith
Inflammatory and infectious
Septal abscess or hematoma
Diphtheria
Typhoid
Syphilis
Tuberculosis
Systemic lupus erythematosus
Sarcoidosis
Wegener's granulomatosis
Toxic
Cocaine
Snuff taking
Chromic acid fumes
Sulfuric acid fumes
Arsenicals
Mercurials
Phosphorus
Neoplasia
Carcinoma
Lymphoma
Midfacial necrotizing lesions
Leukemia
Idiopathic

to absorption of the cartilage and saddling of the nose. Consequently, the hematoma should be drained and the nose packed, together with giving broad-spectrum antibiotics.

Septal Perforation There are many causes of septal perforation, the most common of which is some form of trauma, be it iatrogenic or self-inflicted (Table 33–4) (Figure 33–3). The most common site following septal surgery is in the region of the chondrovomerine suture. Consequently, if a mucoperichondrial tear occurs during the surgery, it is advisable to insert a piece of free cartilage at this point, ideally sutured in position.

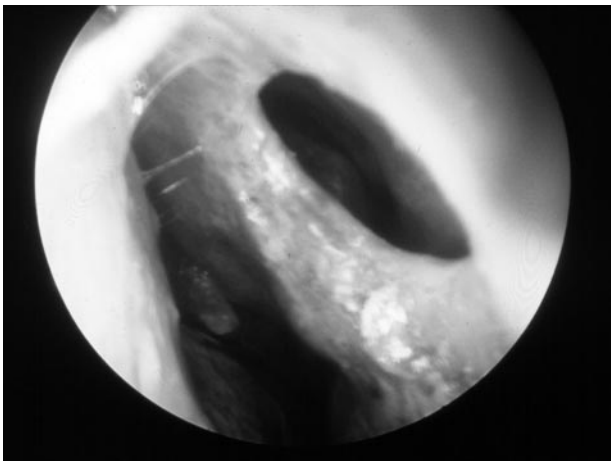


FIGURE 33–3. An endoscopic view of an anterior septal perforation.

Fortunately, many septal perforations are asymptomatic, but those of small diameter may make a whistling sound, and the larger ones may be associated with crusting, spotting of blood, and the sensation of nasal obstruction owing to turbulence of airflow. The more anterior the perforation is, the more likely it is to cause symptoms. Although they will never heal spontaneously, only a minority of perforations require active treatment. In the first instance, a trial of nasal douching with or without a topical corticosteroid may be given.

Many operations have been described to effect closure of septal perforations, the number being inversely proportional to the overall success of such interventions. These have included rotation and advancement of local tissue flaps, in particular "drop down" techniques, free grafts including fascia, split skin, and composites of, for example, pinna cartilage and skin. Bilateral labial flaps from the gingivolabial sulcus tunnelled through into the nose have also been used, but even if successful, all of these procedures suffer from the disadvantages of not providing a ciliated columnar epithelium and being rather bulky.

Alternatively, the perforation may be quickly and easily filled with a commercially available button, which can be customized to the individual perforation and left in place long term.

Infection/Toxic Shock Syndrome Infection after nasal surgery is uncommon but, when it occurs, is very painful and may cause a significant elevation of

body temperature. Cases of toxic shock syndrome have been described as a result of nasal packing. The condition is caused by the toxin produced by a *Staphylococcus* that, when absorbed, produces headache, myalgia, nausea, and vomiting secondary to the pyrexia. Tachycardia, hypertension, hypotension, and erythema of the skin are followed sometime later by desquamation of the skin of the hands. Toxic shock must be dealt with promptly and definitively. Cultures for *S. aureus* must be taken, a β-lactamase-resistant antistaphylococcal antibiotic given systemically, and all packs and stitches removed.

Turbinate Enlargement

Causes The inferior turbinates and, to a lesser extent, the adjacent septum and middle turbinates have a complex submucosal structure, composed of vascular sinusoids under autonomic control. Under normal physiologic conditions, the inferior turbinates change in size with the nasal cycle throughout the day. This may be overridden by a number of factors, such as exercise, posture, emotion, environmental temperature, and humidity. Generally, one is unaware of these changes unless one side of the nose is narrower. A compensatory hypertrophy of the inferior turbinate is often seen on the wider side. A wide range of pathologic processes including allergy, inflammation, and infection can produce symptomatic swelling of the turbinates.

Effects Resistance to airflow primarily occurs in the valve region at the anterior end of the inferior turbinates. To a lesser extent, the middle turbinate changes, obstructing the middle meatus and abutting the septum.

Treatment *Medical* Although topical decongestant sprays can significantly shrink the mucosa, long-term use will lead to rebound congestion and rhinitis medicamentosa. Topical nasal corticosteroids, by contrast, can be used over long periods of time for most forms of perennial rhin(osinus)itis.

Surgical Surgery should be considered only after careful evaluation of the cause of the patient's symptoms and after failure of adequate medication. Most of the techniques can be performed under local or general anesthesia and using head-light, microscope, or endoscopic visualization.[39] As no randomized controlled prospective trial has been performed to determine which of the many techniques is superior in the short or long term, choice devolves to personal preference, and all may be accompanied by hemorrhage and crusting:

- *Lateral outfracture.* Single or multiple fractures of the inferior turbinate bone can be done with or without a submucous incision. The turbinate is pushed closer to the medial wall of the maxilla and may even be pressed into the maxilla (antroconchopexy), but the long-term results are debatable as the turbinate may resume its original position and the procedure does not deal with any soft tissue swelling or its cause.
- *Submucous diathermy.* This has been popular since the early 1900s. A pointed electrode, insulated except for the terminal 3 to 5 mm, is introduced into the anterior end of the inferior turbinate and passed along the whole length. The probe is then withdrawn slowly, over 20 seconds, and the procedure is repeated at two or three points. Hemorrhage and adhesions can occur, and the patient should be warned that there will be blockage owing to reactionary swelling for some weeks after the surgery.
- *Linear electrocautery*. This is done with a red-hot wire electrode to produce a thermal burn or a high-frequency coagulating current using a ball-tip electrode. The technique enjoyed some popularity in the past but has largely been replaced by more precise laser therapy.
- *Turbinectomy.* Partial or subtotal removal of the turbinate offers a permanent solution to hypertrophy. The attachment of the inferior turbinate precludes genuine total turbinectomy unless a portion of the lateral nasal wall is removed. Apart from the occasional mulberry enlargement of the posterior end of the inferior turbinate, it should be remembered that the anterior end of the turbinate offers the greatest resistance to airflow. As with all inferior turbinate surgery, patients must be warned about postoperative hemorrhage, which can be profuse. A secondary hemorrhage rate of 8% is quoted in the literature, which will be severe in 1%.
- *Submucous resection*. Dissection of the mucoperiosteum of the turbinate is described in the original literature, although is difficult to perform in practice owing to the irregularities of the bone. It is best achieved by taking off an anterior sliver and

then nibbling back the exposed bone so that the mucosal edges fall together.

- *Laser turbinectomy.* A variety of cutting and coagulating lasers have been used to "cauterize" and reduce the inferior turbinate, including carbon dioxide, argon, KTP and neodymium:YAG lasers. A cross-hatching of the mucosa is thought to optimize shrinkage while preserving mucosa.
- *Cryosurgery with liquid nitrogen*. This was popular in the past but has largely been superseded by other techniques.
- *Intraturbinal corticosteroid injection.* Enthusiasm for intraturbinal corticosteroid injection diminished after reports of temporary and permanent blindness appeared. These may have occurred owing to retrograde flow of particulate matter into the ophthalmic circulation, possibly associated with inadequate vasoconstriction.

Foreign Bodies

Types and Effects A wide range of objects including metal, plastic, organic materials, and live insects find their way into the nose, either accidentally or deliberately, with size being the only limiting factor. The result is a unilateral chronic purulent nasal discharge, which in a young child should be attributed to a foreign body until proved otherwise. If the foreign material has been present for some time, it may form the nidus for deposition of calcium and magnesium salts, producing a rhinolith that is usually radiopaque.

Treatment The history may be misleading and examination difficult, especially if the patient is uncooperative and the nasal mucous membrane is swollen. Imaging may be helpful but is not exclusive if negative. Therefore, it is often necessary to examine the nose under a short general anesthetic, which should include a throat pack to avoid aspiration of the object.

GRANULOMATOUS CONDITIONS

A range of systemic conditions, including granulomas, vasculitides, connective tissue disorders such as systemic lupus erythematosus, relapsing polychondritis, eosinophilic angiocentric fibrosis, and skin conditions such as pemphigus, pemphigoid, and scleroderma can all manifest in the nose. Consequently, one should have a low threshold of suspicion as one may have the first opportunity to make the diagnosis.

Sarcoidosis

Cause and Effects Sarcoidosis is a systemic condition of unknown cause characterized by noncaseating epithelioid granulomas. It rarely occurs in isolation in the upper respiratory tract, usually being found in association with lower respiratory involvement. The condition has a predilection for certain geographic areas such as the rural southeast United States and Scandinavia and certain ethnic groups such as West Indians and is more common in women.

Nasal symptoms include obstruction, mucopurulent blood-stained discharge, and crusting.[40] There may be associated sinus infection, septal perforation with saddling of the nasal bridge, and violaceous lesions of lupus pernio on the nasal skin. The mucosa is inflamed with crust and old blood, and a "strawberry skin" effect may be seen of the yellowish granulomas against the reddened lining.

Diagnosis includes erythrocyte sedimentation rate, serum calcium, and angiotensin-converting enzyme, although there is no one absolute test for this condition. The Kveim test, which was hitherto the most accurate, has been withdrawn in the United Kingdom owing to concerns over slow viruses. Radiology of the chest is mandatory, and plain radiographs may show punched-out erosion of the nasal bones similar to that seen in dactylitis (Figure 33–4). Nasal biopsy can be useful when mucosal change is

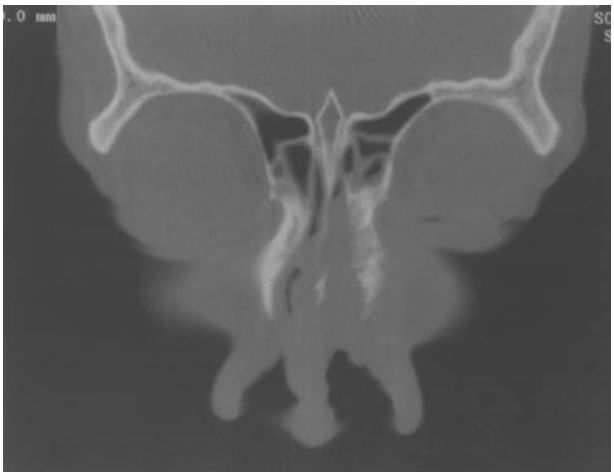


FIGURE 33–4. Coronal computed tomographic scan showing osteitis of the nasal bones in a patient with sarcoidosis.

present but is positive in only 7% when the nose appears normal.[40] The condition must be distinguished from other granulomatous diseases such as tuberculosis, leprosy, berylliosis, Wegener's granulomatosis, and acquired immune deficiency syndrome (AIDS).

Treatment Local treatment with saline douche and topical corticosteroids is helpful, and the nose may improve with systemic treatment that usually includes oral corticosteroids and a variety of cytotoxic agents. Surgery, except for sinus infection, should be avoided as this can exacerbate collapse of the nose.

Wegener's Granulomatosis

Cause and Effects This systemic condition characterized by granulomas and vasculitis is also of unknown cause. An autoimmune basis or infection has been suggested but remains unproven.[41] It was originally described affecting the upper and lower respiratory tracts, together with a focal glomerulonephritis, which rapidly led to renal failure and death. It is now known to affect any part of the body ab initio and to have a variable natural history, progressing inexorably in some individuals over a few weeks or months or to a more insidious limited form in others that may continue for some years before affecting other organs. Whether all of the latter ultimately progress to the full-blown disease is unknown.

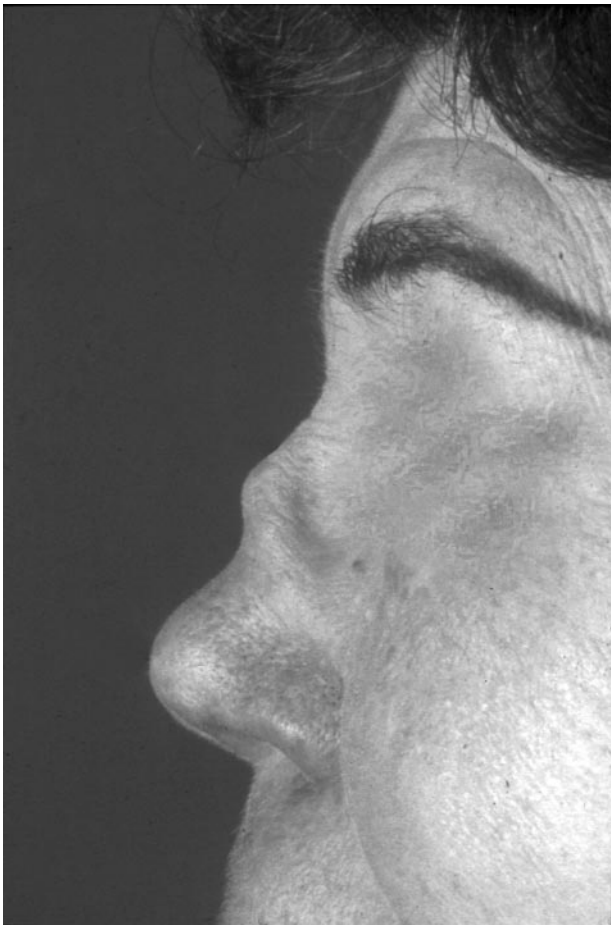


FIGURE 33–5. Photograph showing typical collapse of the nasal bridge in a patient with Wegener's granulomatosis.

Patients often present with a short history of progressive malaise, pyrexia, weight loss, and a disproportionate feeling of "unwellnesss" in comparison with what are often fairly nonspecific findings. In the nose, the swollen, inflamed mucosa produces blockage and blood-stained crusting. There may be destruction of the septum with a characteristic implosion of the nasal bridge (Figure 33–5). The nasopharynx, ears, mouth, larynx, trachea, cranial nerves, and orbit may all be involved.

Diagnosis of the condition relied in the past on clinical acumen supported by a high erythrocyte sedimentation rate, C-reactive protein, and evidence of pulmonary and renal damage. The advent of the C-ANCA (antineutrophil cytoplasmic antigen) greatly assisted diagnosis, being highly specific and sensitive for the condition. However, a negative C-ANCA does not absolutely exclude the condition, particularly in the limited form and/or when oral corticosteroids have been given. Computed tomography of the nose and paranasal sinuses may show some midline destruction and often shows opacification of the sinuses (Figure 33–6). This may be attributable to active granulomatous infiltration, burnt-out disease with fibrotic change, or secondary infection. Unfortunately, biopsy is not diagnostic, at best being described as "consistent with" Wegener's granulomatosis, but is helpful to exclude other pathology such as T-cell lymphoma.

Treatment Medical therapy includes systemic corticosteroids, a range of cytotoxic drugs (eg, cyclophosphamide and azathioprine), immunoglobulin infusions, cotrimoxazole, and plasma exchange. Earlier diagnosis and careful monitoring have significantly improved the outcome for these patients, but the condition remains life threatening in its more severe forms. The nose is again managed with douching and topical corticosteroids. Surgery

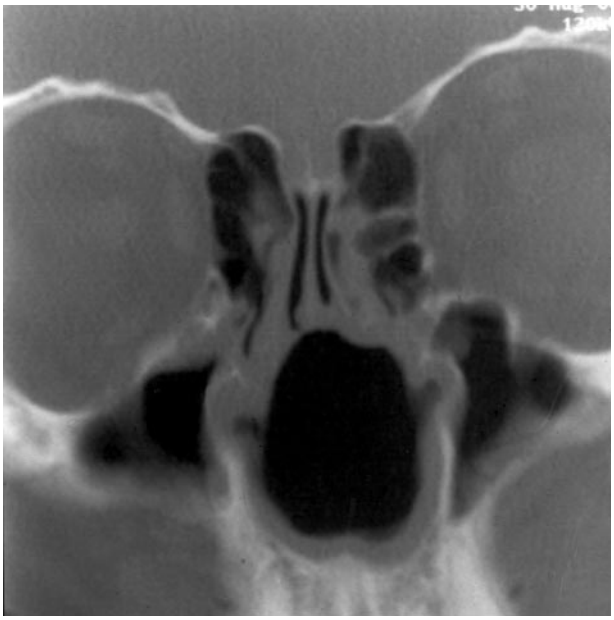


FIGURE 33–6. Coronal computed tomographic scan in a patient with Wegener's granulomatosis showing midline destruction of the septum and loss of the inferior turbinates.

may be required for secondary infection, but attempts to reconstruct the nose should be resisted until the disease has been quiescent for some years.

AMYLOIDOSIS

Apart from amyloid deposits in the skin of the vestibule in generalized amyloidosis, involvement of the nose and sinuses is extremely rare. In contrast to the upper respiratory tract as a whole, only a handful of isolated amyloid deposits have been reported, whereas in generalized plasmacytic malignant lymphoma, 20% of patients may have sinonasal amyloid deposits. The diagnosis may be confirmed histologically by the presence of Congo red reactivity and green birefringence in polarized light. Colchicine has been used in the treatment of the condition.

EPISTAXIS

The nose has an extremely good blood supply from both the external and internal carotid arteries, with crossover between the right and left sides. Many conditions can produce nose bleeds, both local and general (Table 33–5). The majority of local causes are some form of self-inflicted, accidental, or iatrogenic trauma. A range of general causes does not include hypertension per se but rather the underlying problem of atherosclerosis in which changes in the tunica media lead to loss of elasticity and the ability of vessels to constrict.

TABLE 33–5. Causes of Epistaxis

- Local
  - Idiopathic
  - Trauma
    - Self-inflicted
    - Maxillofacial fractures
    - Iatrogenic surgery
    - Drying, septal deviation
    - Foreign bodies
    - Septal perforation
  - Inflammation, allergy
    - Atrophic rhinitis
  - Infective
  - Tumors
    - Benign, eg, angiofibroma
    - Malignant
- General
  - Atherosclerosis
  - Bleeding dyscrasias
    - Coagulation disorders
      - Congenital
        - Hemophilia
        - Christmas disease
        - von Willebrand's disease
      - Acquired
        - Deficiency of clotting factors, eg, liver and renal failure
        - Massive transfusion
        - Vitamin deficiency
        - Anticoagulants
        - Myelosuppressive drugs, eg, chemotherapeutic agents, anti-inflammatories, antibiotics
    - Hemopoietic
      - Leukemia
      - Aplastic anemia
      - Dysproteinemia
      - Lymphoma
      - Widespread metastases
  - Hereditary hemorrhagic telangiectasia
  - Endocrine
    - Vicarious menstruation
    - Pregnancy

Hereditary hemorrhagic telangiectasia is an autosomal dominant, non–sex-linked transmitted condition characterized by telangiectasias that are deficient in muscle and/or elastic tissue (Figure 33–7). The majority of patients experience epistaxis of varying severity, although the lesions can be found anywhere in the body. Although a wide range of treatment options have been described, a treatment algorithm shows how this may be tailored for the individual patient.[42]

Management In most cases, the clinical history and examination, including endoscopy, will provide the diagnosis, although in appropriate cases, hematologic screening and imaging, including arteriography, may be required.

The treatment will be determined by the severity of the bleed and ranges from first aid measures and cauterization to various forms of packing, arterial ligation, or embolization together with correction of any underlying medical problem.

First Aid Measures. Cauterization. Small prominent vessels on the septum have been traditionally cauterized with silver nitrate sticks or electrocautery as an outpatient procedure. However, this is not usually appropriate for significant active bleeding. Cauterization under endoscopic control for posterior epistaxis is being used with some success and may avoid packing and inpatient stay.

Packing A range of packing materials is available from resorbable hemostatic materials to ribbon gauzes soaked in various substances (eg, Vaseline, bismuth), and commercially manufactured balloons and packs. Any packing must be used as atraumatically as possible to avoid damage to the columella and mucosa and subsequent adhesions and perforation of the nasal septum. Patients with packing in place generally require admission to hospital to ensure that the packing does not become dislodged or impact adversely on their respiratory function. While packing is in place, patients require antibiotics to prevent the development of infection and the possibility of toxic shock.

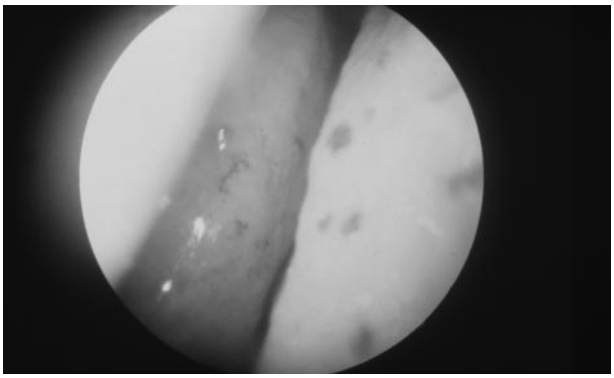


FIGURE 33–7. An endoscopic view of nasal lesions in hereditary hemorrhagic telangiectasia.

Surgery It is often better to intervene earlier rather than later in patients with persistent bleeding from the nose. Endoscopic examination may reveal the precise bleeding spot, which can then be cauterized or lasered. In more severe cases, arterial ligation may be required. This may be directed at the anterior and posterior ethmoidal, maxillary, or sphenopalatine arteries or, in severe cases, the external carotid artery. The ethmoidal vessels may be accessed via a small external incision and represent the internal carotid component of the nasal blood supply. The internal maxillary artery can be reached via the posterior wall of the maxillary sinus through an anterior antrostomy, but finding its many branches can be a time-consuming affair, and many surgeons now prefer the direct approach to the sphenopalatine artery, which can be approached endoscopically as it passes from the sphenopalatine canal into the mucosa of the lateral wall of the nose just beneath the horizontal attachment of the middle turbinate. The external carotid artery can be ligated in the neck on one or both sides.

Embolization In some centers, interventional radiology is the preferred option for reducing the blood supply to the nose by embolization techniques.

RHINOPHYMA

In rhinophyma, a disfiguring enlargement of the nose develops owing to an overgrowth of sebaceous tissue. Tissue overgrowth begins at the tip of the nose and progresses to involve the ala nasi and the columella and is associated with a florid discoloration. The affected tissue may be pared down using a knife, electrodissection, or laser therapy, producing good cosmetic results.

MALIGNANT TUMORS OF THE SKIN OF THE NOSE

Basal Cell Carcinoma

The skin of the nose is the most common site for basal cell carcinoma in the head and neck. It is considerably more common than squamous cell carcinoma, with a peak incidence in the sixth to eighth decades. The most common areas are the side, dorsum, and tip of the nose, and the lesion may be a small nodular growth or form a chronic ulcer with a rolled edge. Although many behave in an indolent fashion, unless treated appropriately, they have a tendency to recur, which may lead to considerable destruction of the face. The side of the nose adjacent to the medial canthal area can prove particularly problematic.

Basal cell tumors are usually very sensitive to radiotherapy but may require surgical excision, particularly for recurrence. A variety of local rotation flaps or full-thickness skin grafts are available.

Squamous cell carcinoma, though less common, may also affect the skin of the nose, again in the sixth to eighth decades. The lesion is generally nodular, resembling a keratoacanthoma, but can ulcerate. When ulcerated, differentiation from a basal cell carcinoma may be difficult (Figure 33–8). Treatment again relies on radiotherapy and/or surgery. Squamous cell carcinoma has a higher propensity for metastatic spread than basal cell carcinoma, and lesions in the midline, particularly on the columella, may disseminate bilaterally, producing a potentially poor prognosis as in continuity, lymphatic resection is not possible.

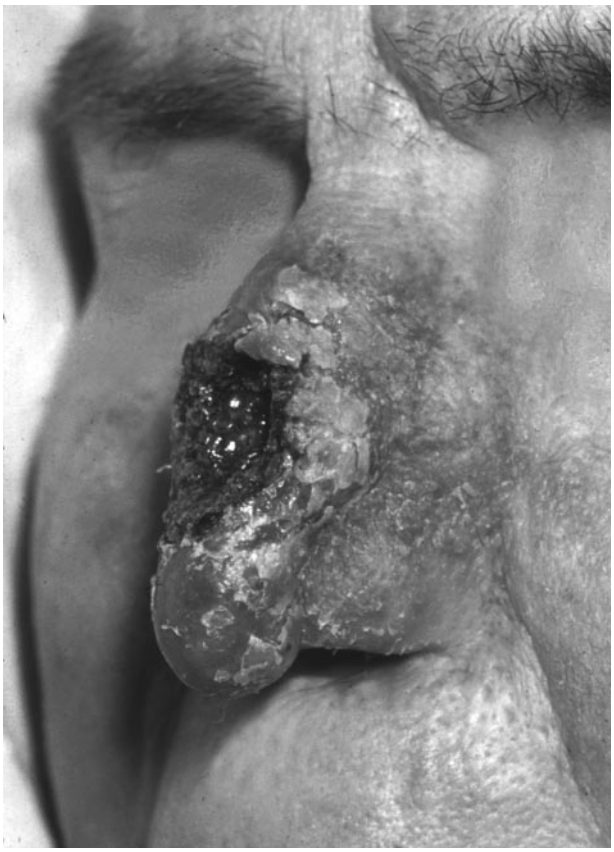


FIGURE 33–8. Photograph showing a squamous cell carcinoma of the bridge of the nose.

REFERENCES

1. Lund VJ, Aaronson D, Boŭsceŭet J, et al. International Rhinitis Management Working Group. The International Consensus Report on the Diagnosis and Management of Rhinitis. Allergy 1994;49 Suppl: 1–34.
2. Bousquet J, Van Cauwenberge P, Khateau N, et al. Allergic rhinitis and its impact on asthma. Geneva: World Health Organization; 2001.
3. Lund VJ, Scadding GK. Objective assessment of endoscopic sinus surgery in the management of chronic rhinosinusitis. J Laryngol Otol 1994;108:749–53.
4. Klossek JM, Dubreuil L, Richet H, et al. Bacteriology of the adult middle meatus. J Laryngol Otol 1996; 110:847–9.
5. Andersen I, Camner P, Jensen PL, et al. Nasal clearance of monozygotic twins. Am Rev Respir Dis 1974;100:301–5.
6. Rutland J, Dewar A, Cox T, Cole P. Nasal brushing for the study of ciliary ultrastructure. J Clin Pathol 1982;35:357–9.
7. Lindberg S, Cervin A, Runer T. Nitric oxide (NO) production in the upper airways is decreased in chronic sinusitis. Acta Otolaryngol (Stockh) 1997; 117:113–7.
8. Holmstrom M, Scadding GK, Lund VJ. The assessment of nasal obstruction—a comparison between rhinomanometry and nasal inspiratory peak flow. Rhinology 1990;28:191–6.
9. Clement PAR. Committee report on standardisation of rhinomanometry. Rhinology 1984;22:151–5.
10. Hillberg O, Jackson AC, Swift DL. Acoustic rhinometry: evaluation of nasal cavity geometry by acoustic deflections. J Appl Physiol 1989;66:295–303.
11. Amoore JE. Odor standards in squeeze bottle kits for matching quality and intensity. Wat Sci Tech 1992; 25:1–9.

12. Doty RL, Shaman P, Dann M. Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function. Physiol Behav 1984;32:489–502.
13. Simmen D, Briner HR, Hess K. Screeningtest des Geruchssinnes mit Riechdisketten. Laryngorhinootologie 1998;77:1–6.
14. Kobal G, Hummel T, Sekinger B, et al. "Sniffin' sticks": screening of olfactory performance. Rhinology 1996;344:222–6
15. Gwaltney JM, Phillips CD, Miller RD, et al. Computer tomographic study of the common cold. N Engl J Med 1994;330:25–30.
16. Shahin J, Gullane PJ, Dayal VS. Orbital complications of acute sinusitis. J Otolaryngol 1987;16:23–7.
17. Lund VJ. Complications of sinusitis. In: Scott-Brown's otolaryngology. Vol 4. Rhinology. London: Butterworth Heineman; 1997. p. 1–11.
18. Silvers WS. The skier's nose: a model of cold-induced rhinorrhea. Ann Allergy 1991;67:32–6.
19. Bascom R, Kesavanatha J, Fitzgerald TK, et al. Sidestream tobacco smoke exposure acutely alters human nasal mucociliary clearance. Environ Health Perspect 1995;103:1026–30.
20. Vinke JG, KleinJan A, Severijnen LW, Fokkens WJ. Passive smoking causes an "allergic" cell infiltrate in the nasal mucosa of non-atopic children. Int J Pediatr Otorhinolaryngol 1999;51:73–81.
21. Annesi-Maesano I, Oryszczyn MP, Neukirch F, Kauffmann F. Relationship of upper airway disease to tobacco smoking and allergic markers: a cohort study of men followed up for 5 years. Int Arch Allergy Immunol 1997;114:193–201.
22. Lacroix JS, Buvelot JM, Polla BS, Lundberg JM. Improvement of symptoms of nonallergic chronic rhinitis by local treatment with capsaicin. Clin Exp Allergy 1991;21:595–600.
23. Schiffman SS, Nagle HT. Effect of environmental pollutants on taste and smell. Otolaryngol Head Neck Surg 1992;106:693–700.
24. Ellegard E, Karlsson G. Nasal congestion during pregnancy. Clin Otolaryngol 1999;24:307–11.
25. Schatz M. Special considerations for the pregnant woman and senior citizen with airway disease. J Allergy Clin Immunol 1998;101(Suppl 2 Pt 2): S373–8.
26. Mazzotta P, Loebstein R, Koren G. Treating allergic rhinitis in pregnancy. Safety considerations. Drug Saf 1999;20:361–75.
27. Girgis IH, Yassin A, Hamdy H, Moris M. Estimation of effect of drugs on the nasal circulation. J Laryngol Otol 1974;88:1163–8.
28. Bauer GE, Hull RD, Stokes GS, Raftos J. The reversibility of side effects of guanethidine therapy. Med J Aust 1973;1:930–3.
29. Proud D, Naclerio RM, Meyers DA, et al. Effects of a single-dose pretreatment with captopril on the immediate response to nasal challenge with allergen. Int Arch Allergy Appl Immunol 1990;93:165–70.
30. Schwartz RH, Estroff T, Fairbanks DN, Hoffmann NG. Nasal symptoms associated with cocaine abuse during adolescence. Arch Otolaryngol Head Neck Surg 1989;115:63–4.
31. Dax EM. Drug dependence in the differential diagnosis of allergic respiratory disease. Ann Allergy 1990;64:261–3.
32. Graf P. Rhinitis medicamentosa: aspects of pathophysiology and treatment. Allergy 1997;52(40 Suppl):28–34.
33. Scadding GK. Rhinitis medicamentosa [editorial]. Clin Exp Allergy 1995;25:391–4.
34. Bousquet J, Metcalfe D, Warner J. Food allergy. Report of the Codex alimentarius. ACI Int 1997;9:10–21.
35. Henriksen S, Gundersen W. The aetiology of ozaena. Acta Pathol Microbiol Scand 1959;47:380–6.
36. Young A. Closure of the nostrils in atrophic rhinitis. J Laryngol Otol 1967;81:515–24.
37. Blom HM, Godthelp T, Fokkens WJ, et al. The effect of nasal steroid aqueous spray on nasal complaint scores and cellular infiltrates in the nasal mucosa of patients with nonallergic, noninfectious perennial rhinitis. J Allergy Clin Immunol 1997;100(6 Pt 1):739–47.
38. Euler AR. Upper respiratory tract complications of gastroesophageal reflux in adult and pediatric-age patients. Dig Dis 1998;16:111–7.
39. Lund VJ. Surgery of the inferior turbinate. Fac Plast Surg 1986;3:227–80.
40. Wilson R, Lund VJ, Sweatman M, et al. Upper respiratory tract involvement in sarcoidosis and its management. Eur J Respir Med 1988;1:269–72.
41. Lund VJ, Cambridge G. Immunological aspects of Wegener's granulomatosis. In: Veldman JE, Passali D, Lim DJ, editors. New frontiers in immunology. Amsterdam: Kugler Publications; 2000. p. 195–207.
42. Lund VJ, Howard J. A treatment algorithm for the management of epistaxis in hereditary haemorrhagic telangiectasia. Am J Rhinol 1999;13:319–22.

CHAPTER 34

Sinusitis and Polyposis

Andrew P. Lane, MD, David W. Kennedy, MD

EPIDEMIOLOGY

Rhinosinusitis is an extremely prevalent disorder that has a significant impact on the quality of life of affected individuals and places a large economic burden on the United States as a whole. It is estimated that approximately 33 million Americans per year get sinusitis, with 26.7 million office visits attributed to sinus and related airway disorders.[1,2] The cost of these visits is nearly 6 billion dollars, on top of which people spend an additional 2 billion dollars annually for over-the-counter sinusitis medications.[2,3] There is a loss in worker productivity related to rhinosinusitis, with 12.5 million workdays lost and 58.7 million days of restricted activity owing to the disease.[4]

In 1995, sinusitis was the fifth leading diagnosis for which antibiotics were prescribed nationwide.[5] When sinusitis recurs, it has been shown that subsequent episodes cost successively more for diagnostic tests and therapy.[6] Rhinosinusitis affects men and women equally, and it is thought to be more common in childhood than in adults. However, the health care expenditures are approximately split 30% for children and 70% for adults.[2] In demographic terms, rhinosinusitis is more common in the Midwest and South than in the northeast and western United States. The incidence is higher in spring, winter, and fall than it is during the summer months.

PATHOPHYSIOLOGY DEFINITION AND CLASSIFICATION OF RHINOSINUSITIS

Acute Rhinosinusitis

Etiology and Pathophysiology The underlying causes of acute rhinosinusitis are multiple and include a variety of host and environmental factors (Table 34–1). Ultimately, however, the common pathway of acute sinusitis is thought to be the presence of bacteria in a sinus cavity with an obstructed ostium. This requires not only blockage of the normal anatomic outflow of the sinus but also a failure of the mucociliary clearance function of the mucosa that would ordinarily remove the bacteria. In addition, abnormalities of the quantity or consistency of the sinonasal secretions can affect mucociliary clearance and promote bacterial growth. Finally, immunodeficiencies, such as selective immunoglobulin (Ig) A or IgG deficiency or human immunodeficiency virus (HIV) infection, lead to an increased incidence of acute sinusitis owing to an inability to clear pathogenic bacteria.

There are multiple factors that can lead to anatomic obstruction of the sinus ostia. In most

TABLE 34–1. Host and Environmental Factors Associated with the Pathogenesis of Rhinosinusitis

Host factors
Genetic/congenital conditions
Cystic fibrosis
Immotile cilia syndrome
Allergic/immune conditions
Anatomic abnormalities
Systemic diseases
Endocrine
Metabolic
Neuromechanisms
Neoplasm
Environmental factors
Infectious/viral bacterial and fungal agents
Trauma
Noxious chemicals
Iatrogenic
Medications
Surgery

cases, it is edema secondary to inflammation that constricts the small sinus openings in a reversible fashion. This may result from a viral upper respiratory infection, allergy, environmental irritants, or barotrauma. Irreversible blockage caused by fixed anatomy may contribute to the problem or be the sole etiologic agent. Examples of anatomic obstruction would include septal deviations, polyps, nasal tumors, foreign bodies, or postsurgical synechiae. Additionally, sinus disease is often related to a host of anatomic variants such as Haller's cells, agger nasi cells, concha bullosa, and paradoxical middle turbinate, although this relation has not been definitively supported by randomized, controlled studies. Whatever the cause, once the ostium becomes occluded, a local hypoxia develops in the sinus cavity, and sinus secretions accumulate. This combination of low-oxygen tension and a rich culture medium of secretions allows exponential bacterial growth to occur within the sinus.

A normal sinus is generally considered to be a sterile environment. Bacteria that find their way into the sinus are rapidly eliminated by the action of ciliated columnar epithelial cells moving a layer of viscous secretions created by epithelial goblet cells and submucosal glands. Abnormalities of the cilia themselves, or of the double-layered mucus blanket within which the cilia beat, may hinder bacterial removal. Ciliary function may be defective secondary to underlying disease states such as primary ciliary dyskinesia or Kartagener's syndrome or to injury by environmental irritants or surgical trauma. In animal models of sinusitis, bacterial products and endogenous mediators of inflammation disrupt epithelial cilary activity.[7] The quality of the mucus in the sinuses is of critical importance in the effectiveness of ciliary clearance. In cystic fibrosis and severe dehydration, for example, the viscosity and elasticity of the mucus are increased by alterations in water and electrolyte transport. Additionally, increased production of mucus, which can occur in response to airborne irritants, allergens, cold air exposure, or viral upper respiratory infections, also may exceed the rate of mucociliary clearance. Any of these situations will lead to the accumulation of mucus in the sinuses, with poor removal of bacteria and the creation of a favorable environment for bacterial growth.

Any systemic disease leading to an immunocompromised state will potentially predispose a patient to acute sinusitis. Chronic illness, such as diabetes, or malnutrition, metabolic derangement, chemotherapy, or long-term corticosteroid therapy will similarly increase the tendency to develop acute sinusitis. The intensive care unit setting is particularly conducive to the development of a nosocomial sinusitis owing to the presence of critical illness in conjunction with nasotracheal and/or nasogastric tubes, which cause mechanical obstruction of sinus drainage. Although fluid and mucosal thickening in the sinuses is common, acute sinusitis is rarely the sole or main cause of fever or sepsis in this patient population.[8] Antral puncture is only indicated in the presence of an endoscopic examination that reveals purulence.[9] Immunocompromised patients with HIV infection or other severe immune deficiencies are prone to sinus infection with unusual pathogens such as fungi and atypical mycobacteria. Bone marrow transplant patients are another group with a high incidence of acute sinusitis and a propensity toward developing life-threatening invasive fungal sinusitis.

Classification Sinusitis is a clinical diagnosis, based largely on history and physical examination findings. In 1997, the Rhinosinusitis Task Force of the American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) attempted to create a uniform diagnostic paradigm for sinusitis by organizing common sinonasal symptoms and signs into major and minor factors.[10] The presence of two or more major factors, or one major and two minor factors, is considered suggestive of sinusitis. These signs and symptoms of sinusitis are discussed in the diagnosis section of this chapter. The Task Force defined five categories of adult rhinosinusitis in their report. These categories are based on duration, history, and physical examination. They do not take into account radiologic findings. Table 34–2 shows a list of major and minor factors associated with the diagnosis of rhinosinusitis that are used in this classification scheme. The first category is "acute," which is defined, in part, as having a duration of less than or equal to 4 weeks. The patient's history must include either two or more major factors or one major and one minor factor. If the patient has nasal purulence on endoscopy, the presence of a strong history is not necessary. Acute rhinosinusitis should be considered if patient symptoms worsen after 5 days or if symptoms persist for more than 10 days.

TABLE 34–2. Major and Minor Factors Associated with the Diagnosis of Rhinosinusitis

Major factors
Facial pain/pressure*
Facial congestion/fullness
Nasal obstruction/blockage
Nasal discharge/purulence/discolored postnasal discharge
Hyposmia/anosmia
Purulence in nasal cavity on examination
Fever (acute rhinosinusitis only)†
Minor factors
Headache
Fever (in nonacute cases)
Halitosis
Fatigue
Dental pain
Cough
Ear pain/pressure/fullness

*Facial pain/pressure alone does not constitute a suggestive history for rhinosinusitis in the absence of another major symptom or sign.

†Fever in acute sinusitis alone does not constitute a strongly suggestive history for acute sinusitis in the absence of another major symptom or sign.

The next category is "subacute," which is defined by a duration of 4 to 12 weeks. "Recurrent acute" is defined as four or more episodes of rhinosinusitis per year with complete resolution between episodes.

Chronic Rhinosinusitis

Rhinosinusitis lasting longer than 12 weeks is classified as chronic. If the symptoms worsen periodically and then return to baseline with treatment, the AAO-HNS Task Force separately categorizes this as acute exacerbations of chronic rhinosinusitis. The pathophysiology of chronic sinusitis shares many similar features with acute sinusitis, but with some important differences. The predisposing conditions for chronic disease parallel those for acute infections in terms of environmental, systemic, and local host factors. However, in addition, multiple episodes of acute sinusitis themselves may cause scarring and mucosal dysfunction that lead to chronic sinus infections. Symptoms in chronic sinusitis are usually more insidious than in an acute situation and therefore may be more difficult to diagnose. Another important distinction is that the underlying process in chronic sinusitis is not necessarily infectious and is often a self-perpetuating inflammatory process. The exact nature of the mucosal dysregulation is unknown but in some cases can be recalcitrant to all medical and surgical therapies currently available. Many times, the only effective treatment is systemic corticosteroids, which provide short-term control of symptoms at the risk of serious complications related to their use.

Whereas acute sinusitis is histologically an exudative process characterized by neutrophilic infiltration and necrosis, chronic sinusitis is a proliferative process remarkable for thickening of the mucosa and lamina propria. The predominant infiltrative cell in chronic sinusitis is the eosinophil, both in the allergic and the nonallergic patient. There is evidence that potent eosinophil-attracting chemokines are produced in the sinus mucosa, elaborated by a variety of cell types under the stimulation of cytokines produced largely by T cells.[11] An increase in the levels of interleukin-4 and -5 in the sinonasal tract promotes the continued proliferation and prolonged life span of eosinophils.[12] A number of other proinflammatory cytokines are up-regulated and participate in the process of directing lymphocyte and granulocyte traffic, while causing further production of cytokines in an autocrine fashion. Degranulation of eosinophils releases a number of destructive enzymes that damage the epithelium. This disrupts the normal barrier function and the mucociliary activity of the mucosa, allowing bacteria and fungi to colonize the sinus cavities. The damage to the epithelium irritates sensory nerve endings, causing pain and stimulating changes in mucus secretion and endothelial permeability via reflex pathways. Corticosteroids disrupt this vicious cycle by interfering with the transcription of many cytokines at the nuclear level. Without these mediators in the sinus mucosa, the inflammatory process is halted, and there is a return to more normal function. However, for reasons that remain poorly understood, the reversal of inflammation is often temporary and recurs once the corticosteroids are withdrawn.

Fungal Rhinosinusitis

Some forms of sinusitis are caused by fungal microorganisms within the sinonasal tract. The fungal infection can be either invasive or noninvasive.

One noninvasive form is a fungal ball growing within a sinus cavity. The invasive forms can present in either an indolent, chronic form or a fulminant, acute form. The latter frequently occurs in immunosuppressed patients, such as diabetics in ketoacidosis, transplant recipients, HIV-infected individuals, and patients undergoing chemotherapy. Without control of the underlying immunosuppression, the process is rapidly progressive and may extend into the orbits and intracranial space despite aggressive surgical and medical management. The disease is uniformly fatal if the immune failure cannot be corrected. Even if the underlying condition is controlled, significant morbidity and mortality may result based on the ultimate extent of the disease. In contrast, chronic invasive sinusitis tends to occur in immunocompetent hosts and generally advances very slowly. Vascular invasion in the chronic form is minimal or nonexistent. The microorganisms involved in fulminant fungal sinusitis are varied, but the *Mucoraceae* family (most often *Rhizopus oryzae*) and *Aspergillus* species are most frequent (Figure 34–1, A and B). A number of fungal microorganisms have been reported as pathogens in chronic invasive disease, with *Aspergillus* being the most typical. *Mucor*, *Alternaria*, *Curvularia*, *Bipolaris*, *Candida*, *Drechslera*, *Sporothrix schenckii*, and *Pseudallescheria boydii* (which is unresponsive to amphotericin B) have specifically been cultured.[13] A further classification of chronic invasive fungal sinusitis was proposed by deShazo et al based on histopathology.[14] The form of the disease that lacks a granulomatous response has been associated with a high rate of persistence, recurrence, and a poor prognosis. However, the presence or absence of a granulomatous histology is not apparent in the clinical presentation and does not currently alter treatment.

In addition to fungal balls, the other noninvasive form of fungal sinusitis is an entity known as allergic fungal rhinosinusitis (AFRS). It is the most common form of fungal sinus disease, although the pathogenesis remains poorly understood. The diagnosis of AFRS depends on five criteria: (1) type I hypersensivity, (2) nasal polyposis, (3) characteristic computed tomographic (CT) scan appearance (hyperdense material in the sinus cavity), (4) positive fungal stain or culture, and (5) no evidence of tissue invasion.[15] The disease is also characterized by increased total IgE, antigen-specific IgE, and peripheral eosinophil count. Allergic fungal rhinosinusitis is often unilateral and can cause bony erosion with orbital or intracranial extension (Figure 34–2). At surgery, there is typically thick, brown-green mucus that has been called "peanut butter"-like. Histologically, this mucus contains eosinophils in sheets, Charcot-Leyden crystals, and fungal hyphae. The critical element that separates AFRS from chronic eosinophilic rhinosinusitis is the presence of allergy to fungus.[16] Allergic fungal rhinosinusitis is not caused by the abnormal presence of fungus in the nose but rather an abnormal response to nonpathogenic fungi that exist in the environment. It has been shown by Ponikau et al that fungi are ubiquitous and can be isolated out of every nose, given sufficiently sensitive culture techniques.[17] A wide range of fungal species have been implicated in allergic fungal

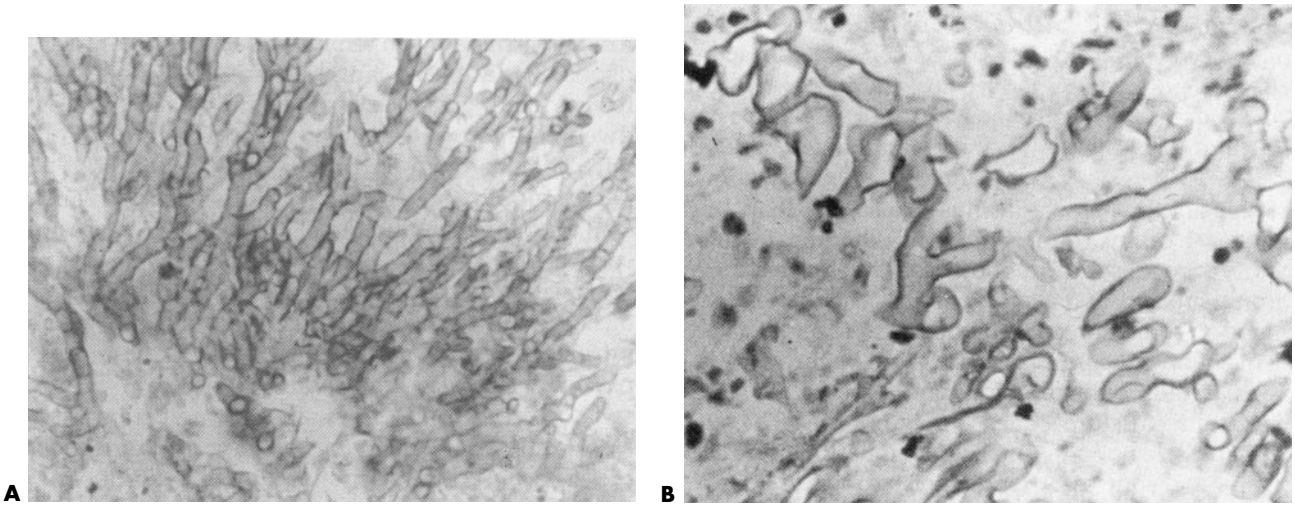


FIGURE 34–1. *A*, Hematoxylin and eosin stain of *Aspergillus*. *B*, Hematoxylin and eosin stain of *Mucor*.

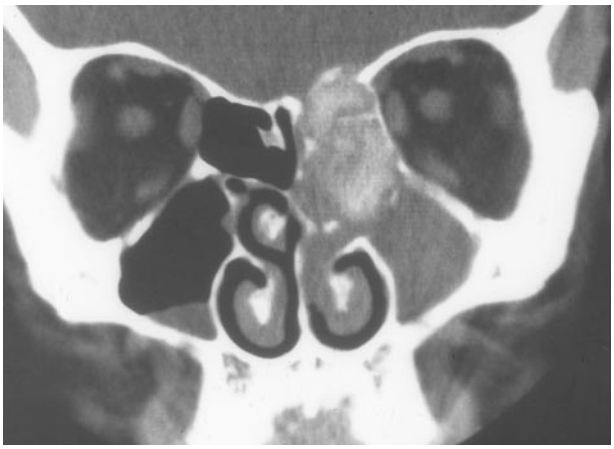


FIGURE 34–2. Coronal computed tomographic scan showing typical appearance of allergic fungal sinusitis on the patient's left side. Incidental note is made of fluid in the right maxillary sinus.

sinusitis, and it is important to identify the specific species if desensitization is to be attempted. The mainstays of therapy are surgical débridement and systemic corticosteroids. Allergy desensitization and the use of antifungal medications are controversial but may be of benefit in some patients.

POLYPOSIS

Etiology

Nasal polyposis is a prevalent condition that is said to affect between 1 and 4% of the general US population.[18] Based on autopsy studies, the true incidence of polyps may be much greater, although most polyps are small and presumably asymptomatic. The etiology of nasal polyps is largely unknown and has long been a topic of debate. Although historically many have believed polyps to be a manifestation of allergy, in part because of the histologic prominence of eosinophils, epidemiologic evidence for this is lacking. The incidence of allergy is not higher in patients with nasal polyps than in the population as a whole, nor do polyp patients have elevated rates of positive allergy skin tests.[19] Nasal polyps are associated with a number of systemic diseases including aspirin intolerance, intrinsic asthma, primary ciliary dyskinesia, and cystic fibrosis. They are frequently observed in chronic rhinosinusitis, including allergic rhinosinusitis, and other chronic sinonasal inflammatory states. Overall, the mechanisms behind polyp formation are believed to be multifactorial. Recent evidence suggests an important role for proinflammatory cytokines, chemokines, and chemotactic factors in the pathogenesis of inflammatory polyps, along with a variety of environmental, genetic, and biochemical factors that have previously been proposed.[20]

Gross Appearance and Histology

When significant intranasal polyposis is present, polyps can easily be seen by anterior rhinoscopy. Grossly, they are translucent to pale gray, pear shaped, smooth, soft, and freely mobile. Polyps arise from the lateral nasal wall and in many cases are limited to the middle meatus, where they can only be visualized endoscopically (Figure 34–3). Inflammatory polyps are usually seen bilaterally. Histologically, polyps are composed of a fibromyxomatous stroma covered by typical respiratory epithelium that may display benign squamous cell metaplasia. The epithelium displays very few nerve endings and submucosal glands, and the basement membrane is thickened. As compared to adjacent lateral nasal wall mucosa, there is a marked abundance of eosinophils and mast cells in inflammatory polyps.

Diagnosis

The typical complaints associated with polyposis are nasal congestion, rhinorrhea, and olfactory dysfunction. A thorough nasal endoscopic examination is

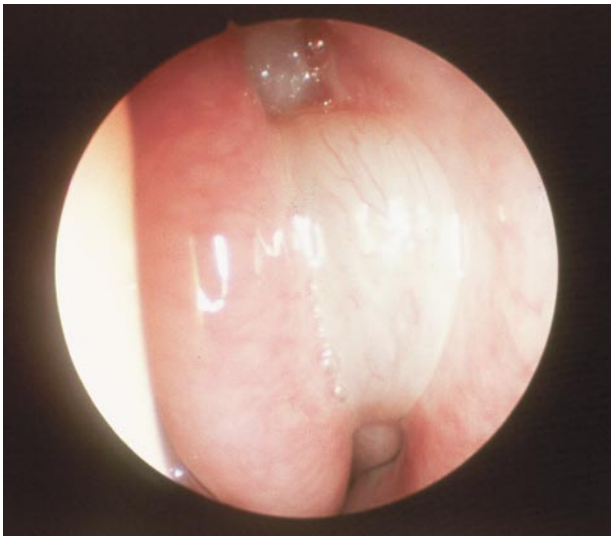


FIGURE 34–3. Polyp in the left middle meatus.

necessary for all patients with polyp disease to assess the gross appearance and sites of origin. When evaluating a patient with nasal polyps, the differential diagnosis of nasal masses should be entertained, including the possibility of systemic disease. For example, the finding of nasal polyps in children should elicit consideration of cystic fibrosis or ciliary dyskinesia. Unilateral polyps should raise concern for allergic fungal sinusitis or inverted papilloma. Sinonasal neoplasms may display a similar endoscopic appearance to polyps; thus, consideration should be given to the possibility of carcinoma, sarcoma, angiofibroma, meningioma, or esthesioneuroblastoma in certain patients. A single, unilateral polyp originating high in the nasal cavity or with a stalk that is not clearly visible may represent an encephalocele or meningocele. Visible pulsations on endoscopy and enlargement of the mass with ipsilateral internal jugular vein compression (Furstenberg's sign) help to confirm the diagnosis. As a rule, if the intranasal mass does not have the characteristic appearance of a polyp, is unilateral, bleeds easily, or has a stalk that is not clearly identified, imaging studies are indicated before proceeding with management (Figure 34–4). High-resolution CT may reveal bony dehiscences or erosion. Magnetic resonance imaging (MRI) is confirmatory of meningoencephalocele and is advisable when evaluating an area of skull base erosion adjacent to an opacified sinus. Magnetic resonance imaging is also useful for identifying the extent of nasal tumors and differentiating tumor from secondary mucosal disease and retained secretions. These radiologic studies should be obtained prior to performing a biopsy of any unusual nasal mass.

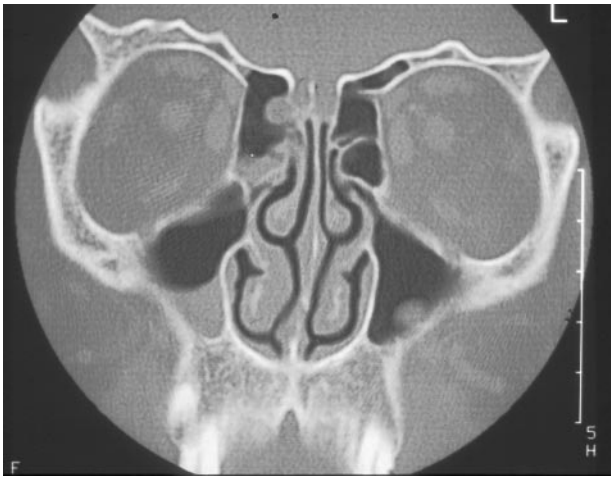


FIGURE 34–4. Coronal computed tomographic scan showing a meningocele in the right ethmoid. The appearance of encephaloceles and meningoceles may mimic that of nasal polyps in some patients.

Management

The most important element in the treatment of nasal polyposis is medical therapy with corticosteroids. Both oral corticosteroids and topical nasal corticosteroids are effective in shrinking polyps and controlling their recurrence. Topical corticosteroids are first-line therapy that should be employed prior to considering surgical intervention. Unless there is a contraindication, a trial of a tapering course of oral corticosteroids is also frequently used prior to surgical resection. Should surgery eventually become necessary, topical corticosteroids and occasionally oral corticosteroids may be needed for long-term maintenance. A variety of topical nasal corticosteroid preparations are available today with similar efficacy in reducing polyps. They differ primarily in the propellant and additives rather than in the active ingredients. However, these factors may make one brand more acceptable than another to a given patient. Head positioning is important during intranasal corticosteroid use to direct the medication in contact with the polypoid mucosa. Since most recurrent polyps after surgery occur in the frontal recess, the head-inverted position is necessary to obtain the maximal benefit in the localized area of disease.

The use of oral corticosteroids in polyposis and sinusitis is somewhat controversial because of the myriad side effects and complications associated with them. However, the effectiveness of oral corticosteroids is remarkable, sometimes with even a short taper. In some patients, recurrence of nasal polyps can be prevented with very low doses given for prolonged periods of time. The dose of corticosteroids prescribed is dependent on the extent of disease. Typically, for moderate to severe polyposis, the upper end of the taper is 30 to 40 mg of prednisone per day for 3 to 4 days. A similar dose is used preoperatively to maximize shrinkage of the polyps and to reduce the mucosal reactivity and vascularity. Also, since asthma is a comorbid condition in many patients with polyps, the corticosteroids help to stabilize bronchial hyperreactivity preoperatively. Antibiotics are often used along with corticosteroids in treatment for polyps occurring secondary to an infectious chronic

sinusitis. Most patients with nasal polyps have some concurrent sinusitis at the time of presentation, and this is likely to be an exacerbating factor. Aggressive treatment of the underlying sinusitis will lead to more rapid resolution of inflammatory polyps.

Surgery for polyps is indicated when medical management fails to give symptomatic improvement or if complications develop. Also, in those polyp patients with asthma, surgical intervention may be considered when repeated sinus infections are causing a worsening of the lower airway disease. The operative technique of polyp removal is the same as other endoscopic sinus surgery and is described in detail later in this chapter. Powered instrumentation using a microdébrider is particularly useful in polyp surgery because it is rapid and spares the mucosa, while the continuous suction allows excellent visualization to be maintained in the face of bleeding. The postoperative monitoring of the sinus cavities is absolutely critical for the success of sinus surgery. Recurring disease should be addressed early in an office setting by débriding small polyps under topical anesthesia. Although complete resolution of mucosal abnormalities is rare, continued medical management and vigilance in follow-up make revision surgery unnecessary in the vast majority of patients. Successful treatment ultimately depends on a commitment by both the patient and the physician to an intensive postoperative course, which may be quite prolonged and involve multiple débridements.

COMPLICATIONS OF RHINOSINUSITIS

The complications of sinusitis can be divided broadly into those involving the orbits and those that involve the intracranial space. In the antibiotic era, such complications have become less commonplace, but they still have the potential for serious morbidity or even mortality. Awareness and early recognition of complications are necessary to minimize adverse sequelae. Fortunately, improved diagnostic modalities and advances in medical and surgical techniques have significantly reduced the risk of blindness or life-threatening intracranial infections.

Orbital complications of sinusitis occur because the orbit is bounded by the sinuses inferiorly, superiorly, and medially. This close proximity leads to early extension of infections that are not treated adequately or in the setting of impaired host immunity. Most orbital complications occur in young children, but those in older children and adults are typically more severe and necessitate surgery. Ethmoiditis most commonly leads to orbital involvement, followed by infections of the maxillary, frontal, and sphenoid sinuses. Infections of the ethmoid can directly erode the thin lamina papyracea or extend through suture lines or foramina into the orbit. The other mechanism of spread is hematogenously via retrograde thrombophlebitis of valveless veins. Chandler et al described a classification scheme for orbital complications of sinusitis in 1970.[21] The first stage, preseptal periorbital cellulitis, consists of eyelid swelling anterior to the orbital septum without involvement of the orbital contents (Figure 34–5, A and B). When the orbital soft tissue becomes involved, the result is orbital cellulitis, a diffuse process of inflammation without abscess formation. Patients with this complication are generally proptotic, with some degree of ophthalmoplegia and chemosis. When pus accumulates between the bone and the orbital periosteum, the result is a subperiosteal abscess (Figure 34–5, C and D). This will displace the orbit inferolaterally and may cause some proptosis. Unrecognized or untreated, the process can expand to cause extraocular muscle impairment, chemosis, and loss of visual acuity. Pus within the orbital tissue is an orbital abscess and can be intraconal or extraconal. In either case, marked proptosis, limitation of extraocular movement, and visual loss are commonly observed when an orbital abscess is present. Finally, the most serious complication is cavernous sinus thrombosis. This may result from extension of ethmoid or sphenoid sinusitis directly or via thrombophlebitis of the ophthalmic vein. Here, proptosis, chemosis, ophthalmoplegia, and decreasing visual acuity are the rule. In the early stages, fever, headache, periorbital edema, or photophobia may herald the onset of cavernous sinus thrombosis. Once it occurs, the process can extend to the opposite side, and bilateral findings are considered a diagnostic hallmark. Cranial neuropathies of nerves II to VI are seen, and pituitary insufficiency may occur. From the cavernous sinuses, the infection can spread rapidly through the dural sinuses, causing a range of intracranial complications.

Intravenous antibiotics are the mainstay of medical therapy and may be combined with topical

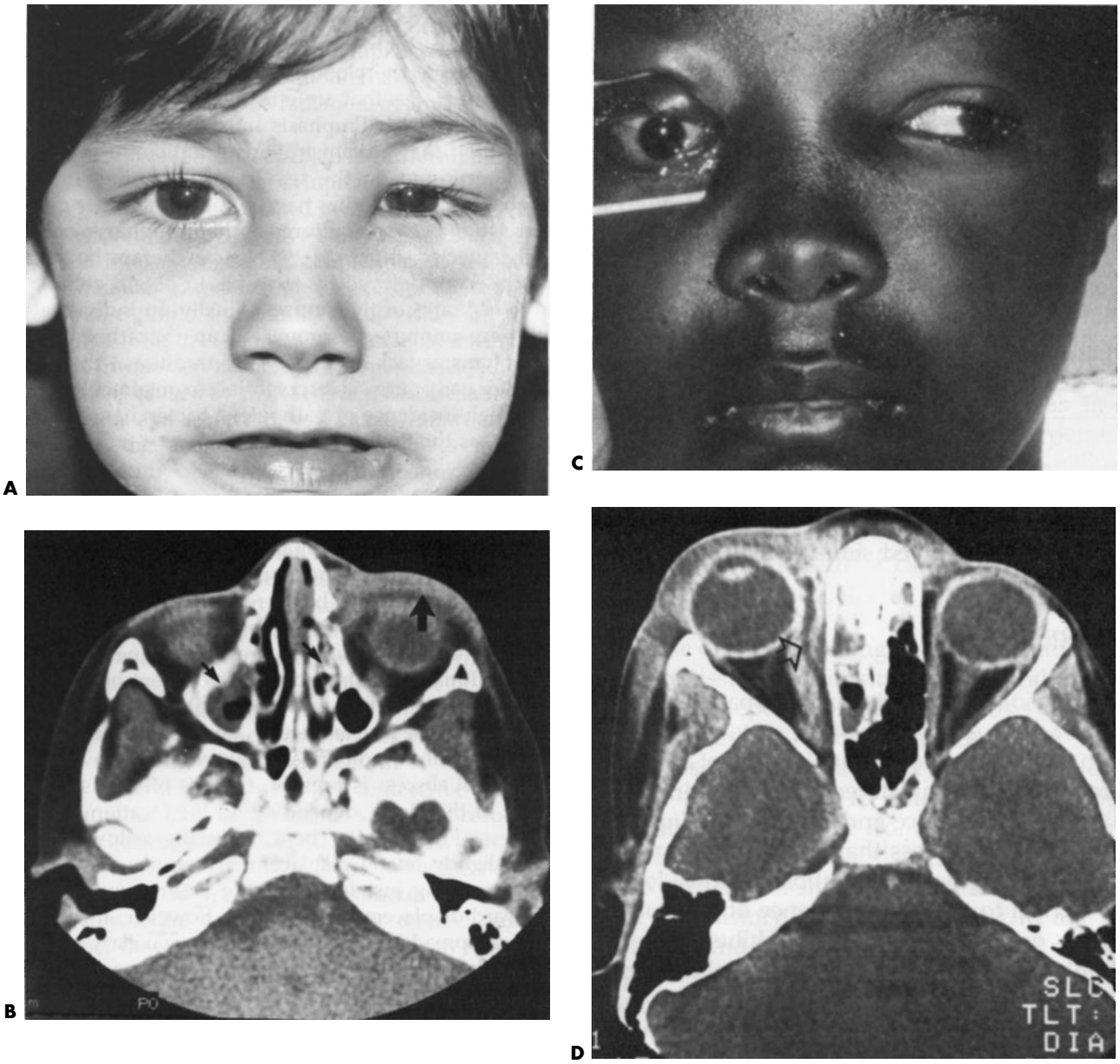


FIGURE 34–5. *A*, Preseptal cellulitis secondary to ethmoid sinusitis. *B*, A computed tomographic (CT) scan of child with a preseptal cellulitis. *Large arrow* points to swelling in the preseptal compartment; *small arrows* point to concomitant ethmoid and maxillary sinusitis. *C*, Right subperiosteal abscess causing ophthalmoplegia, proptosis, and chemosis. *D*, A CT scan of a right subperiosteal abscess. *Arrow* points to the abscess adjacent to the medial rectus muscle. Note swelling of the medial rectus muscle.

decongestants to promote sinus drainage. Corticosteroids are generally not recommended in the face of active infection. When orbital complications are suspected, a complete neuro-ophthalmologic examination is essential. However, even without funduscopic abnormalities, blindness can still result. Optic nerve injuries stem most often from pressure effects within the orbit. The damage may be caused by ischemia or be secondary to direct compression of the optic nerve. Also, permanent visual loss can result from irritation of the optic nerve by the adjacent infection. It is also important to obtain a radiologic study when evaluating a patient with orbital complications of sinusitis. Computed tomographic scans are helpful in identifying the involved sinuses and in defining the location of an abscess. Magnetic

resonance imaging is an excellent complementary study when intracranial extension needs to be evaluated. Serial CT scans may be used to monitor the progress of medical management and to help determine if surgery is needed.

The decision to proceed to surgery is made based on a number of factors and is individualized to the particular patient. Certainly, progressive visual loss demands aggressive management and drainage of the source of infection. Medical therapy is typically not effective once an abscess develops, and so surgical incision and drainage are required at that point. Often the extent of the infection is unclear, even with radiologic imaging, and the decision should be based on the clinical picture. Surgical intervention should be considered when there is disease progression after 24 hours of antibiotics or no improvement after 2 to 3 days of therapy. Ideally, surgery involves approaching both the orbital complication and underlying sinusitis simultaneously. Endoscopic approaches will generally be used to approach the sinuses and can even be employed in experienced hands to drain ethmoid subperiosteal abscesses. The classic technique for managing orbital complications is the external ethmoidectomy approach. Frontal sinus trephination may also be employed for acute frontal sinusitis. These techniques are described in detail later in this chapter.

Intracranial complications of sinusitis occur less frequently than orbital complications but are potentially life threatening if not recognized and treated. Most intracranial infections arise from the frontal sinus, although extension from the other sinuses is possible. The most frequent route of spread is retrograde thrombophlebitis via the diploic veins of Breschet in the posterior table of the frontal sinus. These valveless veins communicate directly with dural veins and empty into the sagittal sinus. The types of complications that may develop include osteomyelitis of the frontal bone, meningitis, epidural abscess, subdural empyema, and intracerebral abscess. Pott's puffy tumor is a well-circumscribed swelling of the forehead caused by anterior extension of frontal sinusitis. The edema of the skin and soft tissue overlies a collection of pus under the periosteum of the anterior table of the frontal sinus.

Abscesses in the epidural space often do not give localizing findings and do not give abnormal lumbar puncture results other than a high opening pressure. Headache and low-grade fever are common findings, and only when the abscess becomes enlarged do effects of increased intracranial pressure become evident. Epidural abscesses can extend by rupturing into the subdural space to create an empyema. Because of the barrier function of the arachnoid, subdural infections do not necessarily progress to meningitis. However, as the empyema spreads from anterior to posterior, inflammation of the leptomeninges will occur, resulting in significant edema. This edema can lead to increased intracranial pressure and eventually cortical ischemia. Early signs include headache, fever, and leukocytosis, whereas later findings may be much more severe such as hemiparesis, hemiplegia, or seizures. When the intracranial pressure becomes high enough, bradycardia, hypotension, and decreased mental status will ensue. Unless treated, transtentorial herniation can follow, causing death. Brain abscesses in the frontal lobe can also begin insidiously and then progress rapidly to fatal herniation. Abscesses in this location are frequently neurologically silent except for subtle mood or personality changes. If the abscess becomes large enough, or if the surrounding edema is extensive, symptoms of increased intracranial pressure may become evident. Rupture of the abscess into the ventricles will quickly lead to death.

When intracranial complications are suspected, the study of choice is a CT scan of the brain and sinuses with and without contrast. Magnetic resonance imaging is a more sensitive tool in the early stages of intracranial infection and can demonstrate enhancement of the dura in meningitis. Magnetic resonance imaging is also excellent for showing dural sinus thrombosis. Lumbar puncture may be helpful in making the diagnosis but needs to be performed with caution in the setting of a potentially increased intracranial pressure. Lumbar puncture is generally nondiagnostic in abscesses and subdural empyema but will clearly indicate meningitis if it is present.

The mainstay of therapy for suspected intracranial complications is intravenous antibiotics capable of crossing the blood-brain barrier. If cultures can be obtained from the affected sinuses, this will guide specific antibiotic choice. A neurosurgical consultation is sought when a procedure may be necessary to drain an intracranial collection. Since these complications often cause seizures, input from a neurologist

may be helpful in determining appropriate anticonvulsant medications. Corticosteroids are usually not used during an active infectious process; however, they are sometimes employed to reduce severe brain edema. Surgery should be directed at the involved sinuses as well as the intracranial process unless the patient's condition limits operative time, in which case, the neurosurgical procedure takes precedence. Epidural abscesses are drained via bur holes without violating the dura. Subdural empyema can be approached with either bur holes or a craniotomy. The morbidity from subdural empyema is high with either technique, and the mortality is between 12 and 18%.[22] Brain abscesses actually have a lower mortality, although the morbidity remains high. These collections can be either serially aspirated or excised, based largely on the accessibility of the lesion and the stability of the patient.

DIAGNOSIS OF RHINOSINUSITIS

History

The diagnosis of rhinosinusitis is made difficult by the similarity in symptomatology with allergic rhinitis and acute viral rhinitis. In rhinosinusitis, the most common complaints are nasal obstruction and nasal congestion. These sensations likely result from the thickening of the sinus and nasal mucosa, along with reactive swelling of the inferior and middle turbinates. Postnasal discharge is also a common symptom reported by sinusitis patients. These symptoms, which largely reflect inflammation of the nasal cavities, are present in allergies and colds as well as sinusitis. One factor that may help in the differentiation is the duration and timing of symptoms. Allergic rhinitis often has a seasonal component and is associated temporally with exposure to known allergens. Upper respiratory tract infections usually last less than 10 days or are at least on their way toward resolution by then. Both allergic rhinitis and colds can trigger rhinosinusitis and may therefore be present in the history prior to onset of sinus symptoms. It is estimated that 0.5 to 2% of upper respiratory tract viral infections go on to cause sinusitis.[23] A history of a "cold" that lasts longer than 10 days, especially with green nasal discharge, is suspicious for sinusitis.

Facial pain, pressure, or fullness can be a more localizing symptom of sinus disease that may help in its identification. Particularly in acute sinusitis, pain over the maxillary or frontal regions can be a prominent feature in the patient's history. Maxillary sinus pain may also be referred to the upper teeth and palate. Ethmoid sinusitis classically causes pain between or behind the eyes. Sphenoid sinus inflammation or infection tends to cause more insidious pain that may be referred to the occipital, vertex, or bitemporal regions of the skull. Of course, there are many other causes of headache and dental pain other than sinusitis; thus, facial or head pain is not a specific finding. Also, pain is a less common finding once sinusitis becomes chronic, except when there is an acute exacerbation. Similarly, acute sinusitis is sometimes associated with systemic symptoms such as malaise, fever, and lethargy, whereas chronic sinusitis typically is not. A common symptom seen in chronic disease more often than in the acute situation is olfactory loss. Sore throat and cough may be present in either case. As discussed earlier, the Rhinosinusitis Task Force of the AAO-HNS has identified major and minor symptoms associated with rhinosinusitis that can be helpful in making the diagnosis (see Table 34–2).

It is important in the history to elicit any symptoms related to orbital or intracranial extension of the infection. Mild periorbital swelling is not uncommon in frontal, ethmoid, or maxillary sinusitis. This swelling is often worse in the morning and improves over the course of the day. However, massive periorbital edema, orbital pain, diplopia, or change in vision may signal the presence of an orbital complication. Change in mental status or meningeal signs can herald an intracranial process related to ethmoid, frontal, or sphenoid sinusitis.

Physical Examination

The external findings in sinusitis may be limited and nonspecific. Periorbital, forehead, or cheek swelling is sometimes apparent, and there may also be associated tenderness to palpation or percussion in these regions. The oral cavity and oropharynx should be examined for dental pathology and for the presence of postnasal discharge. Anterior rhinoscopy can reveal mucosal hyperemia and edema of the septum and inferior turbinate. It may be possible to discern mucopurulent discharge in this manner, although the site of origin is not likely to be visualized. Transillumination of the maxillary and frontal sinuses is

sometimes described and can be of supportive help in diagnosis but has largely been supplanted by more specific techniques of examination.

In recent years, there has been tremendous advance in the use of nasal endoscopes for the diagnosis of nasal and sinus disease. Such endoscopes may be either rigid or of a flexible fiber-optic design. Pretreatment of the nasal cavities with decongestant and topical lidocaine allows a very thorough inspection in the office setting with minimal discomfort to the patient. Septal deformities, polyps, and other anatomic abnormalities impacting sinus drainage can be readily observed endoscopically. The middle meatus can be seen directly, and any purulent discharge may be traced either to the middle meatus or sphenoethmoidal recess. A swab can then be used to obtain this material under endoscopic control, yielding highly accurate cultures that may direct specific antibiotic therapy (Figure 34–6). The characteristics of the sinonasal discharge can give clues as to the nature of the underlying pathology. For example, clear, thin secretions are more typical of viral or allergic inflammation. Thick, opaque, rubbery secretions are often referred to as "allergic mucin" because of the histologic presence of numerous eosinophils and their granular contents. This type of secretion is often seen in chronic sinusitis and may be suggestive of AFRS in certain clinical settings.

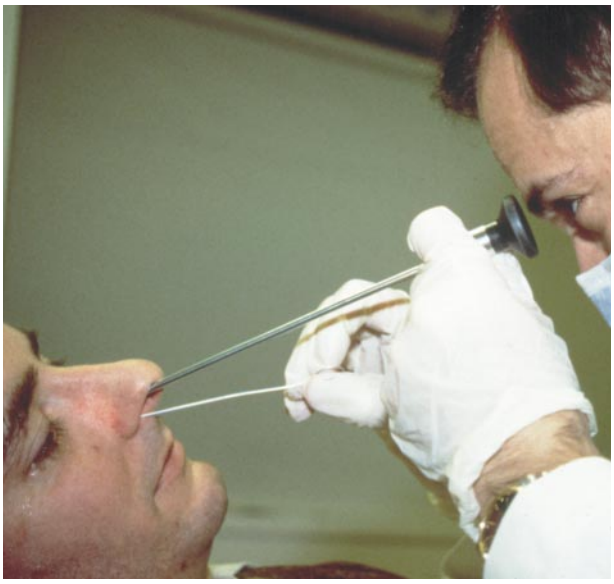


FIGURE 34–6. Technique of sinonasal endoscopy with direct culture of purulent discharge.

Radiology

Since rhinosinusitis is primarily a clinical diagnosis, and considering the improvements in diagnostics brought about by endoscopy, there is a limited role for radiologic studies in the initial assessment of suspected sinusitis. In earlier times, plain x-ray films were considered screening studies for acute disease and are fairly specific for maxillary sinusitis when air-fluid levels are seen. However, for chronic mucosal disease and for imaging of every other sinus, plain films are not useful and therefore are rarely obtained. Certainly, there are situations in which imaging studies are needed for evaluation of the sinuses: for example, clinical sinusitis that fails to resolve with medical therapy or persistent sinus symptoms without endoscopic evidence of disease. The imaging study of choice today is CT with fine coronal sections at the level of the ostiomeatal complex. This technique is excellent in assessing bony detail and thus provides an accurate road map for endoscopic sinus surgery. It is also sensitive in demonstrating mucosal thickening and revealing trapped secretions within the sinus cavities. Unfortunately, the mucosal changes seen by CT are not specific for sinusitis and thus should be interpreted cautiously. Viral respiratory tract infections and allergy will both cause mucosal thickening in the absence of infectious or chronic sinusitis. The mucosal changes associated with viral rhinitis can last for up to 2 weeks after the resolution of symptoms, and changes resulting from an acute bacterial sinusitis can last over a month. Indeed, about 40% of normal people without sinonasal complaints will have abnormalities of the sinus mucosa on CT scan that may be transient and not indicative of true disease.[24]

Radiologic imaging of the sinuses plays an important role when complications are suspected and in differentiating inflammatory from neoplastic disease. Computed tomography is useful in demonstrating bony dehiscences that may predispose a patient to intraoperative complications. When such defects are present in the skull base, they may be important in the intracranial spread of infections or be sites of origin for nasal encephaloceles that may be confused for polyps. Magnetic resonance imaging is a complementary study that is more sensitive than CT in showing soft tissue detail (Figure 34–7). It can

clearly demonstrate dural inflammation that would not be appreciable by CT and shows the communication of encephaloceles with the intracranial space. In invasive fungal sinusitis, MRI can reveal vascular abnormalities that represent fungal infiltration of neurovascular structures, sometimes allowing the process to be detected early when it is more treatable. In the setting of orbital complications of sinusitis or sinus surgery, CT is generally the better study, unless intracranial complications are suspected as well. Another use of MRI is to evaluate symptoms referable to the frontal sinus after it has been surgically obliterated and cannot be monitored endoscopically. Whereas CT merely shows opacification of the sinus, MRI can differentiate fat used in the obliteration from retained mucosa, secretions, and infection. The disadvantage of MRI is mainly that bony anatomy is not seen with this technique, making it useless as a surgical map. Also, to a greater extent than even CT, mucosal disease on MRI is difficult to interpret because of the high sensitivity and the presence of abnormalities in many asymptomatic people.

Staging Systems for Rhinosinusitis

In addition to the establishment of diagnostic criteria for rhinosinusitis, there have been several efforts in the last decade to formulate a standardized staging system to be applied across individuals and used

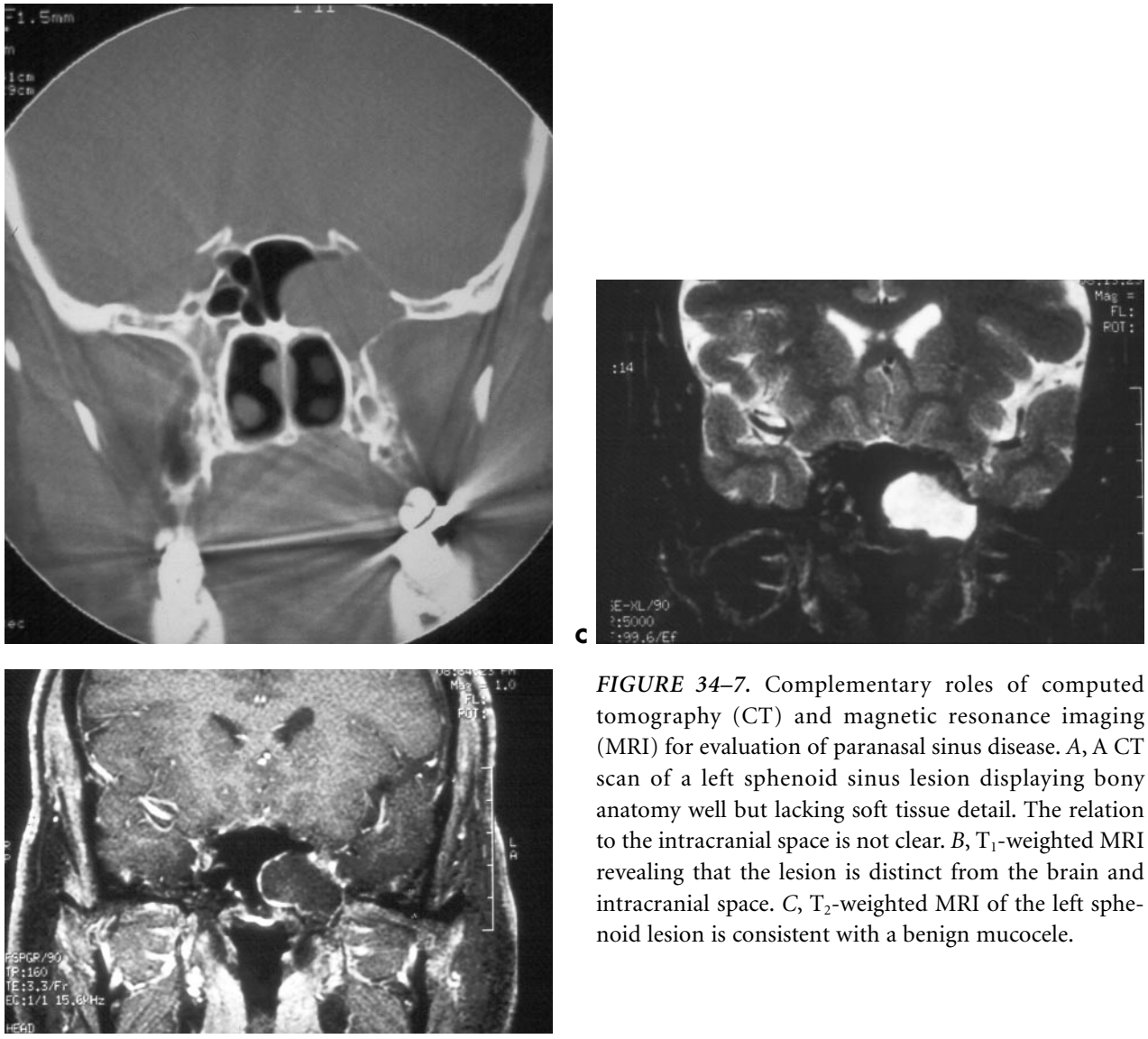


FIGURE 34–7. Complementary roles of computed tomography (CT) and magnetic resonance imaging (MRI) for evaluation of paranasal sinus disease. *A*, A CT scan of a left sphenoid sinus lesion displaying bony anatomy well but lacking soft tissue detail. The relation to the intracranial space is not clear. *B*, T_1-weighted MRI revealing that the lesion is distinct from the brain and intracranial space. *C*, T_2-weighted MRI of the left sphenoid lesion is consistent with a benign mucocele.

to follow the course of the disease. These staging systems have attempted to identify parameters of the history, physical examination, and sinus CT scan with prognostic significance. Most rely heavily on CT findings such as mucosal thickness, anatomic abnormalities, and the location of involved sinuses because these are relatively objective criteria that are readily quantifiable. Unfortunately, the radiographic appearance can be misleading when viewed as an isolated snapshot in time and generally cannot identify the underlying pathophysiology. Better staging information can be obtained by incorporating nasal endoscopy and intraoperative findings into the paradigm. The modified Lund-Mackay System is the staging system for rhinosinusitis currently recommended by the AAO-HNS.[25] This system includes patient symptoms as well as scores derived from radiographic, demographic, endoscopic, and operative findings. Since it is comprehensive and relatively simple to use, the modified Lund-Mackay System may help to unify the various staging techniques that have been described in the literature. Future modifications may involve the inclusion of any biochemical or molecular markers that are demonstrated to have prognostic significance in rhinosinusitis.

PEDIATRIC RHINOSINUSITIS

Pediatric sinus disease is gaining increasing recognition in terms of its prevalence and economic impact. As in adults, the diagnosis of rhinosinusitis in children is made largely on clinical grounds and is similarly difficult to distinguish from viral upper respiratory tract infections. The principal complaints in pediatric sinus disease are nasal congestion, cough, and purulent rhinorrhea. Unfortunately, a good history and physical examination may be unobtainable, especially in very young children. Usually, an otoscope is the most effective instrument to examine the nose of a young child. To the extent possible, attention should be directed toward viewing the region of the middle meatus to look for purulence. Radiologic studies such as CT may be helpful where the physical examination is limited, although the same problems of interpretation exist as do in adults. The pathophysiology of sinus disease in children mirrors that of the adult population, with anatomic, local, and host factors playing important parts. In a child with nasal polyps and recurrent sinusitis, the diagnosis of cystic fibrosis must be strongly entertained. Unique issues in children that may play a causative role in the pathogenesis of sinusitis include adenoid hypertrophy and gastroesophageal reflux disease. There is good evidence that removal of large, obstructive adenoids is beneficial for chronic pediatric sinusitis. However, removal of nonhypertrophied adenoids is controversial for this indication. Antireflux therapy for sinusitis is also somewhat controversial in the absence of other overt manifestations. The mainstay of the medical treatment of pediatric rhinosinusitis is antibiotic therapy directed against the typical sinus pathogens. This is discussed further in the next section as it pertains also to the adult disease.

TREATMENT

Medical Therapy

Acute Rhinosinusitis Once the diagnosis of acute sinusitis is ensured, the goal of therapy is to prevent disease progression and the possibility of serious sequelae. Without treatment, acute sinusitis is often self-limiting, with approximately 40% resolving spontaneously.[26] However, prompt treatment is believed to hasten the resolution of tissue edema and bacterial contamination, restoring ostial patency and sinus ventilation before permanent mucosal damage occurs. There is no debate that antibiotics should be instituted for all cases of acute sinusitis. In three of four randomized trials reviewed by Hueston et al, the use of antibiotics was supported by a more rapid resolution of symptoms.[27] Historically, there has been a dramatic reduction in the incidence of complications secondary to sinusitis since the introduction of antibiotics.

The selection of first-line antibiotics for acute sinusitis is directed by the knowledge of the most common pathogens. Typically, patients will be seen initially by their primary care physician and treated with inexpensive agents such as amoxicillin or trimethoprim-sulfamethoxazole (TMP-SMX). The effectiveness of these agents is hampered by emerging patterns of resistance in many geographic areas, however. Resistance to TMP, by *Streptococcus pneumoniae* in particular, has increased in recent years and has been reported to be greater than to penicillin. For amoxicillin, recent in vitro evidence suggests that current doses may be inadequate, although clinical trials at higher doses have not been completed.[28] In

general, these first-line agents should not be used when resistant microorganisms are suspected.

The choice of a second-line antibiotic is dependent on a number of variables including patient allergies, dosing schedule, proven efficacy, physician experience, and the patient's previous response history, as well as resistance patterns in the community. β-Lactam cephalosporins have long been the most common second-line agents, although macrolides and fluoroquinolones have recently been increasing in popularity. The use of penicillin, erythromycin, cephalexin, cefixime, and tetracycline is condemned for their limited spectrum of activity. In addition, first-generation cephalosporins have poor *Haemophilus influenzae* coverage and should be used cautiously. The most relevant pharmacokinetic parameter of β-lactam antibiotics is the time above minimal inhibitory concentration, which has been shown to correlate with efficacy. In this regard, cefprozil, cefuroxime, and cefpodoxime are the most consistently reliable agents of this class against penicillin-sensitive and intermediately penicillin-resistant strains. These antibiotics are effective in twice-daily dosing and have an adequate spectrum of activity. For β-lactamase-producing strains of *H. influenzae* or *Moraxella catarrhalis*, the use of amoxicillin-clavulanate may be indicated. In this case, a separate prescription for amoxicillin alone must be written to give a double dose of that drug without increasing the clavulanate dose. It should be noted that clavanulate has no effect on resistant *S. pneumoniae* microorganisms because their mechanism of resistance is different. The newer macrolides such as azithromycin and clarithromicin may be acceptable second-line agents, especially in penicillin-allergic patients, although pneumococcal resistance to these agents is on the rise. Any microorganism that is resistant to erythromycin will also be resistant to these newer agents. Finally, the newer fluoroquinolones such as levofloxacin, moxifloxacin, and gatifloxacin are good second-line agents for patients with activity against *S. pneumoniae*. However, as has been observed with other classes of antibiotics, resistance will develop if these drugs are used inappropriately; thus, they are recommended only for moderate-to-severe infections or treatment failures.

Once antibiotics have been instituted, the duration of therapy is controversial. Symptoms should begin to improve within 48 to 72 hours, and it is important to maintain appropriate follow-up to ensure that the complete course of antibiotics is taken. A good guideline is a 10- to 14-day course of therapy, which can be lengthened for persistent symptoms.

Ancillary Therapy. A variety of therapeutic measures can augment the effectiveness of antibiotics in the treatment of acute sinusitis. The goal of these interventions is to restore proper nasal function through improvement in ciliary function and reduction of mucosal edema. Many simple, inexpensive supportive measures are effective because they help to clear crusts and thick mucus. Examples include nasal saline sprays, humidifiers (warm or cool), steam, hot soup, or tea. Mucolytic agents such as guaifenesin also are useful because they lead to thinning of the mucus, which promotes clearance and prevents stasis.[29] Systemic and topical decongestants act on α-adrenergic receptors to cause vasoconstriction and reduction of edema and are therefore appropriate to relieve nasal obstruction, re-establish ostial patency, and ventilate the sinuses. Topical decongestants have minimal systemic side effects and are rapid in onset. However, use of these agents for longer than 3 consecutive days can result in rebound congestion and rhinitis medicamentosa. Systemic decongestants can be used for longer periods of time but may be associated with insomnia, hyperactivity, cardiac stimulation, worsening of hypertension, and negative interactions with concurrent medications. For these reasons, systemic decongestants should be prescribed with caution in some patient groups. The use of decongestants in conjunction with antibiotics has been shown to be more effective than antibiotics alone.

Antihistamines have been used empirically in patients with sinusitis and allergy, although no studies show a clearly beneficial role for these medications. In the setting of acute infectious sinusitis, first-generation antihistamines may actually be counterproductive because of their anticholinergic side effects of mucosal dryness, crusting, and increased mucus viscosity. The newer second-generation antihistamines cause less of these undesirable changes and therefore may be suitable in cases of sinusitis in which allergy plays an important role. Topical nasal corticosteroids are also excellent for the management of allergic rhinosinusistis but have no proven value in the treatment of acute infectious sinusitis. Systemic corticosteroids are

potent anti-inflammatory agents that are effective in the treatment of allergic disease as well as other inflammatory conditions of the nose. Because corticosteroids reduce tissue edema and inhibit inflammatory mediator production, they almost certainly are beneficial in the treatment of acute sinusitis. However, a number of potential complications are associated with corticosteroid use, including hypothalamic-pituitary-adrenal axis suppression, gastric ulcers, psychiatric changes, sleep disturbances, and exacerbation of diabetes, as well as long-term sequelae such as osteoporosis, weight gain, ocular problems, and hypertension. For these reasons, the use of corticosteroids is not widely accepted for acute sinusitis. In general, the risk of adverse side effects when corticosteroids are used conservatively over short, tapered doses is minimal; therefore, these drugs are reasonable adjunctive therapy for acute sinusitis treated primarily with antibiotics.

Chronic Rhinosinusitis In chronic sinusitis, the microorganisms primarily involved are coagulase-positive and coagulase-negative species of *Staphylococcus* and *Streptococcus*. Antibiotic therapy should therefore be directed at these pathogens, although resistance is a constant problem. The use of antibiotics with activity against anaerobes is controversial since some studies show these microorganisms to be of minor importance; however, these agents can be used alone or in combination with other antibiotics when anaerobes are identified in sinus fluid. The duration of antibiotic therapy in chronic sinusitis is not clearly defined but is typically on the order of 4 to 8 weeks. The goals of ancillary therapy for chronic sinus disease are similar to those in the acute situation; however, there are some important differences. When an underlying condition such as allergy, polyposis, fungal infection, or systemic disease is present, the treatment must first be directed toward controlling these processes. In the case of allergic rhinosinusitis, management with antihistamines, topical nasal corticosteroids, and immunotherapy will be of more value than it would in an acute infection. Likewise, systemic corticosteroids are indispensable for the treatment of polyps and sinus inflammation caused by systemic granulomatous or autoimmune diseases. In patients with chronic sinusitis who have undergone previous sinus surgery, there is access for direct irrigation of the sinus cavities with topical antibiotic solutions. This can also be performed with greater difficulty in unoperated patients using indwelling catheters into the maxillary sinuses. There is some evidence to suggest a role for this strategy in chronic sinus disease failing systemic antibiotic therapy. As compared to acute sinusitis, the mechanisms of chronic rhinosinusitis are more obscure and seem to involve many factors other than infection. In fact, much of the infectious component of chronic disease may be secondary to the compromise of the normal mucosal barrier mechanisms by long-term inflammation. In the future, new therapies may target elements of the ongoing inflammatory response in an attempt to allow healing to occur and restore normal mucosal function.

Surgical Therapy

Endoscopic sinus surgery More than 100 years have passed since the first nasal endoscopy was performed by Hirschmann, who used a modified cystoscope. Since then, major advances in optics, biomechanics, and radiographic imaging have allowed evaluation and treatment of paranasal sinus disorders with greater precision.

Naumann recognized the relationship between the middle meatal-anterior ethmoid complex, termed the ostiomeatal unit (Figure 34–8), and the

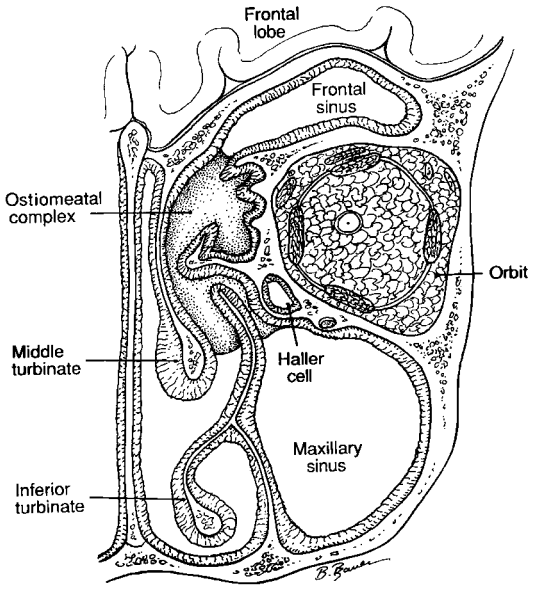


FIGURE 34–8. The ostiomeatal unit (*shaded*). Obstruction of this area leads to disease within the maxillary and frontal sinuses.

pathogenesis of maxillary and frontal sinus disease.[30] After decades of painstaking work, Messerklinger demonstrated that relieving the ostiomeatal unit of obstruction and inflammation could reverse mucosal disease within the frontal and maxillary sinuses, thus rejecting the "irreversibility" theory.[31] Functional endoscopic sinus surgery (FESS) is the natural extension of his labors. Through the work of Kennedy and associates, Stammberger and others advanced endoscopic techniques that are now used to treat an extended list of medical problems in addition to inflammatory processes of the paranasal sinuses (Table 34–3).[32–34]

Even though visualization and instrumentation have improved, endoscopic sinus surgery carries the same risks as traditional sinus surgery. The use of endoscopes has not decreased the complication rate. Surgeons performing endoscopic sinus surgery must understand the anatomy, physiology, technique, and complications related to this surgical procedure.

Technique. The instruments necessary for a safe and complete operation are pictured in Figure 34–9. The majority of the dissection is carried out using the 0-degree endoscope because the angulation of the other telescopes can be disorienting. The 30-degree endoscope is usually required for examination and manipulation of the maxillary sinus ostium and frontal recess. The 70-degree endoscope is excellent for viewing the antrum and for identification of an anterior frontal sinus. Attaching a commercially available endoscope washer facilitates surgery by defogging and clearing secretions from the tip of the telescope.

TABLE 34–3. Possible Indications for Functional Endoscopic Sinus Surgery

Recurrent acute sinusitis
Chronic sinusitis
Nasal polyposis
Fungal sinusitis
Barosinusitis
Advanced techniques
Tumor removal
Antral choanal polyp removal
Dacryocystorhinostomy
Encephalocele repair
Cerebrospinal fluid leak repair
Pituitary surgery
Mucocele removal
Orbital abscess/cellulitis management
Orbital decompression
Septal reconstruction
Choanal atresia repair
Epistaxis control

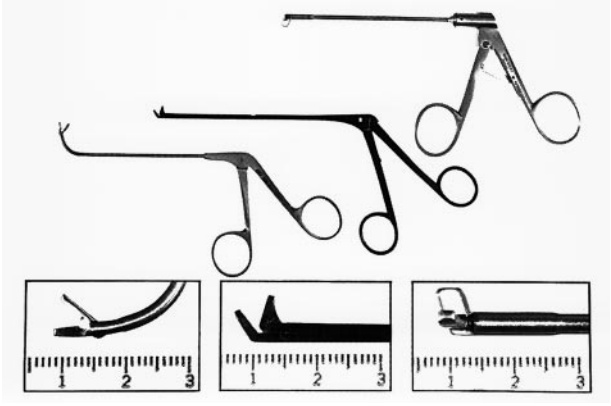


FIGURE 34–9. Micro-through-cutting instruments with varying angles allow mucosa-sparing dissection throughout all regions of the paranasal sinuses. The Stammberger down-biting punch is excellent for enlarging the maxillary ostium.

An important principle in endoscopic sinus surgery is the preservation of mucosa wherever technically feasible. To this end, there have been many improvements in the design of forceps to promote precise removal of tissue without stripping of sinus mucosa. Use of such "through-cutting" instruments greatly facilitates rapid healing and reduces postoperative care needs. Another important innovation has been the introduction of powered instrumentation to endoscopic sinus surgery. In 1994, Setliff first described the application of microarthroscopy-type microdébriders for removal of diseased nasal and sinus tissue. These instruments use an oscillating blade, accompanied by suction and irrigation, to shave away targeted soft tissue while sparing adjacent mucosa (Figure 34–10).[35] The chief advantages of this technique are the continuous suction and a largely bloodless operative field. Powered instrumentation has become widely accepted for the removal of polyps and soft tissue masses. The utility of this technology has been limited by an inability to remove effectively the bony partitions that contribute to the pathogenesis of sinusitis. For this rea-

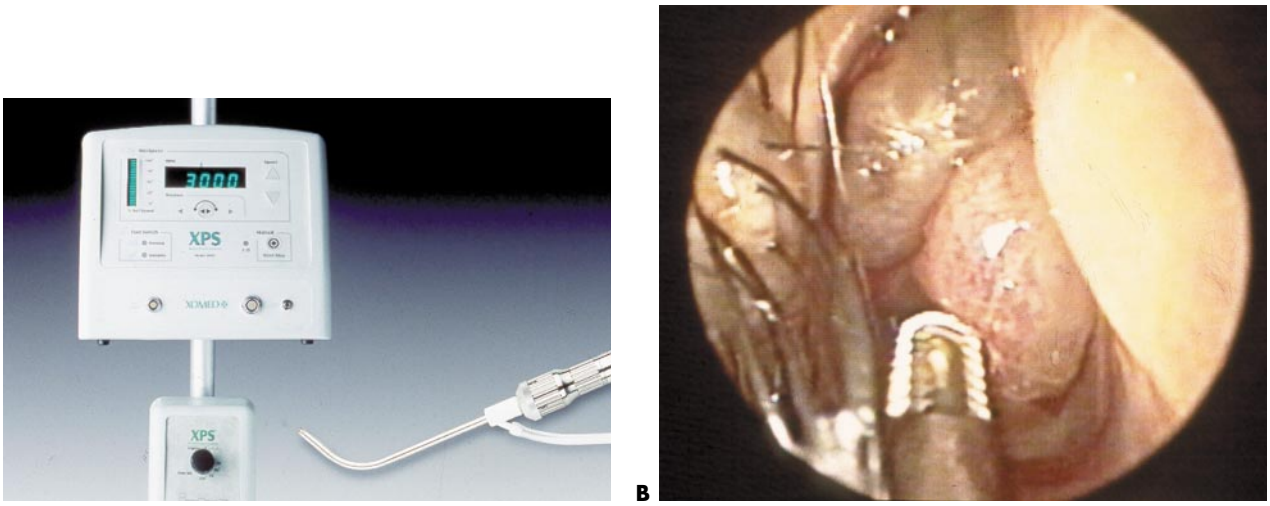


FIGURE 34–10. *A*, Microdébrider used for powered dissection in endoscopic sinus surgery. *B*, Intraoperative endoscopic view of a microdébrider being used to remove nasal polyps.

son, microdébriders and manually operated forceps serve complementary roles and are generally used concurrently in FESS procedures.

Another significant advance in endoscopic sinus surgery has been the development of computer-assisted image-guided technology. Accurate three-dimensional CT reconstruction with real-time intraoperative computer guidance has become widely available in recent years. There are several devices available today that use either electromagnetic or infrared tracking and are relatively user friendly. With some systems, the patient must undergo a preoperative CT wearing a specially fitted headpiece that is worn again in the operating room. The computer uses the headpiece to determine head position and correlate this with the preoperative CT. In other systems, anatomic landmarks are selected from the CT scan and identified on the patient in the operating room, instead of making use of a headpiece. Probes and instruments are then tracked with an infrared or electromagnetic system to determine their position continuously. The surgeon views the sinus CT in multiple planes on the monitor, with the instrument tip position displayed in real time (Figure 34–11). The accuracy of the system is on the order of 2 mm currently, although the technology is constantly undergoing refinement and improvement. Overall, image-guided surgery is helpful in all sinus cases but especially valuable for difficult revision cases with few landmarks or where unusual or distorted anatomy is present.

Anesthesia The operation can be performed under local or general anesthesia. Local anesthesia with intravenous sedation may be preferable because sensory information remains intact along the periorbita and skull base. Allowing the patient to listen to favorite music during surgery reduces the amount of anxiolytics necessary during the procedure. On the other hand, general anesthesia is preferable for patients who are very anxious about being awake and for those who become disinhibited under sedation. There has been a general tendency toward performing more endoscopic sinus surgery under general anesthesia in recent years. One reason for this is that the meticulous nature of mucosa-preserving sinus surgery has increased the length of the procedure to the point that it may be difficult for patients to remain still comfortably. Also, the advent of image-guided surgery has made it necessary for the patient to wear a tight-fitting headpiece, which is not tolerable for long periods of time. It is reasonable to present the advantages and disadvantages of both anesthetic approaches to patients and let them decide which they prefer preoperatively. Patients electing local anesthesia must be informed that they will be put to sleep if they cannot tolerate being awake for the duration of the procedure or if they become disinhibited with sedation.

Before the patient enters the operating suite, oxymetazoline hydrochloride nasal spray is used for initial mucosal shrinking. The nasal mucosa is then further decongested and anesthetized, one side at a

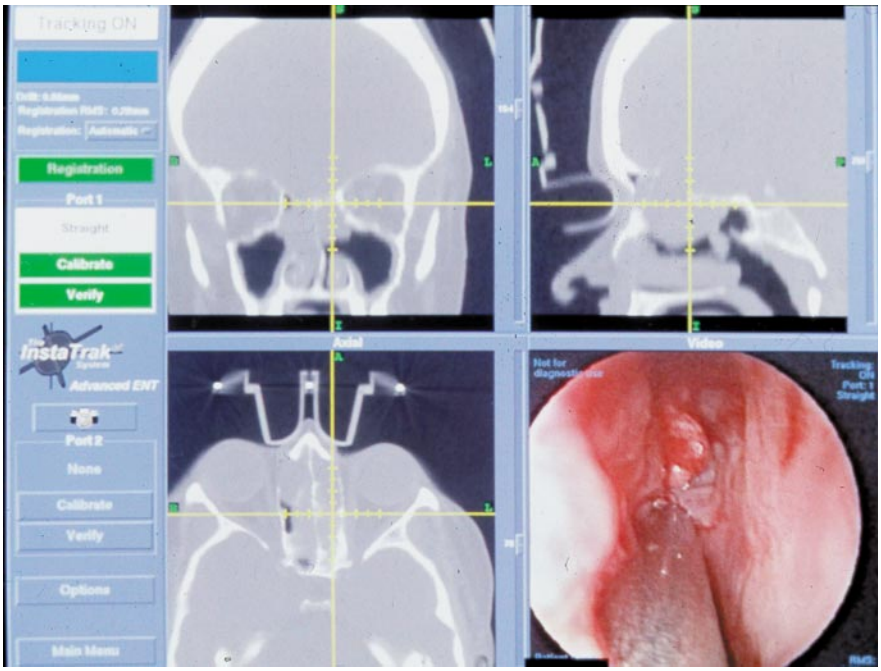


FIGURE 34–11. Use of computer-assisted surgical navigation in endoscopic sinus surgery. In this case, an area suspicious for preexisting dural exposure is identified intraoperatively. The location of the probe is shown at the crosshairs on the triplanar computed tomographic scan images in real time. A skull base defect is verified radiographically and correlated with the endoscopic appearance.

time, with 150 mg of topical cocaine followed by injections of 1% lidocaine with 1:100,000 epinephrine at the sphenopalatine neurovascular bundle (Figure 34–12). Septal abrasions should be avoided; otherwise, visualization becomes difficult from repeated soiling of the telescope. In addition, the sphenopalatine ganglion is anesthetized through the nose or through the greater palatine foramen. Allowing at least 10 minutes for maximal vasoconstriction decreases bleeding. The second side of the nose is anesthetized when surgery on the first side is near completion.

PROCEDURE The surgeon should be comfortable and relaxed during the dissection. The preoperative CT scan should be readily available in the operating room for repeated viewing. The surgeon must perform a meticulous dissection with adequate visualization if complications are to be avoided.

To begin, an infundibulotomy is made using a sickle knife along the anterior and inferior edges of the uncinate process. The uncinate is grasped with Blakesley forceps and removed, taking care not to strip mucosa from the lateral nasal wall. Gentle medial displacement of the uncinate during the

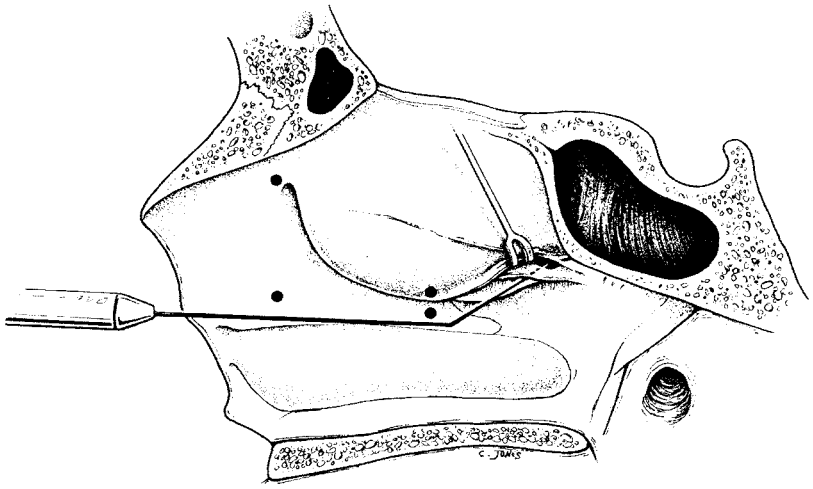


FIGURE 34–12. Recommended injection sites for endoscopic sinus surgery. An angled tonsil needle is used for intranasal injection of the sphenopalatine neurovascular bundle.

infundibulotomy helps define the line of the incision. If possible, the location of the maxillary sinus ostium is noted along with any accessory ostia. Extent of surgery is based on the preoperative assessment; however, for purposes of instruction, a description of a sphenoethmoidectomy with frontal sinusotomy and antrostomy follows.

The bulla ethmoidalis is entered and removed. The medial orbital wall should be identified as early as possible during the dissection. Every attempt at preserving the mucosa along the medial orbital wall is made to improve and hasten postoperative healing. The posterior ethmoid is entered through the inferior and medial area of the basal lamella. The skull base is most easily identified within the posterior ethmoidectomy and serves as the superior boundary for the dissection. The safest way to remove bony partitions within the ethmoid is to feel behind the partitions with up-biting forceps prior to removal (Figure 34–13). This technique assists in avoiding orbital or intracranial penetration.

The sphenoid ostium is located and, if necessary, widened beginning inferiorly and medially. The ostium is usually a few millimeters from the posteroinferior edge of the superior turbinate. The free edge of the superior turbinate may be difficult to appreciate when working in the ethmoid sinus but can be palpated medially. Alternatively, placing an endoscope medial to the middle turbinate usually brings the free edge of the superior turbinate in view. The sphenoid sinus can also be entered by amputating the inferior aspect of the superior turbinate. Care must be taken not to remove this turbinate completely because olfactory neuroepithelium could be injured. Absolutely no blind tissue removal is performed within the sphenoid because this may lead to optic nerve or carotid artery injury.

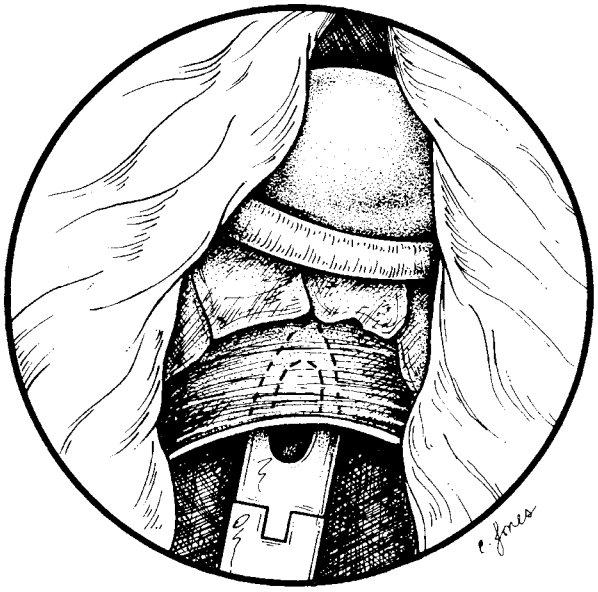


FIGURE 34–13. Blakesley forceps are used to palpate behind bony partitions before removal.

At this point, the dissection is continued anteriorly along the skull base, with care not to injure the anterior and posterior ethmoid arteries. The mucosa along the skull base is also preserved. The frontal recess is dissected only if required because manipulation of this region can lead to stenosis of the frontal ostium. In addition, frontal recess surgery increases the requirement for postoperative care.

Operating within the frontal recess is the most challenging segment of the operation. Changing to the 30- or 45-degree endoscope is helpful for dissecting this area. The dome of the ethmoid should be identified and followed anteriorly as the skull base ascends. The skull base represents the posterior limit of dissection in this area. All motions should be made away from the skull base in an anterior direction. Anteromedial ethmoid cells and agger cells can be removed with a frontal recess curette. Eicken curved suction is helpful in this area. The size of the frontal ostium should be recorded for postoperative reference.

Finally, attention is turned toward the maxillary sinus ostium. The 30-degree telescope is usually required to identify the natural ostium of the maxillary sinus. If necessary, the ostium is enlarged posteriorly and inferiorly. Beginning from within the ostium, the mucosa of the posterior fontanelle is split and then removed with the aid of a back-biting instrument. The remaining uncinate is removed at this time to prevent postoperative scarring in this area.

At the conclusion of the operation, an expandable sponge covered with a water-soluble antibacterial ointment is placed lateral to the middle turbinate. Additional packing is rarely required. Most patients are released the day of surgery and return to the clinic on the first postoperative day for removal of the sponge.

Postoperative Care The overall success of FESS is in large part attributable to appropriate postopera-

tive care. The goal during this period is to promote mucosal generation within the sinus cavities. To facilitate this generation, sinus cavities are examined and cleaned of blood, fibrin clots, and crusts at regular intervals following surgery. Endoscopes should be used, as opposed to blind manipulation, to prevent injury to healthy mucosa and for directed débridement. This is especially true for the frontal recess.

Systemic antibiotics are used in the postoperative period, and their use may be prolonged when severe, chronic inflammation is present. Agents covering the typical sinus flora, such as amoxicillin/clavulanate, cephalosporins, or quinolones, are the most common antibiotics chosen in uncomplicated cases. When there is known exposed osteitic bone, broader coverage with clindamycin and TMP-SMX may be indicated. If postoperative endoscopy reveals purulence, this material should be cultured and antibiotic therapy tailored appropriately. Patients may be instructed to irrigate the nose with saline solution twice daily. For the majority of patients with chronic sinusitis, surgery alone does not result in a permanent disease resolution. Long-term, culture-directed, systemic antibiotics and prolonged use of topical nasal corticosteroids are frequently required. Occasionally, antibiotics are added to the nasal irrigation solution for patients with recalcitrant chronic sinusitis. In some circumstances, such as in aspirin-sensitivity triad disease failing multiple previous sinus surgeries, long-term oral corticosteroid therapy using a low-dose, alternating-day regimen may be required to prevent recurrence.

It is critical to the ultimate success of endoscopic sinus surgery that meticulous, sometimes aggressive, débridement be performed in the weeks following the procedure. At each postoperative visit, the nasal cavities are sprayed with tetracaine and a nasal decongestant, and cocaine may be applied with nasal applicators as needed. Since the mucosa is particularly sensitive in the early postoperative period, the patient is instructed to take an oral narcotic/analgesic before arriving. The goal of these postoperative débridements is to clear crusts, osteitic bone fragments, and forming scar tissue before these factors create persistent inflammation and disease recurrence. In some instances, diseased sinus cells that were missed in the operation will be recognized and need to be opened. As in the operating room, mucosal preservation is paramount in postoperative débridement and sinus cavity revision. Local anesthetic injection may be necessary to allow the patient to tolerate the intranasal manipulation comfortably. Endoscopic surveillance and proactive treatment of residual disease eliminate the need for most revision surgery and lead to long-term success.

Open Sinus Procedures Although FESS can be used effectively for most medically recalcitrant sinus disease, there are still occasional circumstances under which open sinus procedures may be indicated.

Maxillary Sinus. Caldwell-Luc Operation. In cases of symptomatic maxillary sinus mucoceles, antrochoanal polyps, mycetoma, or foreign bodies not accessible via an intranasal endoscopic approach, the traditional open procedure has long been the Caldwell-Luc procedure. This procedure may be performed under general anesthesia or locoregional anesthesia via blocks of the infraorbital, sphenopalatine, and posterosuperior alveolar nerve. Topical 4% cocaine can be placed intranasally in the region of the anterior ethmoid nerve and the medial wall of the maxillary sinus.

Technique. The procedure begins with a gingivobuccal incision, made from the second molar to the ipsilateral canine tooth (Figure 34–14). The incision should be made just above the apex of the buccogingival sulcus to preserve an inferior flap of soft tissue that will facilitate closure. Dissection proceeds sharply through the submucosal tissue and the maxillary periostium down to bone. A periostial elevator is used to raise the periosteum and overlying soft tissue superiorly to the level of the infraorbital foramen, being careful not to injure the neurovascular bundle. Medially, the periosteum is elevated to the pyriform aperture. Once adequate exposure is achieved, the maxillary sinus in entered through its anterior wall using an osteotome or cutting bur. The antrum is entered superior to the roots of the canine and premolar teeth. A sphenoid punch and a Kerrison rongeur are used to remove additional bone and enhance exposure. Bone can be removed superiorly, medial and lateral to the infraorbital nerve, preserving the infraorbital foramen itself.

The approach gives good visualization of all portions of the maxillary sinus and allows complete removal of infectious material and masses that

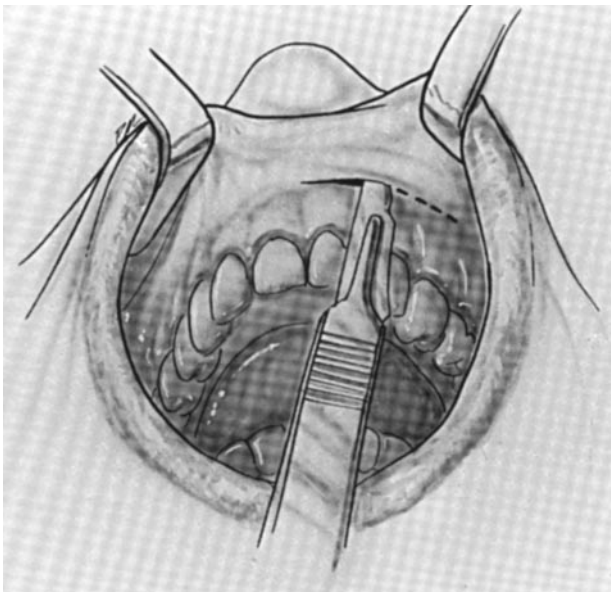


FIGURE 34–14. Mucosal incision for a Caldwell-Luc procedure.

would not be easily reached through a middle meatus antrostomy. Stripping of the mucosa with forceps or curettes is to be condemned and will result in permanent sinus dysfunction. Furthermore, creation of an inferior meatus antrostomy, which is typically performed in conjunction with the Caldwell-Luc procedure, is nonphysiologic in terms of mucociliary clearance patterns and may result in recirculation of purulent material. At the conclusion of the maxillary sinus procedure, the gingivobuccal incision is closed with either interrupted or running absorbable suture.

Complications The most common complication of the Caldwell-Luc approach is paresthesia or anesthesia of the cheek, teeth, and gingiva secondary to traction on the infraorbital nerve. Fortunately, this is self-limited in the majority of cases and resolves in less than 6 months. However, as many as 30% of patients will report some degree of persistent dysesthesia in the infraorbital nerve distribution.[36] Other complications include hematoma formation, maxillary tooth root injury, and superior alveolar nerve damage. There have also been rare reports of oroantral fistula and epiphora from nasolacrimal duct trauma after the Caldwell-Luc operation. It should be noted that the Caldwell-Luc procedure generally results in fibrosis and new bone formation in the maxillary sinus, leading to antral contracture. This can affect interpretation of postoperative films and make revision surgery more difficult.

Ethmoid Sinuses. External Ethmoidectomy. The external ethmoidectomy has largely been supplanted by the endoscopic approach; however, there remain some situations in which this approach may prove useful. For example, excellent exposure may be provided externally for biopsies of certain orbital lesions or lesions of the ethmoid or frontal sinuses. Also, this approach may be employed for rapid and safe access for orbital complications of acute ethmoid or frontal sinusitis. Repair of cerebrospinal fluid leaks originating from the cribriform plate or ethmoid roof may be addressed via an external approach, depending on the preference of the surgeon.

Technique The external ethmoidectomy incision generally begins at the inferior margin of the medial aspect of the eyebrow, curving gently downward midway between the medial canthus and the anterior aspect of the nasal bones (Figure 34–15). The incision is carried successively through the skin, subcutaneous tissues, and periosteum. Bipolar cautery is used to control bleeding from the angular vessels, and care is taken to preserve the supraorbital bundle. The periosteum is elevated both medially and laterally to facilitate eventual closure. Attention is given to protecting the trochlea and the attachment of the medial canthal ligament during periosteal elevation. The lacrimal sac is elevated atraumatically from the lacrimal fossa.

Dissection proceeds in the subperiosteal plane beyond the posterior lacrimal crest, exposing the medial wall of the orbit. The anterior ethmoid neurovascular bundle is encountered approximately 24 mm posterior to the anterior lacrimal crest. The anterior ethmoid artery, found in the frontoethmoid suture line, is clipped or electrocoagulated and divided. A malleable retractor is placed laterally for exposure and to protect the periorbita. Inadvertent entry into the periorbita will result in herniation of orbital fat into the field. Although it is not routinely necessary to dissect further posteriorly, the posterior ethmoid artery is located approximately 12 mm posterior to the anterior ethmoid artery and approximately 6 mm anterior to the optic foramen. The surgeon must be extremely careful in ligating the posterior ethmoid artery as the exact distance to the optic nerve is variable.

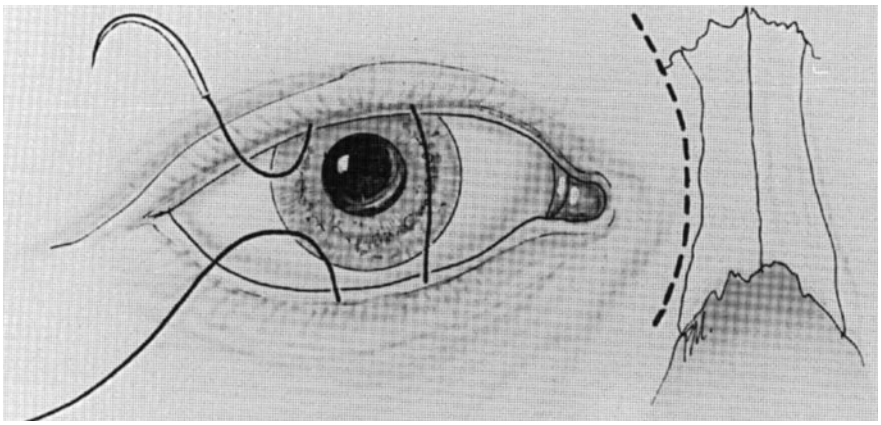


FIGURE 34–15. Incision for external ethmoidectomy.

Once the lacrimal bone, frontal process of the maxilla, lamina papyracea, and orbital process of frontal bone have been widely exposed, the ethmoid is entered through the lacrimal fossa with a mallet and gouge or drill. The opening is circumferentially enlarged with the Kerrison rongeur. When the frontoethmoid suture is reached, no further bone should be removed superiorly as this represents the level of the anterior cranial fossa. The lamina papyracea is taken down to allow complete exenteration of the ethmoid cells. Working intranasally, an uncinectomy and anterior ethmoidectomy will link the ethmoid cavity to the middle meatus. The middle turbinate should not be removed.

In some cases, the frontal sinus may be disrupted during the external ethmoidectomy procedure. To promote drainage of the frontal sinus, Silastic stents have sometimes been placed extending from the frontal sinus to the anterior aspect of the inferior turbinate and secured with a permanent suture. A period of at least 6 weeks is needed to produce a mucosally lined tract between the frontal sinus and the middle meatus. Even with stenting, stenosis tends to occur, and the orbital contents have a tendency to medialize into the frontal recess and interfere with sinus outflow.

At the end of the procedure, the nose is packed with 1 cm gauze or preformed surgical sponges impregnated with antibiotic ointment. The periosteum and subcutaneous tissues are closed with 3-0 absorbable sutures, and skin closure is accomplished with 6-0 nylon or mild chromic sutures.

Complications. A poorly designed incision can result in cosmetic problems such as a hypertrophic scar, medial canthal blunting, or medial canthal webbing. Webbing can be addressed with a Z-plasty at a later time or may be prevented by performing one at the time of the initial surgery. Injury to the lacrimal system can also occur, with epiphora as a consequence. If dissection proceeds into the orbit, serious complications such as cranial nerve injury or visual loss are possible. Retrobulbar hematoma and cerebrospinal fluid leak are other less common but potentially catastrophic complications of the external ethmoidectomy.

Frontal Sinus. Open approaches to the frontal sinus are indicated for chronic, complicated frontal sinusitis that has not responded to trephination or conventional endoscopic sinus surgery. Other indications include intracranial or orbital extension of disease, mucocele or mucopyocele, and osteomyelitis of the frontal bone.

Trephination. This approach to the frontal sinus is useful in the face of acute infection when mucosal bleeding may hamper an intranasal endoscopic approach. It is safe and rapidly performed and can decompress a pus-filled frontal sinus prior to a more definitive procedure. The procedure is also useful in conjunction with FESS to help locate the natural drainage pathway of the frontal sinus and to visualize the ostium from above and below.

Technique. An incision is made beneath the medial portion of the eyebrow, taking care to bevel the incision parallel to the hair follicles. The incision is carried through skin, subcutaneous tissue, and the periosteum over the frontal sinus floor. Once the

incision is completed, the wound can be retracted superiorly and medially over the anterior table of the frontal sinus. A drill is used to make a hole into the sinus in a controlled fashion (Figure 34–16). If the hole is made with a diameter greater than 4 mm, an endoscope can be inserted to visualize the interior of the sinus. Purulence is removed for culture, and the sinus is irrigated. The wound is closed with deep absorbable sutures and nylon or prolene on the skin. A catheter may be left in the trephination and brought out through the wound to allow for irrigation in the postoperative period.

Frontal Sinus Obliteration with Osteoplastic Flap. Frontal sinus obliteration is an old procedure that has been revived and refined for use in recalcitrant cases of frontal sinusitis. After endoscopic surgery, the frontal sinus ostium is sometimes very difficult to keep open from within the nose when osteoneogenesis is present. More extensive intranasal drill-out procedures may be successful in establishing long-term patency of a frontal sinus drainage pathway, but many times it is not possible to maintain an opening into the sinus with any endoscopic or open surgical technique. In these patients, frontal sinus obliteration may be the only viable option. The technique has the advantage of unparalleled visualization of the entire frontal sinus and elimination of the need to reconstruct the nasofrontal drainage outflow. The negatives of the technique include the external scar, potential for forehead hypesthesia, and the loss of a connection between the nose and sinus through which endoscopic surveillance can be performed.

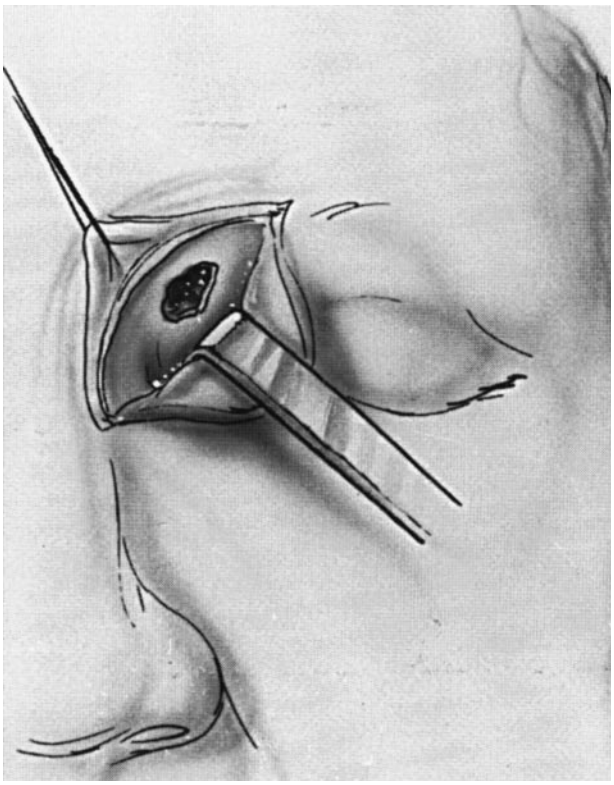


FIGURE 34–16. Frontal sinus trephination.

Technique The osteoplastic flap frontal sinus obliteration can be performed through either a gull-wing suprabrow incision or a bicoronal approach with the incision behind the hairline. Typically, the type of incision is determined by the patient's hairline and the presence of suitable forehead creases in which to place the scar. A 6-foot Caldwell radiograph of the sinuses is obtained preoperatively, and a template of the frontal sinus is fashioned and sterilized. The patient's face and hair are prepared for a bicoronal approach, and the abdomen is prepared for harvest of abdominal fat.

The bicoronal incision is made approximately 2 cm behind the hairline and curved laterally toward the ears on both sides (Figure 34–17). Since the scalp is very vascular, hemostasis must be maintained with Raney clips and careful cautery (avoiding hair follicles) to prevent excessive blood loss. The flap is elevated above the level of the pericranium down to the

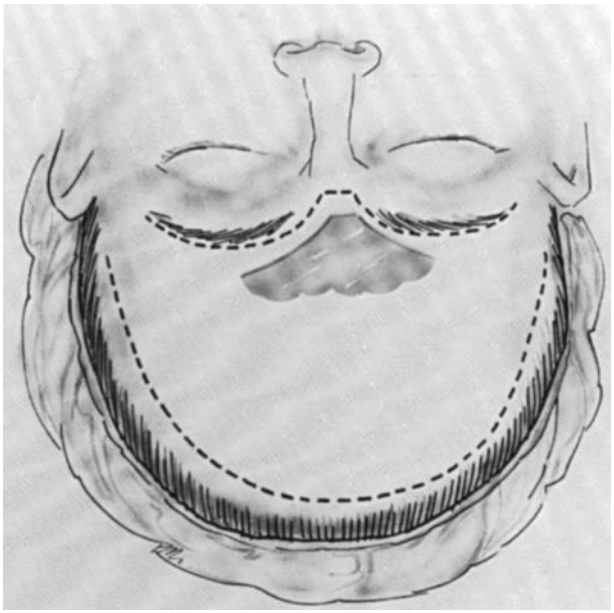


FIGURE 34–17. Incisions used for the osteoplastic approach to the frontal sinus.

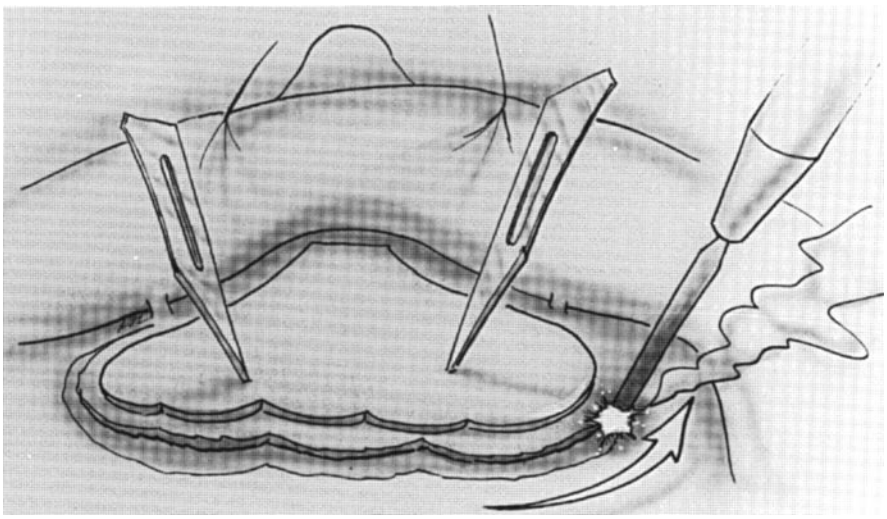


FIGURE 34–18. Periosteal incison being made around margins of the template. The *arrow* indicates the direction of the periosteal incision made with the cautery handpiece.

root of the nose and superior orbital rims. Near the orbits, the supraorbital neurovascular bundle should be identified and protected. With the flap elevated, the template is placed to show the borders of the frontal sinus, which are marked with a pen. Other techniques used to define the sinus boundaries are transillumination of the sinus through a trephination or the use of surgical navigation. Once the borders are marked, the periosteum is incised along the marked edges and elevated 1 to 2 mm (Figure 34–18). The periosteal flap remains attached inferiorly at the orbital rims to preserve a blood supply.

An oscillating saw is used to make the bone cuts through the anterior table of the frontal sinus. The cuts are beveled at about a 45-degree angle directed toward the sinus, both to prevent inadvertent penetration outside the sinus and to assist in repositioning the flap at the end of the procedure. Even with this degree of beveling, care must be taken to avoid cutting the posterior table of a very shallow sinus. With the bony cuts completed, the intersinus septum is cut and the flap is elevated with an osteotome. The bone flap, with the periosteum attached superficially, is reflected inferiorly to expose the sinus (Figure 34–19). To perform an obliteration successfully, it is absolutely critical to remove all of the mucosa from within the sinus. The gross mucosa is relatively easily stripped from the sinus and its recesses. However, rests of mucosa exist within the small foramina of Breschet in the bone of the anterior and posterior tables. Therefore, all of the bony surfaces of the frontal sinus must be meticulously burred down under microscopic visualization before placing the fat graft. The frontal ostium is occluded with pericranium or temporalis muscle to obliterate the connection to the nasal cavity. Abdominal fat is usually harvested through a periumbilical incision and cut to fit snugly within the confines of the sinus. When the graft is in place, the bone flap is replaced and fixed with small titanium plates to recreate the normal frontal contour. Drains are placed in the incision, and the wound is closed in layers.

Complications The most significant disadvantage of a frontal sinus obliteration procedure is the loss of the ability to image the sinus following the procedure. This can make assessment of postop-

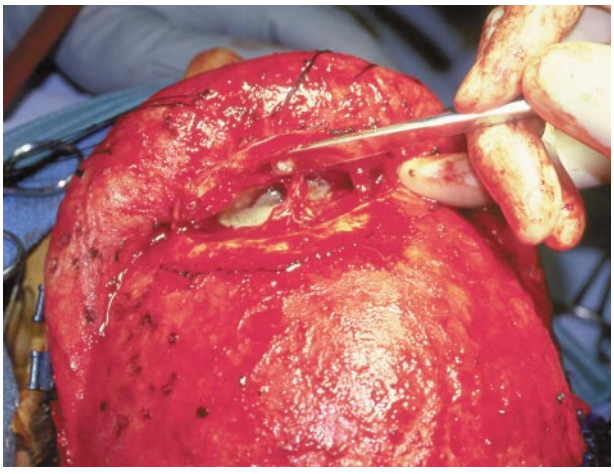


FIGURE 34–19. The osteoplastic flap approach to frontal sinus obliteration. Here the osteoplastic flap has been elevated, and the frontal sinus can be seen.

erative frontal sinus complaints difficult and delay detection of some late complications. Perioperative complications of the osteoplastic frontal sinus obliteration procedure include hematoma, infection or abscess of the bicoronal or abdominal wound, and dural injury from bone cuts outside the sinus. The major long-term complications of obliteration are pain or altered sensation, visible bony defects, and mucocele formation. Pain and sensory abnormalities often result from trauma or sectioning of the supraorbital nerve. Bone defects can occur from depression or protrusion of the osteoplastic flap or the development of hyperostosis. Mucoceles take many years to develop but can ultimately lead to extensive operations to secondarily remove the entrapped mucosa. The frontal recess is the most likely place for mucosa to have been incompletely removed, but there are always crevasses around the periphery of the frontal sinus that can harbor rests of mucosa. Incomplete separation of the nasal cavity from the obliterated sinus can eventually result in an infected graft and fat necrosis.

Sphenoid Sinus. Although the sphenoid sinus is readily accessible endoscopically through the nose, external approaches are frequently used to achieve wider exposure for resection of masses or for pituitary surgery. Some techniques, such as the external ethmoidectomy and transantral approaches, are seldom used today because of the oblique angle toward the sphenoid. Most commonly, the open procedures involve operating through the nasal septum to the face of the sphenoid. This can be accomplished via a transnasal septoplasty approach, an open rhinoplasty-type incision, or a sublabial incision. The transnasal approach has become more popular with the increased use of the endoscope; however, the technique requires a relatively large nose unless an alar incision is used. The external rhinoplasty incision gives unparalleled access and visualization but has the potential to leave a noticeable columellar scar. The sublabial approach is the one most often employed because it is easy, leaves no nasal scars, does not depend on nasal size, and allows the speculum to be placed in the midline. The drawbacks involve oral contamination of the wound and difficulties with oral incisions in denture wearers.

Sublabial, Transseptal Technique. After induction of general anesthesia, the upper gingivobuccal sulcus, nasal septum, and floor of the nose are injected with 1% lidocaine with 1:100,000 epinephrine. A scalpel is used to make an incision in the upper labial sulcus, leaving a cuff of mucosa attached to the gingiva to facilitate closure. The incision is carried down to the maxillary bone, and the periosteum is elevated to the inferior margin of the pyriform aperture until the anterior nasal spine is exposed. The right septal periosteum is incised at the caudal margin, and mucoperichondrial flaps are elevated along the right side of the septum and along both floors of the nasal cavities. The septum is divided at the bony-cartilaginous junction, and mucoperiosteal flaps are elevated bilaterally from the bony septum. The pockets along the nasal floors are then connected with the septal pockets, and the cartilaginous septum is dislocated from the maxillary crest toward the left. The perpendicular plate of the ethmoid is then removed with forceps, taking care not to place any torque on the superior attachment of the bone to the ethmoid roof. Once the rostrum of the sphenoid is reached, fluoroscopy or surgical navigation may be used to verify proper positioning. The anterior face of the sphenoid can then be entered directly in the midline or by first locating the natural ostia and using these as the superior and lateral landmarks. For pituitary surgery, the speculum is then inserted, and the procedure is generally turned over to neurosurgery. After pituitary tumor resection, the sphenoid mucosa is often stripped away, and the sinus is obliterated with fat or bone grafts. The septal flaps are carefully reapproximated, and the septum is returned to the midline in the maxillary crest. The sublabial incision is closed with absorbable suture, and the nose is packed.

External Rhinoplasty Technique. The sublabial approach was first described by Cushing and was then reintroduced and popularized by Hardy in 1968.[37] Although it has become the standard, drawbacks to this approach include limited exposure secondary to the overhanging upper lip, loss of nasal tip projection from resection of the anterior nasal spine, postoperative dental hypesthesia, and difficulty with denture fitting in older patients. The external rhinoplasty approach was designed to avoid these limitations of the traditional sublabial technique.[38] In this procedure, the nasal columella, dome, septum, and nasal floors are injected with 1% lidocaine with 1:100,000 epinephrine. A columellar

flap is created with an external rhinoplasty incision and elevated onto the dome of the nose. The medial crural ligaments are divided to expose the caudal edge of the septal cartilage. A mucoperichondrial flap is then raised on the left side of the septum beyond the bony-cartilaginous junction and onto the maxillary crest and floor of the nose. At this point, the cartilage is separated posteriorly and inferiorly so that posterior and inferior tunnels can be elevated on the right side and connected. Once the speculum is inserted, the perpendicular plate of the ethmoid is resected and the vomer followed to the anterior face of the sphenoid sinus. The sphenoid is entered, removing the mucosa, intersinus septum, and posterior wall to reach the pituitary. After the hypophysectomy is performed by neurosurgery, the sphenoid is obliterated, and the septal flaps are replaced. The medial crura is repaired with 4-0 chromic suture. The columellar incision is closed with 5-0 mild chromic, and the nose is packed.

Complications The nasal complications are typically related to the septoplasty portion of the procedure. These include septal perforation, saddle deformity, and tip deformity. Epistaxis and wound infection are also possible nasal problems postoperatively. Potentially serious neurologic and vascular complications may occur in sphenoid surgery since the carotid artery and optic nerves travel in the lateral wall of the sinus. Even if the optic nerve is not injured directly during surgery, overpacking of the sinus with fat can cause optic chiasmal compression and visual loss. Another possible complication is cerebrospinal fluid leak, which should be treated when it is recognized intraoperatively. During the postoperative period, patients must be closely monitored for evidence of change in mental status or signs of active bleeding. The cause of these findings will generally be discovered through radiologic evaluation, and the proper intervention can be planned and undertaken by the operative team.

FUTURE DIRECTIONS

There have been many recent advances in the medical and surgical approaches to sinus disease. Computerized guidance systems continue to improve and have the potential to assist the surgeon greatly in operating safely at the skull base for difficult sinus disease. An improved understanding of the necessity for mucosal preservation in endoscopic sinus surgery has reduced the incidence of postoperative scarring and sinus dysfunction requiring revision surgery. With the current armamentarium available for early diagnosis and aggressive treatment, most sinus disease can be controlled and serious complications averted.

However, considering the prevalence of rhinosinusitis, its economic cost, and quality of life impact on those who suffer with it, there is much about the pathophysiology of the disease that has yet to be learned. In particular, chronic rhinosinusitis that is recalcitrant to current methods of medical and surgical management awaits a new therapeutic approach. As with many areas in medicine, the future of diagnosis and treatment in sinus disease will likely use the rapidly expanding tools of molecular biology and the unfolding knowledge of the human genome. Already there has been the discovery that some patients with recurrent sinusitis have alterations in the membrane ion channel gene that is responsible for cystic fibrosis.[39] Perhaps there are other subtle, undiscovered differences at the genomic level that predispose patients to chronic sinonasal inflammation. As the underlying mechanisms of chronic sinus disease are uncovered, more directed pharmacologic therapy may be devised that spares the patient the side effects and complications of systemic corticosteroids. Ultimately, specific genetic markers associated with an increased risk of sinusitis may be identified. With this information, early genetic screening will allow preventive treatment or even gene therapy to be instituted prior to the onset of symptoms. This has already begun with other types of diseases, and it is only a matter of time before the methodology is applied to rhinosinusitis.

REFERENCES

1. National disease and therapeutic index. Plymouth Meeting (PA): IMS Inc., 1994. p. 963–76.
2. Benson V, Marano MA. Current estimates from the National Health Interview Survey, 1992. Vital Health Stat 1994;189:12.
3. Ray NF, Baruniak JN, Thamer M, et al. Health care expenditures for sinusitis in 1996: contributions of asthma, rhinitis, and other airway disorders. J Allergy Clin Immunol 1999;103:408–14.

4. Brook I, Gooch WM, Jenkins SG, et al. Medical management of acute bacterial sinusitis. Recommendation of a clinical advisory committee on pediatric and adult sinusitis. Ann Otol Rhinol Laryngol 2000;109(5 Pt 3, Suppl):2–20.
5. McCaig LF, Hughes JM. Trends in antimicrobial drug prescribing among office-based physicians in the United States. JAMA 1995;273:214–9.
6. Ober NS. Respiratory tract infections: consider the total cost of care. Drug Benefit Trends 1998; 10:23–9.
7. Khan R, Dolata J, Lindberg S. Effects of inflammatory mediators on ciliary function in vitro. Rhinology 1995;33:22–5.
8. Borman KR, Brown PM, Mezera KK, Jhaveri H. Occult fever in surgical intensive care unit patients is seldom caused by sinusitis. Am J Surg 1992; 164:412–5.
9. Vandenbussche T, De Moor S, Bachert C, Van Cauwenberge P. Value of antral puncture in the intensive care patient with fever of unknown origin. Laryngoscope 2000;110(10 Pt 1):1702–6.
10. Lanza DC, Kennedy DW. Adult rhinosinusitis defined. Otolaryngol Head Neck Surg 1997;117(3 Pt 2):S1–7.
11. Bachert C, Wagenmann M, Rudack C, et al. The role of cytokines in infectious sinusitis and nasal polyposis. Allergy 1998;53:2–13.
12. Wardlaw AJ. Molecular basis for selective eosinophil trafficking in asthma: a multistep paradigm. J Allergy Clin Immunol 1999;104:917–26.
13. Stringer SP, Ryan MW. Chronic invasive fungal rhinosinusitis. Otolaryngol Clin North Am 2000; 33:375–87.
14. deShazo RD, O'Brien M, Chapin K, et al. A new classification and diagnostic criteria for invasive fungal sinusitis. Arch Otolaryngol Head Neck Surg 1997; 123:1181–8.
15. Bent JP, Kuhn FA. Diagnosis of allergic fungal sinusitis. Otolaryngol Head Neck Surg 1994;111:580–8.
16. Ferguson BJ. Definitions of fungal rhinosinusitis. Otolaryngol Clin North Am 2000;33:227–35.
17. Ponikau JU, Sherris DA, Kern EB, et al. The diagnosis and incidence of allergic fungal sinusitis. Mayo Clin Proc 1999;74:877–84.
18. Kirsch JP, White JA. Nasal polyposis. La State Med J 1996;142:11–4.
19. Drake-Lee AB, Lowe D, Swanston A, Grace A. Clinical profile and recurrence of nasal polyps. J Laryngol Otol 1984;98:783–93.
20. Shin SH, Park JY, Jeon CH, et al. Quantitative analysis of eotaxin and RANTES messenger RNA in nasal polyps: association of tissue and nasal eosinophils. Laryngoscope 2000;110:1353–7.
21. Chandler JR, Langenbrunner DJ, Stevens ER. The pathogenesis of orbital complications in acute sinusitis. Laryngoscope 1970;80:1414–28.
22. Hoyt DJ, Fisher SR. Otolaryngologic management of patients with subdural empyema. Laryngoscope 1991;101:20–4.
23. Kennedy DW. A 48-year-old man with recurrent sinusitis. JAMA 2000;283:2143–50.
24. Havas TE, Motbey JA, Gullane PJ. Prevalence of incidental abnormalities on computed tomography scans of the paranasal sinuses. Arch Otolaryngol Head Neck Surg 1988;114:856–9.
25. Lund VJ, Kennedy DW. Staging for rhinosinusitis. Otolaryngol Head Neck Surg 1997;117(3 Pt 2):S35–40.
26. Kaliner MA, Osguthorpe JD, Fireman P, et al. Sinusitis: bench to bedside. Current findings, future directions. J Allergy Clin Immunol 1997;99 Suppl: S829–48.
27. Hueston WJ, Eberlein C, Johnson D, Mainous AG. Criteria used by clinicians to differentiate sinusitis from viral upper respiratory infection. J Fam Pract 1998;46:487–92.
28. Aronovitz GH. Antimicrobial therapy of acute otitis media: review of treatment recommendations. Clin Ther 2000;22:29–39.
29. Wawrose SF, Tami TA, Amoils CP. The role of guaifenesin in the treatment of sinonasal disease in patients infected with the human immunodeficiency virus (HIV). Laryngoscope 1992;102:1225–8.
30. Naumann H. Pathologische Anatomie der chronischen Rhinitis and Sinusitis. In: Proceedings of the VIIIth International Congress of Otorhinolaryngology. Amsterdam: Excerpta Medica; 1965. p. 80.
31. Messerklinger W. Endoscopy of the nose. Baltimore: Urban & Schwarzenberg; 1978.
32. Kennedy DW, Goodstein ML, Miller NR, Zinreich SJ. Endoscopic transnasal orbital decompression. Arch Otolaryngol Head Neck Surg 1990;116: 275–82.
33. Kennedy DW, Josephson JS, Zinreich SJ, et al. Endoscopic sinus surgery for mucoceles: a viable alternative. Laryngoscope 1989;99:885–95.
34. Stammberger H. Functional endoscopic sinus surgery. Philadelphia: BC Decker; 1991.

35. Setliff RC. The hummer: a remedy for apprehension in functional endoscopic sinus surgery. Otolaryngol Clin North Am 1996;29:95–104.
36. McDonald TJ, Pearson BW. Follow-up on maxillary artery ligation for epistaxis. Arch Otolaryngol 1980; 106:635–8.
37. Hardy J. Transsphenoidal hypophysectomy. J Neurosurg 1971;34:582–94.
38. Koltai PJ, Goufmam DB, Barnes SM, Steiniger JR. Transsphenoidal hypophysectomy through the external rhinoplasty approach. Otolaryngol Head Neck Surg 1994;111:197–200.
39. Wang X, Moylan B, Leopold DA, et al. Mutation in the gene responsible for cystic fibrosis and predisposition to chronic rhinosinusitis in the general population. JAMA 2000;284:1814–9.

CHAPTER 35

Headache and Facial Pain

James M. Hartman, MD, Richard A. Chole, MD, PhD

The craniofacial region is the most common location in which pain drives patients to seek medical attention.[1] The wide diversity of causes and the extreme overlap of historical features coupled with nonspecific physical findings serve to challenge the physician's ability to diagnose and ultimately relieve or control patients' craniofacial pain. However, an understanding of the different types of craniofacial pain provides a powerful tool to meet the challenge.

INNERVATION FOR PAIN

Nociceptors serve as the sense organs in which noxious stimuli create a response that excites afferent nerve fibers that provide the brain with information about location, intensity, quality, and duration of the response.[2] Neurochemicals responsible for this excitation include serotonin and substance P as well as other neurotransmitters. These afferent fibers are carried to the central nervous system in cranial nerves V, IX, X, and XI and the first three cervical nerves.

Pain-sensitive innervation of facial structures is extensive, whereas intracranial pain sensation is limited to specific structures. The extracranial tissues innervated for pain sensation include the muscles of the head and neck, the scalp and facial skin, sinonasal mucosa and perichondrium, temporomandibular joint (TMJ) synovium and capsule, tooth pulp, the external and middle ear, orbital contents, salivary glands, cervical spine, and craniofacial periosteum. Intracranial structures with nociceptive neurons include major arteries, specifically the internal carotids, vertebral basilars, middle meningeals, ophthalmics, and the circle of Willis, and the major venous sinuses. Nonvascular structures with pain nerve fibers are represented by cranial nerves V, VI, VII, IX, and X and the upper three cervical nerves themselves, as well as the pituitary fossa and the base of skull dura mater. The remaining intracranial structures are asensate to pain, including the brain, most of the dura, the ventricles, and the cranium. The afferents converge on nuclei in the brainstem where multiple synaptic connections occur including transmission to the ipsi- and contralateral thalamus and the somatosensory cerebral cortex. The thalamic nuclei appear to play a role in the affective response to pain, whereas the cortical centers' role is in localization and intensity recognition of noxious stimuli.[2] The extensive convergence of afferent neurons in brainstem nuclei, however, limits the brain's ability to distinguish between sources. The result is referral of pain to tissues with a past experience recognized as pain.[3] Specific patterns of referral of pain are common; for example, from the TMJ and muscles of mastication, radiation is to the ear, cheek, and temple[4]; from the tonsillar fossa and supraglottic larynx to the middle ear; and from the maxillary sinus to the maxillary teeth, whereas pain from the sphenoid sinus is more often referred to the vertex or occiput, and, of course, angina is sometimes referred to the jaw.

PATHOPHYSIOLOGY

Multiple mechanisms resulting in excitation of nociceptive neurons (ie, generating the perception of pain) are partially understood. One common mechanism is sustained muscle contraction resulting in tension headache. The exact source of excitation is unclear but may be the result of ischemic changes or the production of nitric oxide.[5] Another common scenario is vasodilation of intracranial arteries stimulating trigeminal sensory pathways, which release vasoactive peptides that increase the pain response.[6] The vasodilation seems linked to subtypes of serotonin 5-hydroxytryptamine (5-HT) receptors in vessel walls. Other subtypes of 5-HT receptors, such as 5-HT_{1B}, yield vasoconstriction and inhibit the pain response.[7] It is this basis on

which triptan antimigraine agents have their effect by selectively binding to 5-HT_{1B} receptors. An inflammatory mechanism is thought to be responsible when neuropeptides such as substance P are released with mucosal inflammation. These result in neural excitation, resulting in the perception of pain.[8] Neural inflammation following injury or tumor invasion has an excitatory effect. Direct nerve pressure may induce nociceptor activity, as seen in foraminal stenosis. Many agents that result in vasodilation can trigger headache including hypoxia, carbon monoxide, caffeine withdrawal, acute alcohol withdrawal, oral contraceptives, hypoglycemia, and antihypertensive and other vasodilators such as nitroglycerin and monosodium glutamate found in Chinese food.[9] Finally, cerebral mechanisms are also thought to play a role. For example, some migraineurs have been found to have defective release of endogenous opiates,[10] and lowered cortical pain thresholds occur in chronic tension headaches.[11]

CLASSIFICATION

Understanding headache and facial pain is essential to facilitate diagnosis and treatment. To this end, definitions and features of clinical syndromes were organized by the International Headache Society (IHS).[12] This classification, with inclusion of diagnostic criteria for headaches, cranial neuralgias, and facial pain, was created in 1988 and has facilitated the diagnostic approach and management of craniofacial pain across many medical fields. Table 35–1 lists the classification of most headaches, neuralgias,

TABLE 35–1. International Headache Society Classification of Headache and Facial Pain

- Migraine type
  - Without aura (common migraine)
  - With aura (classic migraine)
  - With prolonged aura (complicated migraine)
  - Ophthalmoplegic
  - Retinal
- Tension type
  - Episodic (muscle contraction headache)
  - Chronic (chronic daily headache)
  - Oromandibular dysfunction (myofascial pain dysfunction syndrome) (temporomandibular joint pain dysfunction syndrome)
- Cluster (Horton's cephalalgia)
- Post-traumatic headache
- Vascular intracranial disorder
  - Transient ischemic attack–associated headache
  - Intracranial hematoma
  - Subarachnoid hemorrhage
  - Unruptured aneurysm
  - Giant cell arteritis (temporal arteritis)
  - Carotid or vertebral artery pain
    - Dissection
    - Carotidynia
  - Cerebral venous thrombosis
  - Acute arterial hypertension
    - Pheochromocytoma
    - Malignant hypertension (accelerated)
    - Preeclampsia and eclampsia
- Nonvascular intracranial disorder
  - Benign intracranial hypertension (pseudotumor cerebri)
  - Post–lumbar puncture headache
  - Cerebrospinal fluid fistula headache
  - Intracranial infection
    - Meningitis
    - Encephalitis
    - Brain abscess
    - Subdural empyema
  - Intracranial neoplasm
- Headache from substance exposure or withdrawal
  - Acute exposure
    - Nitrate- or nitrite-induced headache (hotdog headache)
    - Monosodium glutamate–induced headache (Chinese restaurant syndrome)
    - Carbon monoxide–induced headache
    - Alcohol-induced headache
  - Chronic exposure
    - Ergotamine-induced headache
    - Analgesics abuse headache
    - Oral contraceptives use
  - Acute withdrawal
    - Alcohol (hangover)

Continued

TABLE 35–1. Continued

- Chronic withdrawal
  - Analgesics
  - Ergotamine
  - Caffeine
  - Narcotics
- Headache with extracranial infection
  - Viral
  - Bacterial
- Headache from metabolic disorder
  - Hypoxia
    - High-altitude headache
    - Sleep apnea headache
  - Hypoglycemia
- Headache or facial pain associated with craniofacial disorder
  - Cranial disorder
    - Osteomyelitis
    - Multiple myeloma
    - Paget's disease
  - Cervical spine disorder (cervicogenic headache)
  - Eye disorder
    - Acute glaucoma
    - Refractive errors
  - Sinonasal disorder
    - Acute sinus headache
    - Rhinogenic headache
  - Odontomandibular disorder
    - Periodontitis
    - Pulpitis
  - Glossitis (burning mouth syndrome)
  - Temporomandibular joint disease
- Cranial neuralgia
  - Compression of cranial nerve or cervical root 1, 2, or 3
  - Inflammation of cranial nerves
    - Acute herpes zoster
    - Chronic postherpetic neuralgia
    - Tolosa-Hunt syndrome
    - Gradenigo's syndrome
    - Raeder's paratrigeminal neuralgia
  - Trigeminal neuralgia (tic douloureux)
    - Idiopathic
    - Compression of trigeminal ganglion
      - Vascular
      - Tumor
      - Cholesteatoma
      - Aneurysm
      - Multiple sclerosis
  - Glossopharyngeal neuralgia
  - Occipital neuralgia
  - Anesthesia dolorosa
    - Postsurgical after trigeminal rhizotomy
- Unclassifiable facial pain (atypical facial pain)

Formerly used nomenclature is in parenthesis.

Adapted from Headache Classification Committee of the International Headache Society[12] with permission from Blackwell Science Ltd.

and facial pains. Greater detail can be found in the IHS classification publication.

PATIENT EVALUATION

History is the essential tool to establish a diagnosis for headache or facial pain. It must be comprehensive to ensure accuracy of diagnosis. Once the features of the cephalalgia are discovered, one can refer to the IHS classification to determine which specific diagnostic criteria are fulfilled (Table 35–2).[13–15]

Serious or even life-threatening conditions associated with headache usually will present with distinct characteristics of the headache and nervous system or systemic complaints (see Table 35–3 for specific serious conditions in which headache is a presenting symptom[16]). Warning features of the headache include sudden onset and highest severity, steady crescendo of intensity, awakening with headache, and exacerbation by coughing or straining. Dangerous associated systemic symptoms include fever, sudden vomiting, declining mental status, syncope, or seizures. Specific physical ailments such as stiff neck, tender or enlarged temporal artery, unilateral rhinorrhea, skull percussion tenderness, or focal neurologic defects such as visual field loss, diplopia, hemiparesis, and facial nerve palsy should be alarming. Finally, a history of preex-

TABLE 35–2. Pertinent Elements of a Headache History[13–15]

- Temporal profile
  - Course
    - Duration of problem
    - Frequency of headaches
  - Duration of each headache
- Location
  - Unilateral or bilateral
  - Focal or diffuse
  - Radiates to other sites
- Quality
  - Intensity
  - Characteristics
    - Throbbing (vascular)
    - Stabbing (neuritic)
    - Pressure (viscus or chamber derived)
    - Aching or burning (muscular)
- Prodrome presence (hours to days prior to headache)
  - Mood changes
  - Fatigue
  - Nausea
  - Excessive yawning
  - Chilling
  - Anorexia or food craving
  - Stiff neck
  - Urinary frequency
  - Diarrhea
  - Fluid retention
- Aura presence
  - Visual disturbances
    - Scotomata
    - Photophobia
  - Motor disturbances
    - Ataxia
    - Hemiparesis
    - Aphasia or dysarthria
    - Diplopia
  - Sensory disturbances
    - Hypo- or hypersensitivity to touch
    - Paresthesias
    - Phonophobia
    - Tinnitus
    - Vertigo
    - Confusion
- Associated symptoms
  - Nausea and vomiting
  - Lacrimation
  - Nasal congestion
  - Rhinorrhea
  - Tender or pulsating scalp
  - Vertigo
  - Syncope
  - Confusion
  - Sweating
- Precipitating factors
  - Smell provocation
  - Chewing
  - Head movement or sustained position
  - Coughing
  - Yawning
  - Straining
  - Position change
  - Stress
  - Sleep change
  - Missed meals
  - Menses
  - Weather change
  - Alcohol intake
  - Specific food ingestion
  - Medication use
- Past medical and dental history
  - Trauma to head or neck
  - Surgery to head or neck
  - Dental procedures
  - Meningitis
  - Hypertension
  - Diabetes
  - Glaucoma
  - Rheumatoid arthritis
  - Sinusitis
- Medication use
  - Over-the-counter analgesics
  - Prescription analgesics
    - Vasodilators
    - Oral contraceptives
    - Insulin
  - Herbal remedies

Continued

TABLE 35–2. Continued

- Family history of headache
  - Migraine
  - Tension type
  - Temporomandibular joint disorders
  - Subarachnoid hemorrhage
  - Substance abuse
- Social history
  - Smoking habits
  - Alcohol intake
  - Illicit drug use
  - Occupation
- Review of systems
  - Fever
  - Weight change
  - Sleep habit change
  - Appetite change
  - Depression or anxiety
  - Decreased range of motion
    - Jaw
    - Neck
  - Bruxism
  - Chest pain
- Motivation (secondary gain)
  - Disability
  - School phobia

isting malignancy or immunosuppression may herald a serious condition.[17]

The physical examination is guided by history and clinical suspicions. Routine measures should include a check of blood pressure and temperature; assessment of mental status and cranial nerve function; and a visual examination inclusive of pupil size, extraocular movement, visual fields, and funduscopy. Range of motion testing of the jaw and neck should be performed along with palpation of the TMJ and muscles of mastication and paraspinal musculature searching for trigger points for neuritic or muscular pain. Abnormalities may be observed in the external or middle ear, the sinonasal passages, the mouth, and dentition, or herpetic lesions may be seen in dermatomes. Palpation of the temporal arteries should also be performed. Asymmetric skull percussion tenderness may suggest a subdural process. Laboratory testing, including cerebrospinal fluid (CSF), should be limited and guided by clinical impression. Radiologic imaging with computed tomography (CT), magnetic resonance imaging (MRI), or magnetic resonance angiography (MRA) may be indicated when organic pathology is suspected or to reassure the patient.

SPECIFIC CLINICAL SYNDROMES

The characteristic features of many craniofacial pain diagnoses are presented along with specific treatment recommendations (see Table 35–1 for specific conditions in the classification system).

Migraine-Type Headaches

Migraines comprise the second most common form of headache, with a prevalence of 18% in women and 6% in men and costing billions of dollars for treatment and lost productivity.[18] Its peak age of onset is in the second or third decade, but children and even infants may also be affected.[19] Migraine headaches usually occur as recurrent episodes of severe, throbbing, unilateral head pain of sudden onset lasting 4 to 72 hours. However, 40% of migraineurs have bilateral pain. The headache often strikes after awakening in the morning. In children, migraine headaches usually end between 2 and 48 hours. Routine activities of daily living exacerbate the symptoms. Coexisting symptoms with the headache are common including nausea, vomiting, photophobia, and phonophobia. Stress is the usual precipitating factor. A family history of migraine is often present.

Migraine is divided into several types: those without aura (formerly *common migraine*) and those with aura (formerly *classic migraine*) are the two most common by far. Migraine without aura is the most common type. It lacks any preheadache warning symptoms. Migraine with aura is also common, comprising 20% of migraine sufferers. The typical

TABLE 35–3. Life- or Organ-Threatening Causes of Headaches

Subarachnoid hemorrhage
Intracranial aneurysm
Meningitis
Encephalitis
Major artery dissection
Giant cell arteritis (temporal arteritis)
Acute glaucoma
Hypertensive encephalopathy
Carbon monoxide poisoning
Benign intracranial hypertension (pseudotumor cerebri)
Cerebral venous thrombosis
Preeclampsia and eclampsia
Cerebral vascular accident
Mass lesion
Neoplasm
Abscess
Intracranial hematoma
Cerebrospinal fluid fistula

Adapted from Edlow JA, Caplan LR. Avoiding pitfalls in the diagnosis of subarachnoid hemorrhage. N Engl J Med 2000;342:29–36.

aura lasts less than 1 hour and immediately precedes the onset of cephalalgia. It may be precipitated by menses, pregnancy, oral contraceptives, certain foods, or bright lights. *Migraine with prolonged aura* has symptoms of the aura that last through and beyond the headache up to 7 days. Features of aura include focal neurologic symptoms including visual changes such as scotomata or sensory and motor disturbances such as paresthesia, hemiparesis, or asphasia. Brainstem dysfunction may also occur, resulting in ataxia, lethargy, diplopia, tinnitus, vertigo, or dysarthria. *Ophthalmologic and retinal migraine* are rare but result in paresis of ipsi- or contralateral oculomotor nerves or monocular blindness secondary to retinal ischemia, respectively.[20]

Prodromes and postdromes are also common features, occurring in approximately 60% of migraine patients.[13] They experience premonition phenomena hours to days before the onset of head pain. Specific feelings include mood changes, fatigue, anorexia, or autonomic symptoms such as diarrhea or frequent urination. The postdrome essentially consists of general fatigue, irritability, and food cravings or anorexia. The IHS classification requires that the following criteria be met to diagnose migraine: For migraine without aura, at least five attacks lasting 4 to 72 hours, with at least two of the following features present: unilaterally, pulsatile quality, moderate to severe intensity, and exacerbation by physical activity. The headache must be associated with nausea, vomiting, photophobia, or phonophobia. For migraine with aura, at least two attacks with at least three of the following must occur: one or more reversible aural symptoms occurs, the symptom has a gradual onset over more than 4 minutes, the aura resolves in less than 1 hour, and the time between aura and headache is less than 1 hour.[12] Some migraine sufferers will complain of aural symptoms and not experience headache.

Treatment of migraine disorders includes initiation of supportive and educational measures such as a headache diary to decrease the frequency of attacks. Medical therapy falls into two major categories: symptomatic and preventive treatment. The former attempts to abort attacks at their onset or control pain during the headache. Specific measures prove useful including the application of ice, isolation in a dark quiet room, using biofeedback or acupuncture, and inducing sleep, as well as avoiding known triggers (Table 35–4).[13,21]

Symptomatic treatment must be individualized to provide adequate relief with minimal risk and side effects. Stratified care for which a variety of treatment choices are available to the patient has proven successful.[22] Relief should not be limited to headache but also constitutional symptoms such as nausea and vomiting or insomnia. Factors influencing choice of drugs should include the frequency that medications will be required, contraindications, route of administration, prior medication successes and failures, or the need for breakthrough headache therapy.[13] The earlier therapy is initiated, the more likely a good or complete response will occur. Nonspecific migraine therapy includes analgesics and judicious use of opiates for only severe attacks. Specific therapy uses compounds that are known agonists of the 5-HT_1 receptor and include ergot derivatives and synthetic triptan compounds. Often antiemetics or metoclopramide should be given first

TABLE 35–4. Triggers for Migraine

Ingested food or drink
Aged cheese
Alcohol
Artificial sweeteners
Caffeine
Chocolate
Fermented foods
Nitrites
Red wine
Odor exposure
Cigarette smoke
Cleaning agents
Exhaust
Paints
Perfumes
Hormonal variations
Hormonal replacement therapy
Menses
Oral contraceptives
Pregnancy
Irregular schedule
Exercise
Meals
Sleep
Stress or anger
Flashing lights

Reproduced from Saper JR[13] by permission of Blackwell Science, Inc.

TABLE 35–5. Symptomatic Pharmacologic Treatment Options for Migraine[13,21,23–27]

Mild headaches
Acetaminophen, oral or rectal
NSAIDs: naproxen or ibuprofen, oral butalbital and acetaminophen with or without caffeine, oral
Acetaminophen and isometheptene and dichloralphenazone (Midrin), oral
Moderate headache
Indomethacin, oral or rectal
Ergotamine tartrate, oral, sublingual, or rectal
Dihydroergotamine mesylate (Migranol), intranasal spray or subcutaneous
Sumatriptan (Imitrex or Imigran), subcutaneous autoinjector, oral, or intranasal spray
Naratriptan (Amerge), oral
Rizatriptan (Maxalt), oral
Zolmitriptan (Zomig), oral
Lidocaine, intranasal spray
Butorphanol, intranasal spray
Butalbital and acetaminophen and codeine with or without caffeine, oral
Severe headache
Dihydroergotamine mesylate, as for moderate, also intramuscular or intravenous
Ketorolac, intramuscular
Diphenhydramine, intravenous
Prochlorperazine, intravenous
Chlorpromazine, intravenous
Haloperidol, intravenous
Opiates: morphine sulfate, intramuscular or intravenous

NSAID = nonsteroidal anti-inflammatory drug. Adapted from Saper JR,[13] Wilkinson M,[21] Silberstein SD and Young WB,[23] Young WB,[24] Logemann CD and Rankin LM,[25] Maizels M,[26] and Finkel AG et al.[27]

to avoid worsening or triggering nausea. To avoid chronic ergotamine-induced headache, it should not be taken more than 2 days a week.[23] In addition to nausea, angina may be induced by these compounds; thus, their use in patients with risk of coronary artery disease is contraindicated (see Table 35–5 for agents available for abortive therapy[13,21,23–27]).

The recurrence rate for headache in a 24-hour period is reported for some drugs: intranasal dihydroergotamine mesylate has a 14% relapse, whereas triptans relapse between 22 and 45%,[25] and intranasal lidocaine relapses 40% of the time.[26] It is important to realize that severe attacks can be so disabling that occasional referral to the emergency department or admission for pain management is necessary. Finally, in difficult to treat cases, referral to a headache center or neurologist is appropriate. For children, Annequin and associates reviewed the use of various agents and found good effectiveness for both acetaminophen and ibuprofen. Oral dihydroergotamine mesylate and sumatriptan in any form were also effective.[19] Opiates should be avoided. Again, abortive therapy should be initiated early in the attack.

Prophylactic therapy is indicated when the frequency of episodes is greater than two per week.

Dosages should be titrated within recommended levels, and therapeutic trials of any agent should last at least 3 weeks. Specific prophylactic agents are listed in Table 35–6.[7,19,27–29] Preventive medical therapy in children has had some success, although fewer reports exist. Among specific agents used for children, propranolol is most frequently used. Other regimens make use of metoprolol, valproate, flunarizine, cyproheptadine, and pizotifen, a serotonin receptor antagonist.[19]

Tension-Type Headache

Tension-type headaches include a variety of previously described craniofacial pain conditions such as tension headache, stress headache, muscle contraction headache, fibromyalgia, chronic daily headache, myofascial pain syndrome, and TMJ dysfunction syndrome. This category of headaches is by far the most common, afflicting approximately 80% of the adult population, but these headaches usually do not result in a physician evaluation. The IHS delineated both episodic and chronic varieties of tension-type headaches and noted whether there is an associated pericranial muscle disorder, more specifically, muscular tenderness or increased electromyographic activity. It is more likely to occur in women, and usually a family history of headaches is present. The headaches are characteristically bilateral, with a tightening or band-like sensation in the frontotemporal region around the head and spreading to the occipital region or trapezius muscles. The onset is gradual, whereas the quality is dull, nonthrobbing, and constant, sometimes lasting for weeks. These headaches typically do not have associated autonomic symptoms as migraine headaches do, although it is possible when the headaches are severe. The cephalalgia is triggered or exacerbated by stress or anxiety in most patients. Some evidence has discounted the idea that persistent muscle contraction is the cause of the pain. Instead, new theories suggest an abnormality in central pain control.[30] The subdivision of *episodic tension-type headaches*, previously referred to as muscle contraction headaches, requires certain diagnostic criteria, specifically at least 10 prior episodes of headaches less than 15 days a month, lasting 1 half hour to 1 week and having at least two of the following characteristics: a tightening quality, mild-to-moderate intensity, bilaterality, and no exacerbation by routine exertion.[30]

The other category of tension-type headache is chronic and was previously labeled chronic daily headache. Some debate still exists as to whether all chronic daily headaches are tension type. The IHS criteria for *chronic tension-type headaches* are the same as the episodic variety, with the exception that headache frequency is greater than or equal to 15 days per month for at least 6 months. The IHS also recognizes that migraine headaches can transform into a chronic daily headache with features more similar to tension-type headaches. This may represent an evolution of the headache process or often is related to overuse of treating drugs. According to Silberstein and Lipton, chronic daily headache must be divided into both tension type and migrainous type, as well as other entities.[31] They note that the chronic tension type constitutes about 25% of all daily headache patients.

Treatment of tension-type headaches uses nonpharmacologic and medical therapy. First, chronic medication overuse must be eliminated, a step that will improve many cephalalgia sufferers. Depression should be treated if present. Nonmedical remedies include reassurance, muscle relaxation, simple muscle exercises, stress management, biofeedback, and physical therapy with thermal modulation, ultrasonography, or electrical stimulation. Pharmacotherapy varies between episodic and chronic varieties. Most patients who suffer episodic tension-type headache will respond to analgesics, such as aspirin or nonsteroidal anti-inflammatory drugs (NSAIDs). Antidepressants, especially amitriptyline, have also

TABLE 35–6. Prophylactic Medications for Migraine[7,19,27–29]

Antidepressants: amitriptyline, nortriptyline, doxepin, or trazodone
Ergotamine in a sustained-release form (Bellergal)
Beta blockers: metoprolol, atenolol, or propranolol
Nonsteroidal anti-inflammatory drugs (NSAIDs): naproxen, ibuprofen, or aspirin
Calcium channel blockers: verapamil, nifedipine, nimodipine, or flunarizine
Anticonvulsants: divalproex, valproic acid, gabapentin, or topiramate
Others: methysergide, fluoxetine, cyproheptadine, pizotifen, riboflavin, or lithium

proven effective. Agents capable of fostering dependence are best avoided. These include caffeine, opiates, and benzodiazepines.

The use of medications in the management of chronic tension-type headaches is more diverse. Injection of local anesthetics into trigger points may provide sustained relief.[32] Amitriptyline has been used most often, with proven success.[31] Sodium valproate, topiramate, L-5-hydroxytryptophan, and L-arginine have also been used with reported success. Injection of botulinum toxin (Botox) into pericranial muscles has recently been shown to be safe and effective, although the mechanism of relieving pain is poorly understood.[31]

Cluster Headaches

Cluster headaches, previously known as Horton's cephalalgia, are less common than tension-type or migraine-type headaches; however, the severity of the headaches is much greater. They affect men more often than women. They are unilateral, excruciating, and located around the eyes or in the maxilla and usually occur in middle age. The episodes typically last minutes to 3 hours. They tend to occur several times per day for several weeks, and then long headache-free periods occur; hence, they occur in clusters. Ten percent of patients will have chronic episodes. No aura or nausea occurs, but the cephalalgia is associated with unilateral lacrimation, rhinorrhea, and injected conjunctiva. Ptosis and miosis may also occur. The headaches often wake a patient at the same time each night. Typically, the patient "searches for relief" by pacing or rocking back and forth. Alcohol and histamine injections are capable of precipitating attacks, and attacks are more common in the spring or fall.[33] The IHS classification of cluster headaches requires meeting the following diagnostic criteria: at least five episodes of classic cluster headache (characteristics as outlined above) occurring with a frequency of every other day up to eight times per day.[12]

Intracranial lesions, associated with the cavernous sinus or in other locations, as well as viral meningitis and skull base invasion by nasopharyngeal carcinoma, may produce cluster-like headaches that are more apt to be chronic. Therefore, a patient with atypical features of cluster headache should raise suspicions of an organic lesion, especially when periodicity of headaches is lacking or an incomplete response to known efficacious medications is noted.[33] The etiology is unclear, but several mechanisms are suspected. Circadian hormonal fluctuation suggests a hypothalamic dysfunction, whereas excitation of a nerve plexus in the carotid sheath and adventitia may increase trigeminal discharge rates, resulting in facial pain.[34] A subset of patients demonstrate an autosomal dominant transmission in their families.

Variants of cluster headache also occur. *Sluder's "lower half" neuralgia* refers to cluster headaches limited to the maxillary area with autonomic symptoms.[34] It is more common in women in their thirties and forties. *Chronic paroxysmal hemicrania* (*Sjaastad's syndrome*) is a variant with shorter, more frequent attacks, occurs more often in women, and is absolutely responsive to indomethacin.[12] *SUNCT syndrome* (an acronym for short-lasting unilateral neuralgiform headache attacks with conjunctival injection, tearing, sweating, and rhinorrhea) is a cluster-like headache, mostly in men, in which attacks last 10 to 60 seconds. It may be precipitated by touch to the nose or periorbital area, chewing, or ingestion of citrus fruits.[35]

Treatment of cluster headaches has abortive measures and prophylactic therapy. Along with providing reassurance for the patient, episodes can be ended with inhalation of 100% oxygen for 10 minutes. Administration of ergotamine sublingually or rectally works within 1 hour. Dihydroergotamine intramuscularly or intravenously provides relief in 30 minutes or 10 minutes, respectively. Sumatriptan works best when given subcutaneously, taking about 10 minutes to take effect. These four agents all work by inducing a central vasocontriction.[33] Intranasal 4% lidocaine may also be effective. At least three episodes of cephalalgia should be treated with any one agent before declaring it a failure and moving to another agent. Analgesics are too slow to be effective, and corticosteroids have not been proven to be beneficial.

Prophylactic medications are the mainstay of treatment. In general, they should be used during the months when the cluster headaches are occurring and gradually discontinued after at least a 2-week headache-free period.[34] Effective prophylaxis has been achieved with a variety of agents and is most often achieved using calcium channel blockers such as nifedipine or verapamil, low-dose ergotamine in sustained-release form, lithium carbonate, and methysergide for no greater than 4 months to avoid the rare complication of retroperitoneal fibro-

sis. Multidrug therapy may be necessary. Corticosteroids over 3 to 4 weeks may be helpful. Newer approaches have studied valproate, intranasal capsaicin, pizotifen, and phototherapy with bright light, with all finding significant success. Histamine desensitization may offer relief to patients with chronic cluster headaches refractory to treatment.[33] Invasive approaches for refractory cases include trigeminal nerve blockade and blockade of the sphenopalatine ganglion. The most effective procedure for typical cluster headache is radiofrequency ablation of the trigeminal ganglion.[34] When sphenopalatine ganglion (Sluder's) neuralgia is the diagnosis, removal of a precipitating septal spur or coagulation of the ganglion may be beneficial.

Post-traumatic Headache

Head trauma has been associated with the onset of acute or chronic headaches. These may be associated with postconcussive vertigo, decreased concentration, memory loss, and depression as part of the post-traumatic syndrome.[36] The pain is frontal, occipital, and constant, yet percussion of the skull is not painful, as is sometimes seen in subdural hematomas.[27] The IHS classification requires the onset of headache in less than 14 days after injury or regaining consciousness. The acute subclass lasts less than 8 weeks and the chronic greater than 8 weeks. Treatment is composed of reassurance, physical therapy and simple exercises, short courses of NSAIDs or over-the-counter analgesics, tricyclic antidepressants to modulate pain, injection of local anesthetic into trigger zones, carbamazepine for neuropathic pain, and avoiding recurrent head trauma.[36]

Headache Associated with Vascular Disorders

The IHS classifies eight vascular disorders that are associated with headache. The diagnosis hinges on historical, physical, and radiographic data, as well as evaluation of the CSF when indicated. Also, the headache must occur in close temporal relation to the onset of the underlying disorder. Two of the eight disorders, acute ischemic cerebrovascular disease or venous thrombosis, will not be reviewed here.

Intracranial Hematoma *Intracranial hematomas* (intracerebral, subdural, or epidural) often have headache as a presenting symptom. It is moderate to severe and on the side of the hematoma. Ipsilateral skull tenderness to percussion is common. Hemotympanum or mastoid bruising may occur. Associated symptoms include altered level of consciousness and sometimes nausea and vomiting, as well as focal central nervous system changes. There is usually an antecedent history of head trauma, although the trauma may be temporally remote. Anticoagulated patients may relay only minor head bumps. Diagnosis is confirmed by CT of the head. Treatment typically involves a craniotomy for drainage, but spontaneous resolution may occur over time. Once the hematoma is resolved, the headache ceases to occur.

Subarachnoid Hemorrhage *Subarachnoid hemorrhage* accounts for 1 to 4% of all headaches evaluated in the emergency department.[16] It typically presents as the worst headache of the patient's life. Besides the severe intensity, the cephalalgia is sudden and may gradually worsen; it is bilateral and associated with neck stiffness, fever, transient loss of consciousness, diplopia, or seizures. A prodromal or warning headache that is severe occurs in as many as 50% of patients several days or weeks before the major hemorrhage. Physical examination may reveal a changed mental status, restlessness, nuchal rigidity, retinal hemorrhages, papilledema, and focal or general neurologic signs such as cranial nerve III or VI palsy, aphasia, hemiparesis, or ataxia. If aneurysmal bleeding is the source, onset is in minutes; with arteriovenous malformation leaking, the onset is often more insidious, occurring in less than 12 hours. Rupture of a moyamoya lesion may also produce subarachnoid hemorrhage. The neurologic sequelae include devastating brain injury or death, but sometimes only minimal change occurs.

Diagnosis may be difficult as blood seen on CT may be mistakenly attributed to trauma if a fall occurred during any loss of consciousness rather than being suspected to be from a ruptured aneurysm. Hypertension may occur with the hemorrhage and can be confused with hypertensive crisis. Even electrocardiogram findings are common after subarachnoid hemorrhage and may incorrectly suggest cardiac disorders.[16] The evaluation should, however, begin with a noncontrast head CT with 3 mm cuts to find collections of blood. A lumbar puncture and evaluation of CSF should be per-

formed next if the CT was negative for blood or mass effect. Abnormal findings include elevated CSF pressure and red blood cells in the CSF. As traumatic taps may lead to the presence of red blood cells, their presence should be correlated with the CSF pressure. Xanthochromia is a much more specific finding for subarachnoid hemorrhage. If any test indicates that a subarachnoid hemorrhage has occurred, angiography and neurosurgical consultation are warranted urgently for treatment of the aneurysm or arteriovenous malformation. Radiographically detected unruptured aneurysms and vascular malformations should also receive neurosurgical evaluation to prevent catastrophic hemorrhage.[37]

Giant Cell Arteritis *Giant cell arteritis*, often called temporal arteritis, is associated with polymyalgia rheumatica. Essentially, they are different manifestations of the same condition. The cause is unknown, but speculation relates immune responses, infectious triggers, and genetic susceptibility.[38] The condition mostly affects women over 50 years of age. It presents as a new-onset, daily, intermittent or continuous, temporal localized headache that is moderate to severe, burning sharp or throbbing, unilateral or bilateral and lasts months to years.[17]

Patients with giant cell arteritis complain of aching of the jaw during chewing, weight loss, generalized fatigue, and low-grade fever and often have extremity pain attributable to coexistent polymyalgia rheumatica. Visual symptoms include blurring, scotomata, and even sudden blindness. Physical findings consist of palpably thickened or tender scalp arteries that may have a diminished or absent pulse. Palpation of muscle does not reveal trigger points, and muscle strength is typically normal. An erythrocyte sedimentation rate (ESR) is almost always elevated over 50 mm per hour in patients with giant cell arteritis. If the diagnosis is suspected, a temporal artery biopsy is essential to confirm the diagnosis. Inflammation in the artery is often segmental; therefore, a 5 cm long segment of artery should be excised; even then, false negatives may occur.[38] Histologic examination of involved vessels reveals necrosis in the vessel's media with infiltration of macrophages and giant cell formation. Color duplex ultrasonography is being investigated as a diagnostic tool to potentially replace biopsy. Treatment should begin immediately to avoid sudden blindness, a complication found in up to 30% of untreated patients. This is secondary to involvement of the ophthalmic artery. Prednisone is the drug of choice and should be administered in high daily doses of 60 mg. Pain reduces dramatically within days. Once headache has gone into remission and the ESR corrects, prednisone can be tapered over a period of weeks to a daily maintenance dose of 5 to 10 mg while continuing to check for a rise in the ESR.[38] Often treatment must be continued for several years to avoid the complication of blindness.[15] Sometimes the segmental excision of the temporal artery relieves the patient of localized tenderness. However, this does not preclude the need for treatment with corticosteroids.

Carotid Artery Pain *Carotid artery pain* may occur secondary to *dissection or idiopathic carotidynia* or after endarterectomy. Carotid artery dissection is characterized by head and neck pain on the side of the condition following trivial trauma in a young or middle-aged patient. It may cause referred otalgia,[39] Horner's syndrome, lower cranial nerve palsies, and pulsatile tinnitus. Hours later, signs of cerebral or retinal ischemia develop.[40] Dissection can be demonstrated by ultrasonography, MRI with MRA, intra-arterial angiography, or surgical findings. For treatment, early anticoagulation is essential to avoid the 15% risk of mortality. This is followed by warfarin for 6 months. If ischemic signs persist after acute heparinization, surgical intervention may be warranted.[40] Carotidynia is a self-limited inflammation of the carotid fascia or adventitia. It may be induced by viral infection. Physical findings include carotid tenderness, swelling, or intense pulsations. Structural evaluation of the artery is normal. Treatment with NSAIDs may alleviate the pain until spontaneous resolution occurs.

Arterial Hypertension *Arterial hypertension* can be acute or chronic. Chronic hypertension is not typically associated with headache. On the other hand, acute hypertension certainly may cause cephalalgia. The acute rise in blood pressure can be secondary to administration of pressors, vasoactive glomus tumors including pheochromocytoma, malignant hypertension (including hypertensive encephalopathy), or preeclampsia and eclampsia. Typically, headache begins with diastolic pressures greater than 115 mm Hg. It is throbbing, associated with nausea, and does not respond to analgesics.[27]

Treatment involves controlling and correcting the source of the hypertension. Once this is achieved, the headache resolves within 7 days.[12]

Headache Associated with Nonvascular Intracranial Disorders

The IHS classifies six intracranial conditions as nonvascular causes of headache. The symptoms, signs, and diagnostic investigations of four of these are presented below.

High CSF pressure associated headache may be caused by benign intracranial hypertension, formerly known as pseudotumor cerebri or otitic hydrocephalus, and high-pressure hydrocephalus that sometimes follows head trauma. *Benign intracranial hypertension* is defined by the findings of papilledema: an otherwise normal neurologic examination (with a rare exception being abducens palsy); no intracranial mass, venous sinus thrombosis, or ventricular enlargement on head CT or MRI; a normal protein, white cell count, and culture in the CSF; and, most importantly, an increased intracranial pressure of greater than 200 mm of water on lumbar puncture or intraventricular pressure monitor.[12] The associated headache is intermittent, retro-orbital, with gradual worsening, and is exacerbated by coughing or Valsalva's maneuver. Associated symptoms include unilateral or bilateral pulsatile tinnitus,[41] visual dimming occurring for a minute or two at a time, eventual constriction of visual fields, and possible blindness. It may be triggered by otitis media or mastoiditis, irregular menses, recent rapid weight gain, corticosteroid withdrawal, or ingestion of vitamin A, tetracycline, or nalidixic acid. Treatment consists of weight reduction, a low-sodium diet, and diuretics, specifically acetazolamide or furosemide. Lumbar puncture can be used and repeated to reduce CSF pressure acutely to relieve headache. Cerebrospinal fluid diversion by a ventriculoperitoneal shunt may be necessary for refractory cases or when visual fields fail to improve.[27] Transnasal, transethmoidal optic nerve decompression with release of the optic nerve sheath and annulus of Zinn may also be necessary to correct chronic visual field disturbances. *High-pressure hydrocephalus* differs from benign intracranial hypertension by the identification of enlarged ventricles on head CT or MRI.

Low CSF pressure also commonly precipitates a headache. This may be in response to a lumbar puncture or CSF fistula. The headache appears or worsens when the head is upright and resolves after lying down. The pain is constant when the head is upright and is located in the occipital region or vertex. Nausea is common, and vomiting relieves the headache during the regurgitation. Oculomotor or abducens palsies have been reported. Resolution is rapid (usually within several days) following lumbar puncture unless a persistent fistula occurs at the tap site. If this is suspected, a blood patch of the epidural space at the site of puncture is curative. This condition may occur without lumbar puncture and should raise suspicions of post-traumatic, iatrogenic, or spontaneous CSF fistula. If this occurs, site identification by intrathecal metrizamide CT or myelography should be performed and a surgical repair initiated.

Intracranial infection includes meningitis, encephalitis, brain abscess, or subdural empyema, which may present with a generalized, bilateral, severe, constant headache often with nuchal rigidity, positive Kernig's and Brudzinski's signs, and fever. The onset is rapid, usually over several hours.[17] Ambiguous presentations may occur in immunosuppressed patients, atypical infections such as tuberculosis, fungal meningitis, neurosyphilis, or epidural abscess. If a head CT has ruled out an intracranial mass, culture and antigenic studies of CSF from lumbar puncture are diagnostic.[27] Treatment requires appropriate intravenous antibiotic therapy and surgical drainage of abscesses.

Primary or metastatic *intracranial neoplasms* present with discrete or generalized, intermittent or continuous, mild to moderate, unilateral, dull headache that worsens over time in 30% of patients. Early morning headache is a warning symptom, as is exacerbation by Valsalva's maneuver or cough. Sudden vomiting, cranial nerve dysfunction, and seizures may occur.[17,27] Control of pain may be initially achieved with non-narcotic analgesics, but, eventually, narcotics or neuralgia medications become necessary. Removal of the neoplasm, when feasible, is usually curative of the headache.

Headache Associated with Substances or Their Withdrawal

Headaches are often caused by ingestion or withdrawal of specific substances. The onset of a headache during carbon monoxide poisoning and

the hangover after drinking alcohol serve as prime examples. Several particular conditions warrant greater description. Headache induced by chronic intake or withdrawal after chronic use of some medicines is a common phenomenon. Patients typically have a history of a different type of headache, usually migraine, in the past that did not occur daily. Escalating regular use of acute headache remedies to greater than 2 days per week may "transform" migraine into a chronic daily headache.[42] The cephalalgia occurs daily (often early in the morning) and is generalized, bilateral, dull, and of moderate severity. It may be associated with nausea and be brought on with mild exertion. Tolerance to the offending analgesic seems to allow less and less relief to the patient, resulting in the use of increased dosages. Finally, headache-free intervals cease to exist. Both ergotamine and dihydroergotamine have been implicated, as have non-narcotic and narcotic analgesics. Common examples include aspirin, acetaminophen, NSAIDs, barbiturates, codeine, hydrocodone, oxycodone, and propoxyphene. The IHS noted that use of ergot derivatives for 3 months puts the patient at risk, as does taking more than 50 g of aspirin or more than 100 tablets of an analgesic per month. Resolution within a month after discontinuing the substance is typical.

Treatment requires understanding on the part of the patient along with complete elimination of the problem medication for at least 2 months. No substitution to another analgesic should occur. Avoidance of tyramine and caffeine in the diet is recommended, whereas use of antidepressants is recommended and helpful. Nonpharmacologic modalities should be employed, more specifically, biofeedback, transcutaneous electrical nerve stimulation (TENS), and physical therapy. Inpatient withdrawal may be necessary. Prophylactic therapy for the primary headache should also be instituted to decrease the need for ongoing abortive medications.

Headache from withdrawal of a chronically used substance may also occur. It typically follows a period of abstinence from the agent of less than 12 to 48 hours and is relieved by ingestion of the substance. Persistent elimination of the substance allows resolution of the headache within 14 days. Frequent examples consist of caffeine, ergot derivatives, and narcotics.[12] The headache is diffuse, dull, and mild to moderate in severity. It may be associated with nervousness, restlessness, nausea and vomiting, insomnia, and tremor. Persistent abstinence is the mainstay of treatment but may require hospitalization to prevent the patient from consuming more of the offending medication.

Headaches Associated with a Metabolic Disorder

Hypoxia, as seen in high-altitude headache, is a known source of cephalalgia. Descent and oxygen therapy are the treatment. A more common cause of chronic hypoxia is found in sleep apnea syndrome, in which morning headache is common. It is diffuse, dull, and mild and may be associated with short-term memory loss and inability to concentrate. The diagnosis is suggested by more significant symptoms of the syndrome, specifically fatigue, daytime somnolence, and apneas witnessed by a bed partner. Resolution of the headache occurs quickly following correction of the hypoxia by continuous positive airway pressure or surgical intervention.

Headache or Facial Pain Associated with Craniofacial or Cervical Disorders

When pathology afflicts any extracranial organ of the head or the skull itself, pain may ensue. Facial or head pain is localized to the affected structure but may radiate or be referred. If the pathology is corrected, headache resolves in less than 1 month. Examples affecting the cranial bones include osteomyelitis, multiple myeloma, and Paget's disease.

Cervicogenic headache is a relatively common condition created by a disorder afflicting the cervical spine, often following neck trauma. Women are more frequently affected than men. It is characterized by fluctuating or constant headache that is unilateral, unchanging, dull, and nonpulsatile. It tends to start at the occiput and radiate to the frontal region, where the pain is most intense. If it is bilateral, one side is usually of worse intensity. Pain may also radiate into the shoulder or arm. Attacks are precipitated by awkward positions or neck movement or palpation of trigger points such as the greater occipital nerve or over the C2 vertebral body.[43] Associated symptoms of nausea and photophobia or phonophobia are present in 50% of patients. Neck examination reveals a reduced range of motion of the spine and possibly muscle tenderness, spasm, hypertrophy, or atrophy. Criteria for its

diagnosis have been proposed by Sjaastad and Fredriksen.[44] These criteria essentially define all of the features of the syndrome as represented above.

Cervical radiographic data reveal either abnormalities on flexion and extension films, abnormal posture, or evidence of bone or joint pathology. Ipsilateral injection of local anesthetics for blockade of the greater occipital nerve or the C2 root will relieve pain not only in the innervation zones but also in the frontotemporal region.[44] Treatment centers around physical therapy and stretching exercises, use of NSAIDs, and chronic nerve blockade for refractory cases.

Sinonasal disorders are a frequent source of headaches but are probably overcredited by the population as a whole. However, frontal headache and facial pain are two of the three major symptoms suggesting the presence of sinusitis, the other being purulent nasal drainage. Acute sinusitis is a leading cause of facial pain, second only to dental disorders. The IHS classified headache from sinus origin as when it is associated with purulent drainage and they begin together; when plain radiography, CT, MRI, or transillumination reveal opacification or air-fluid levels in one or more sinuses; and when the sinusitis is treated, the headache resolves.[12] The cephalalgia is pressure-like or dull, of moderate intensity, unilateral or bilateral, periorbital, worsened by bending or with Valsalsa's maneuver, worse in the morning, and associated with purulent nasal drainage, nasal obstruction, altered sense of smell, exacerbating asthma, cough, malaise, and dizziness. Nausea and visual change are infrequent associated symptoms.[8,39,45]

Pain with the inflammatory process lasts days or weeks with fluctuating severity. The location and extent of the sinusitis do not correlate well with the severity or site of pain.[45] However, a few tendencies are common. The area overlying an infected sinus is often tender. Pain for sinusitis referred to upper maxillary teeth is typically originating in the maxillary sinus. Occipital or vertex pain from sinusitis is most likely to represent sphenoid sinus disease. Any infected sinus can refer pain to the frontal, retroorbital, and temporal regions. Physical examination finds tenderness to percussion over involved sinuses. Transillumination is infrequently useful to delineate the presence of sinusitis. Other physical findings center around inflammatory changes in the nose. Radiographic imaging is helpful to confirm an uncertain diagnosis of sinusitis or to evaluate response to treatment. As plain films have poor sensitivity, a screening coronal sinus CT has become the imaging study of choice. Treatment requires medical therapy to reverse the inflammatory cause of the headache and includes antibiotics, decongestants, and possibly corticosteroids, with analgesics used in the interim for symptomatic relief until the sinusitis is reversed.

Rhinologic headaches other than those caused by sinusitis also occur, albeit much more rarely. Nasal anatomic abnormalities occasionally associated with facial pain include impacted septal deviations or spurs, hypertrophic turbinates, and even an occasional large maxillary retention cyst.[3] However, most sinonasal neoplasms do not cause pain. The cause of rhinogenic pain may be related to direct sensory stimulation from two mucosal surfaces contacting each other. The pain is usually dull and aching and unilateral. An anatomic abnormality can be seen on examination after decongesting the nose. Topical lidocaine can be used as a diagnostic test as the pain should diminish or disappear after application if the cause is mucosal contact. Medication may effectively control turbinate hypertrophy, but surgical approaches are effective in managing the anatomic defects.[39]

Elongated styloid process, better known as *Eagle's syndrome*, consists of recurrent throat and neck pain exacerbated by swallowing. Otalgia, dysphagia, and foreign-body sensation are often present and are secondary to an elongated styloid process impinging on the carotid plexus or branches of cranial nerve IX. It may develop subsequent to tonsillectomy because of inflammatory changes. Palpation of the styloid process in the tonsil fossa exacerbates the dull pain, whereas injection of a local anesthetic in a diagnostic maneuver provides relief. A lateral plain film of the head reveals a styloid process greater than 2.5 cm long.[46] Reassurance and medical therapy with NSAIDs or corticosteroid injection around the stylohyoid ligament may be successful. Otherwise, treatment involves surgical shortening of the calcified stylohyoid ligament (elongated styloid process) through the tonsillar fossa.[39]

A burning sensation in the oral mucosa is an uncommon but not rare complaint. The condition is known as *glossodynia, burning mouth syndrome*, or oral dysesthesia. It most often occurs in women and is chronic, with a prevalence of about 5%.[47] Local,

systemic, and psychogenic factors play a role in causation. Local factors include mucosal lesions, oral candidiasis, local chemical irritation, and xerostomia. Systemic factors consist of iron, zinc, or vitamin B_{12} deficiency and the use of female replacement hormones, sedatives, antihypertensives, or antidepressants. Anxiety and depression are associated with glossodynia. Physical examination is usually normal without a mucosal lesion. Treatment includes reassurance; discontinuing smoking and alcohol-containing mouthwashes; drinking cooler coffee; therapeutic trials of nystatin oral rinse or dexamethasone solution to rinse and expectorate; iron, zinc, or vitamin B_{12} replacement therapy; oral lubrication or use of salivary stimulants; and using antidepressants.

Temporomandibular joint disorders are frequent causes of facial pain and headache. These disorders are divided into two major groups: those with demonstrable organic disease (uncommon) and those of myofascial origin from masticatory muscles (very common). The IHS established the following diagnostic criteria for TMJ disease, at least two of which must be met: pain with jaw motion, noise with jaw movement, decreased range of motion, and a tender joint. These are associated with abnormal TMJ radiographs and pain that is mild to moderate and centered over the TMJ.[12] Women are far more likely to suffer from TMJ pain. The pain is located preauricularly and in the temporal region of the head; is intermittent, lasting hours to days; is mild to moderate in intensity; can be unilateral or bilateral; and is exacerbated by jaw movement. Associated features include otalgia, bruxism, trismus, TMJ crepitus, and jaw locking. Usually, no muscular trigger point exists, and there are no nausea or visual changes. Etiology of the discomfort includes diseased joints that can occur on the basis of arthritis (the most common cause), traumatic causes such as fractures or dislocations, internal derangements of the intra-articular disk, or developmental defects.

Physical findings include audible joint clicking on one side or both and decreased range of motion. Normal vertical opening between central incisors ranges from 42 to 55 mm.[17] Palpation should cover the head and neck musculature and TMJs including intraoral pterygoid palpation, noting any asymmetry, tenderness, spasm, hypertrophy, atrophy, triggering of referred pain, or crepitus when opening or closing. If pathology is suspected, a TMJ radiographic series may reveal a joint abnormality. The usual indications for imaging include suspected fractures, degenerative joint disease, ankylosis, or tumors.

Management of TMJ disease includes reassurance, education, NSAIDs, restriction of jaw opening, a soft diet, and physical therapy. Occasionally, biofeedback or corticosteroid injection may be helpful. If pain is chronic, tricyclic antidepressants may be beneficial. Because of the potential for dependency in chronic conditions, narcotics are best avoided. Acute disk or condylar displacements may require manual reduction to relieve pain and restore range of motion or relieve a joint locked open or closed. Often sedation or general anesthesia is necessary to relax the muscles to facilitate the manual reduction. The use of a bite appliance may alleviate chronic headache or muscle pain secondary to bruxism or joint pain secondary to anterior disk displacement.[48] Arthrocentesis, sclerotic therapy, and arthroscopic surgery are reserved for chronic dislocations, disk derangements, or condylar anomalies.

Oromandibular dysfunction involves pain in the TMJ without organic disease. Other terms describing this condition include myofascial pain dysfunction syndrome and TMJ pain dysfunction syndrome. The etiology of pain seems to have its origin in myofascial tissues or central processes disinhibiting sensory pain pathways. The IHS system classified this as a tension-type headache that does not purely meet the criteria of episodic or chronic tension-type headaches. A diagnosis requires at least three of the following: the creation of noise with jaw movement, limited range of motion, pain during jaw use, locking of the jaw, or a history of clenching or grinding of the teeth.[12] The pain is typically a continuous, dull, aching, poorly localized, usually unilateral discomfort of mild-to-moderate degree involving the ear, angle of the jaw, and temple. The pain rarely wakes the patient, and the location of pain may shift. The patterns of referred pain often do not make neurologic sense. It may be associated with swelling, numbness, stiffness, or erythema. The hallmarks of this myalgia are the presence of trigger points that, when palpated, reproduce the referred pain.[49] Examination will reproduce the pain when masticatory and neck muscles are palpated. Full range of jaw motion may be lost with this disorder. A trigger point injection of an anesthetic (like 1% procaine) can confirm that point's significance by relieving the pain.[49]

Therapy consists of reassurance, modification to a soft diet, physical therapy, trigger point local

anesthetic injection, and medications. Amitriptyline for at least 2 months is an effective choice, whereas NSAIDs may or may not be. Occlusal splints relieving bruxism may help, as might attention to abnormal occlusal factors that can precipitate the disorder.[32] Because of the self-limited nature of the condition, invasive approaches should be avoided.

Cranial Neuralgias

Cranial neuralgias are conditions affecting nerves with sensory functions in the head and neck, resulting in severe stabbing or throbbing pain in the distribution of the involved nerve. They can involve any cranial nerve with sensory fibers or cervical roots one, two, or three. The conditions are subdivided into persistent painful disorders and paroxysms of pain (tic-like) disorders.

Traumatic injury to a nerve or inflammatory changes can result in chronic neuralgia. Trauma or surgery to the cranium or face may result in entrapment neuritis or formation of a neuroma, typically 2 to 6 months later. The occurrence is greatest in the third branch of the trigeminal nerve because of the frequency of injury related to mandible fracture or tooth extraction. Symptoms include hypersensitivity and pain to light touch, pain in an area of skin that has lost its sensory innervation, and aggravation by cold or emotional duress. The pain is sharp and lancinating. A central cause of a neuritic pain can occur as *anesthesia dolorosa* following surgical ablation of the trigeminal ganglion. The condition is characterized by sharp pain and numbness in the distribution of any or all branches of the trigeminal nerve after trigeminal rhizotomy or trauma. Treatment uses anticonvulsant medications, in particular, carbamazepine, or sometimes baclofen or clonazepam. Amitriptyline has also been employed with success.[36] Injection of local anesthetics into trigger points or surgical resection or repair of neuromas is necessary for refractory cases.

Inflammatory conditions are also well known to cause neuralgia. A prime example is *acute herpes zoster* of a branch of the trigeminal nerve, the seventh cranial nerve, or cervical roots. The ophthalmic branch of cranial nerve V is the most commonly involved nerve. An intense burning or stabbing pain is in the distribution of the involved nerve and is followed within 1 week by a herpetic eruption in the skin distribution of the same nerve. Motor divisions of involved nerves may be paretic. The pain subsides within 6 months of the onset, but the motor palsies have a poor prognosis for complete recovery. Treatment of the acute phase consists of a 10-day course of prednisone and acyclovir. Nonsteroidal anti-inflammatory drugs or opiates may be necessary to control pain. Acute herpes zoster is common in lymphoma patients, so a new outbreak should raise suspicions about that possible comorbidity.

Chronic postherpetic neuralgia exists when herpes zoster pain persists for more than 6 months. It remains in the distribution of the originally afflicted nerve. It is more likely to occur in patients who are over 60 years old when the acute infection starts, and in this group, it is less likely to spontaneously resolve.[12] Nonsteroidal anti-inflammatory drugs and opiates fail to relieve the neuralgia. Instead, anticonvulsants are more effective and may be paired with tricyclic antidepressants or baclofen to enhance their efficacy.

Other inflammatory neuralgia syndromes include the following: (1) *Tolosa-Hunt syndrome*: episodic unilateral orbital pain with palsies in one or more of cranial nerves III, IV, and VI. Episodes last about 8 weeks untreated but generally resolve within 3 days after starting corticosteroids. It has been related to granulation tissue at the cavernous sinus or superior orbital fissure or in the orbit.[12] (2) *Gradenigo's syndrome*: progressive, intense, unilateral, retro-orbital pain with abducens palsy and otorrhea secondary to petrous apicitis of the temporal bone. (3) *Raeder's paratrigeminal neuralgia*: retro-orbital pain with intact trigeminal sensation and an incomplete Horner's syndrome, resulting in an ipsilateral dilated pupil. The paratrigeminal space housing the gasserian ganglion of cranial nerve V, cranial nerves III and IV, and the carotid artery may be encroached on by abnormal vascular structures and produce the syndrome.[9]

Trigeminal neuralgia (TGN), previously called tic douloureux, is the most common cranial neuralgia and most often affects adults over 50 years old and women more than men. Typically, recurrent episodes of unilateral, excruciating, stabbing pain occur most often in the distribution of the maxillary and mandibular branches of the trigeminal nerve. During an episode, ipsilateral twitching may occur—hence the name "painful tic." They occur without warning, while patients are awake, and recur several to many times a day, with each episode lasting seconds to minutes. Numbness does not occur.

Light touching of the face may precipitate an attack, as can movement of the trigger zone by talking, chewing, or shaving. Physical findings include an intact neurologic examination and the presence of a trigger zone most often located in the nasolabial fold, lips, or gums. Diagnostic evaluation should include an MRI of the head and a lumbar puncture if the MRI is normal. The former is to rule out central causes including vascular anomalies or aneurysm, tumor, cholesteatoma, or multiple sclerosis (MS). Patients with MS are predisposed to TGN, and 2% of patients with TGN have multiple sclerosis.[17] The latter is to rule out MS, neurosyphilis, and cryptococcal or tuberculous meningitis. Treatment of idiopathic TGN includes patient education and reassurance with pharmacologic control. Carbamazepine will provide symptomatic relief acutely in the majority of patients.[17] Other agents shown to be effective include gabapentin, phenytoin, baclofen, sodium valproate, and chlorphenesin. Any drug should be employed for at least 2 weeks, and medications may be used in combination. Tricyclic antidepressants and NSAIDs may also be beneficial. Surgical radiofrequency ablation, specifically trigeminal rhizotomy, is recommended for patients refractory to medical therapy, whereas temporary blockade may be indicated for patients experiencing intense activation of the disease. See Chapter 20 for vascular compression syndromes and their management.

Glossopharyngeal neuralgia is a disorder similar to TGN but affecting cranial nerve IX instead. Men in their forties or fifties are more likely to have this disorder. The pain characteristics include repetitive, brief attacks of lancinating pain located in the soft palate, base of the tongue, pharynx, and ear. The trigger point is located in the tonsil fossa and can be provoked by swallowing or yawning. Associated symptoms may include hiccoughing, nausea, vertigo, tinnitus, aural fullness, hearing loss, dysgeusia, bradycardia, and syncope. Evaluation and management are identical to those for TGN.[39]

Superior laryngeal neuralgia is characterized by unilateral lancinating attacks of pain at the side of the larynx over the thyrohyoid membrane. Pain radiates to the ear and submandibular region and is precipitated by swallowing, straining the voice, playing a musical instrument, or turning the head. The pain paroxysms last minutes to hours for days or weeks at a time. Dysphonia may be present secondary to involvement of the external branch of the superior laryngeal nerve affecting the cricothyroid muscle. Injection of local anesthetic as a nerve blockade confirms the diagnosis when pain is relieved.[39] The condition is managed with anticonvulsants as in TGN or with surgical sectioning of the superior laryngeal nerve.

Occipital neuralgia is a paroxysmal stabbing aching pain over the occiput in the distribution of the greater or lesser occipital nerve combined with reduced sensation in the same area. Associated symptoms include visual disturbances, dizziness, nausea, tinnitus, and scalp paresthesias. The physical findings include a positive Tinel's sign or palpable tenderness of the involved occipital nerve.[50] Confirmation of the diagnosis is made by obtaining relief from pain after an injection of local anesthetic is performed to block the nerve. Management consists of heat, physical therapy, rest, TENS units, NSAIDs, baclofen, or carbamazepine. Cervical collars are of questionable benefit. Invasive measures for refractory cases include local blockade with bupivacaine in combination with an injectable corticosteroid, alcohol blockade, or even nerve section.[50]

Unclassifiable Facial Pain (Atypical Facial Pain)

Atypical facial pain fails to fit the profile of any specific craniofacial condition. The prosopalgia (facial pain) is unilateral or bilateral, is located in the face and upper neck, and is a constant ache or burning with excruciating exacerbations. It does not localize to anatomic regions or have relation to specific structures as TMJ arthralgia or myofascial pain does. It lasts for years, is exacerbated by stress, and is associated with paresthesias of the face or mouth.[17] Comorbidities include chronic fatigue, depression, personality disorders, irritable bowel syndrome, and other idiopathic pain disorders. The cause is unknown, but theories propose central causes for which neuroregulation of pain input is disinhibited, systemic causes for which tyramine metabolism is deficient,[51] and psychogenic causes for which personality disorders are linked to the development of chronic pain conditions.

Treatment involves a multidisciplinary approach. Each patient should be reassured, receive only essential dentistry, and sparingly use analgesics. Tricyclic antidepressants are the mainstay of therapy; they should be taken at night and titrated up to a desired response. Other options include serotonin

reuptake inhibitors, monoamine oxidase inhibitors, benzodiazepines, and phenothiazine. Once a good response has occurred, therapy should be maintained until the patient has been pain free for several months before it is discontinued.[51] Psychiatric evaluation may be beneficial. Refractory cases might respond to one of the several neurosurgical procedures.

REFERENCES

1. Dalessio D, Silberstein S. Wolff's headache. New York: Oxford University Press; 1993.
2. Sessle BJ. Neural mechanisms and pathways in craniofacial pain. Can J Neurol Sci 1999;26 Suppl:S7–11.
3. Chow JM. Rhinologic headaches. Otolaryngol Head Neck Surg 1994;111:211–8.
4. Wright EF. Referred craniofacial pain patterns in patients with temporomandibular disorder. J Am Dent Assoc 2000;131:1307–15.
5. Jensen R, Olesen J. Tension-type headache: an update on mechanisms and treatment. Curr Opin Neurol 2000;13:285–9.
6. Hargreaves RJ, Shepheard SL. Pathophysiology of migraine—new insights. Can J Neurol Sci 1999;26 Suppl:S12–9.
7. Hamel E. The biology of serotonin receptors: focus on migraine pathophysiology and treatment. Can J Neurol Sci 1999;26 Suppl:S2–6.
8. Schor DI. Headache and facial pain—the role of the paranasal sinuses: a literature review. Cranio 1993;11:36–47.
9. Ballenger JJ. Headache and neuralgia of the face. In: Ballenger JJ, Snow JB, editors. Otorhinolaryngology head and neck surgery. Baltimore: Williams & Wilkins; 1996. p. 158–62.
10. Ansemi B, Baldi E, Cassaci F, Salman S. Endogenous opinoids in cerebrospinal fluids and blood in idiopathic headache sufferers. Headache 1998;20:294.
11. Jensen R. Pathophysiological mechanisms of tension-type headache: a review of epidemiological and experimental studies. Cephalalgia 1999;19:602–21.
12. Headache Classification Committee of the International Headache Society. Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. Cephalalgia 1988;8 Suppl 7:1–96.
13. Saper JR. Diagnosis and symptomatic treatment of migraine. Headache 1997;37 Suppl:S1–14.
14. Smetana GW. The diagnostic value of historical features in primary headache syndromes: a comprehensive review. Arch Intern Med 2000;160:2729–37.
15. Lance JW. Headache and face pain. Med J Aust 2000;172:450–5.
16. Edlow JA, Caplan LR. Avoiding pitfalls in the diagnosis of subarachnoid hemorrhage. N Engl J Med 2000;342:29–36.
17. Montgomery MT. Extraoral facial pain. Emerg Med Clin North Am 2000;18:577–600.
18. Stewart WF, Lipton RB, Celentano DD, Reed ML. Prevalence of migraine headache in the United States. JAMA 1992;267:64–9.
19. Annequin D, Tourniaire B, Massiou H. Migraine and headache in childhood and adolescence. Pediatr Clin North Am 2000;47:617–31.
20. Solomon S. Migraine diagnosis and clinical symptomatology. Headache 1994;34 Suppl:S8–12.
21. Wilkinson M. Migraine treatment: the British perspective. Headache 1994;34 Suppl:S13–6.
22. Lipton RB, Stewart WF, Stone AM, et al. Stratified care vs step care strategies for migraine. JAMA 2000;284:2599–605.
23. Silberstein SD, Young WB. Safety and efficiency of ergotamine tartrate and dihydroergotamine in the treatment of migraine and status migrainosus. Neurology 1995;45:577–84.
24. Young WB. Appropriate use of ergotamine tartrate and dihydroergotamine in the treatment of migraine: current perspectives. Headache 1997;37 Suppl: S42–5.
25. Logemann CD, Rankin LM. Newer intranasal migraine medications. Am Fam Physician 2000;61: 180–6.
26. Maizels M, Scott B, Cohen W, Chen W. Intranasal lidocaine for treatment of migraine. JAMA 1996; 276:319–21.
27. Finkel AG, Mann JD, Lundeen TF. Headache and facial pain. In: Bailey BJ, Calhoun KH, Deskin RW, et al, editors. Head and neck surgery-otolaryngology. 2nd ed. Philadelphia: Lippincott-Raven; 1998. p. 287–304.
28. Becker WJ. Evidence based migraine prophylactic drug therapy. Can J Neurol Sci 1999;26 Suppl:S27–32.
29. Truelove EL. The chemotherapeutic management of chronic and persistent orofacial pain. Dent Clin North Am 1994;38:559–88.
30. Silberstein SD. Tension-type headaches. Headache 1994;34 Suppl:S2–7.
31. Silberstein SD, Lipton RB. Chronic daily headache. Curr Opin Neurol 2000;13:277–83.
32. Benoliel R, Sharav Y. Craniofacial pain of myofascial origin: temporomandibular pain and tension-type headache. Compend Contin Educ Dent 1998;19: 701–21.

33. Mendizabal JE, Umana E, Zweifler RM. Cluster headache: Horton's cephalalgia revisited. South Med J 1998;91:606–17.
34. Mathew NT. Cluster headache. Neurology 1992;42 Suppl 2:22–31.
35. Pareja JA, Sjaastad O. SUNCT syndrome: a clinical review. Headache 1997;37:195–202.
36. Benoliel R, Eliav E, Elishoov H, Sharav Y. Diagnosis and treatment of persistent pain after trauma to the head and neck. J Oral Maxillofac Surg 1994;52: 1138–47; discussion 1147–8.
37. Raps EC, Rogers JD, Galetta SL, et al. The clinical spectrum of unruptured intracranial aneurysms. Arch Neurol 1994;51:447–8.
38. Meskimen S, Cook TD, Blake RL. Management of giant cell arteritis and polymyalgia rheumatica. Am Fam Physician 2000;61:2061–8, 2073.
39. Levine HL. Otorhinolaryngologic causes of headache. Med Clin North Am 1991;75:195–210.
40. Guillon B, Lévy C, Bousser MG. Internal carotid artery dissection: an update. J Neurol Sci 1998;153:146–58.
41. Sismanis A. Pulsatile tinnitus. Am J Otol 1998; 19:472–7.
42. Mathew NT, Kurman R, Perez F. Drug induced refractory headache—clinical features and management. Headache 1990;30:634–8.
43. Vincent MB. Cervicogenic headache: clinical aspects. Clin Exp Rheumatol 2000;18 Suppl:S7–S10.
44. Sjaastad O, Fredriksen TA. Cervicogenic headache: criteria, classification and epidemiology. Clin Exp Rheumatol 2000;18 Suppl:S3–6.
45. Tarabichi M. Characteristics of sinus-related pain. Otolaryngol Head Neck Surg 2000;122:842–7.
46. Balbuena L, Hayes D, Ramirez SG, Johnson R. Eagle's syndrome (elongated styloid process). South Med J 1997;90:331–4.
47. Hakeberg M, Berggren U, Hagglin C, Ahlqwist M. Reported burning mouth symptoms among middle-aged and elderly women. Eur J Oral Sci 1997; 105: 539–43.
48. Major PW, Nebbe B. Use and effectiveness of splint appliance therapy: review of literature. Cranio 1997:15:159–66.
49. Graff-Radford SB. Facial pain. Curr Opin Neurol 2000;13:291–6.
50. Kuhn WF, Kuhn SC, Gilberstadt H. Occipital neuralgias: clinical recognition of a complicated headache. A case series and literature review. J Orofac Pain 1997;11:158–65.
51. Feinmann C, Peatfield R. Orofacial neuralgia. Diagnosis and treatment guidelines. Drugs 1993;46: 263–8.

CHAPTER 36

Neoplasms of the Nose and Paranasal Sinuses

Ara A. Chalian, MD, David Litman, MD

The poor prognosis for patients with nasal and paranasal sinus malignancy has led many investigators to focus their attention on it. Unfortunately, no true consensus has been reached regarding many aspects of treatment for these disease entities. This stems from the wide variety of tumors found in the area, the relatively short clinical course of patients afflicted with these neoplasms, and the rarity with which the tumors occur. However, with recent developments in fiber optics, radiographic imaging, and a growing consensus on staging and nomenclature, a vast amount of clinical information is now becoming available.

PATHOLOGY

Despite the rarity of nasal and paranasal sinus neoplasms, many histologically diverse types are described in the literature (Table 36–1). The following discussion focuses on the most common pathology seen in this area.

Whereas most sinus neoplasms are malignant, in the nasal cavity, there is a fairly even distribution between benign and malignant disease.[1] The majority of these benign lesions are papillomas that may be subdivided into three categories: fungiform (50%), inverted (45%), and cylindrical cell (5%).[2]

TABLE 36–1. Neoplasms of the Nose and Paranasal Sinuses

Benign	*Malignant*
Epithelial tumors	Epithelial tumors
Papillomas	Squamous cell carcinomas
Adenomas	Adenocarcinomas
	Adenoid cystic carcinomas
Nonepithelial tumors	Mucoepidermoid carcinomas
Osteomas	Melanomas
Fibromas	Esthesioneuroblastomas
Chondromas	Teratomas
Hemangiomas	Teratocarcinomas
Nerve sheath tumors	
	Nonepithelial tumors
	Rhabdomyosarcomas
	Neurogenic sarcomas
	Leiomyosarcomas
	Fibrosarcomas
	Angiosarcomas
	Hemangiopericytomas
	Osteogenic sarcomas
	Chondrosarcomas
	Lymphomas
	Extramedullary plasmacytomas
	Giant cell tumors

Fungiform papillomas present as exophytic septal lesions. Histologically, they have a nonkeratinizing squamous epithelium covering a fibrovascular stroma. In contrast to the other varieties of papilloma, they are not associated with malignant degeneration.[2]

Inverted papillomas (IPs) are characterized by a squamous or transitional cell epithelium surrounding a fibrovascular stroma with endophytic growth. They are most often found on the lateral nasal wall and have been reported to be associated with squamous cell carcinoma in 5 to 15% of patients.[3]

Cylindrical cell papillomas, also known as oncocytic schneiderian papillomas, share with IPs both their localization to the lateral nasal wall and their association with malignant disease. Pathologically, the epithelial lining contains multiple layers of eosinophilic cells, goblet cells, and microcysts full of mucin.[2]

Additional benign tumors of epithelial origin include adenomas, cholesteatomas, and dermoids. Adenomas of the sinonasal tract are histologically identical to those arising in other areas of the body. They occur most commonly in the fourth to seventh decade and usually involve the nasal septum. Excision is curative in over 90% of all cases. Similarly, cholesteatomas and dermoids pathologically resemble those seen in other areas of the body and are relatively rare in the sinonasal region. Cholesteatomas are composed of desquamated epithelium, whereas dermoid tumors contain specialized dermal appendages.[4] When they do occur in the sinuses, they are typically in the frontal or ethmoid sinuses.[4]

Fibrous dysplasia has been described as occasionally arising within the maxilla. It occurs in both polyostotic and monostotic forms, with monostotic being more common. Growth is slow, and treatment consists of local excision for lesions that are physically obstructive or cosmetically deforming.[4] The surgery typically is conservative and involves a sculpting excision of the lesion to restore the tumor site to its natural contour.

Malignant epithelial neoplasms constitute the majority of sinonasal tumors, representing 45 to 80% of all sinus neoplasia.[5,6] Of these, squamous cell carcinoma is the most common. Sixty percent of these neoplasms arise in the maxillary sinus, 20 to 30% in the nasal cavity, 10 to 15% in the ethmoid sinuses, and 1% in the frontal and sphenoid sinuses.[2] As in other areas of the head and neck, the tumor may present with varying degrees of differentiation.

Glandular carcinomas are the next most common tumors, comprising 4 to 15% of all sinus neoplasms.[6] Of these, adenocarcinomas are the most common, representing 5 to 19% of nasal and paranasal sinus tumors.[7–9] There is a lack of consistency in the literature regarding the classification of these tumors. In general, these tumors are classified as low grade on the basis of uniform glandular architecture and uniform cytologic characteristics. High-grade adenocarcinomas have a predominantly solid growth pattern and moderate to prominent nuclear pleomorphism.[5] The prognostic significance of the two types is debatable, but it is felt that the low-grade adenocarcinomas have a propensity toward local recurrence and rarely metastasize, whereas high-grade adenocarcinomas more frequently display regional and distant metastases.[5]

Adenoid cystic carcinomas are somewhat less common than adenocarcinomas but still comprise a sizable portion of sinonasal tumors in most studies. The tumor consists of groups of small cells arranged in one of several patterns described as tubular, cribriform, and solid.[5,10] Low-grade adenoid cystic carcinomas are defined as those with less than 30% solid component.[5] Conversely, high-grade carcinomas consist of tumors with a greater than 30% solid component (Figure 36–1). High-grade carcinomas are characterized by a shorter survival period and more frequent incidence of distant metastases.[5] Both subtypes have a predilection for perineural invasion.[5,11]

Mucoepidermoid carcinoma (Figure 36–2) is an extremely rare form of glandular carcinoma. It is composed of a combination of squamous cells and glandular, mucus-producing, basal cells.[2] They are notable for their propensity toward distant metastases.[10]

Although rare, two tumors of neuroectodermal origin are occasionally encountered. These are sinonasal melanoma and olfactory neuroblastoma (esthesioneurocytoma, esthesioneuroblastoma). Melanoma of the sinuses represents less than 7% of all sinonasal malignancies. It does not differ significantly from that in other parts of the body, although 30% have been noted to be amelanotic.[5,12,13] Melanoma may have plasmacytoid, spindle cell, or epithelioid cytologic features.[12] Additionally, archi-

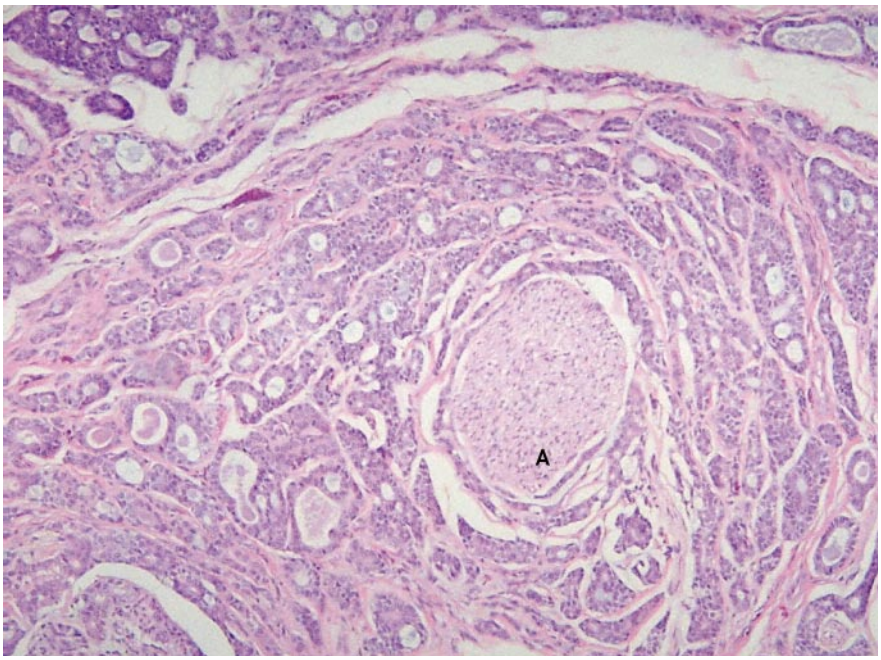


FIGURE 36–1. Photomicrograph of adenoid cystic carcinoma showing low-grade histology and cribriform and tubular architecture with nerve (A), demonstrating perineural invasion. (Hematoxylin-eosin stain, ×20 original magnification.) Courtesy of Michael D. Feldman, MD.

tectural patterns vary and may be sarcoma-like, epithelioid, or pleomorphic.[12] Immunohistochemical staining is helpful and should include anti-S100, HMB-45, and antivimentin.[2,12]

Olfactory neuroblastoma (Figure 36–3) is the second tumor of neuroectodermal origin that may arise in the paranasal region. This neoplasm presents with one of two cell patterns. It may have either a nesting pattern of small cells surrounded by stroma or diffuse sheets of tumor cells with a lack of stroma.[2] Immunohistochemical staining is helpful as esthesioneuroblastoma displays a sustentacular cell pattern of S-100 staining as well as positive staining for neurofilament protein, neuron-specific enolase, and synaptophysin.[2,12,14] Electron microscopy aids in the diagnosis of this tumor by allowing one to identify cells containing microtubules and neurofilaments that are specific to this neoplasm.[2]

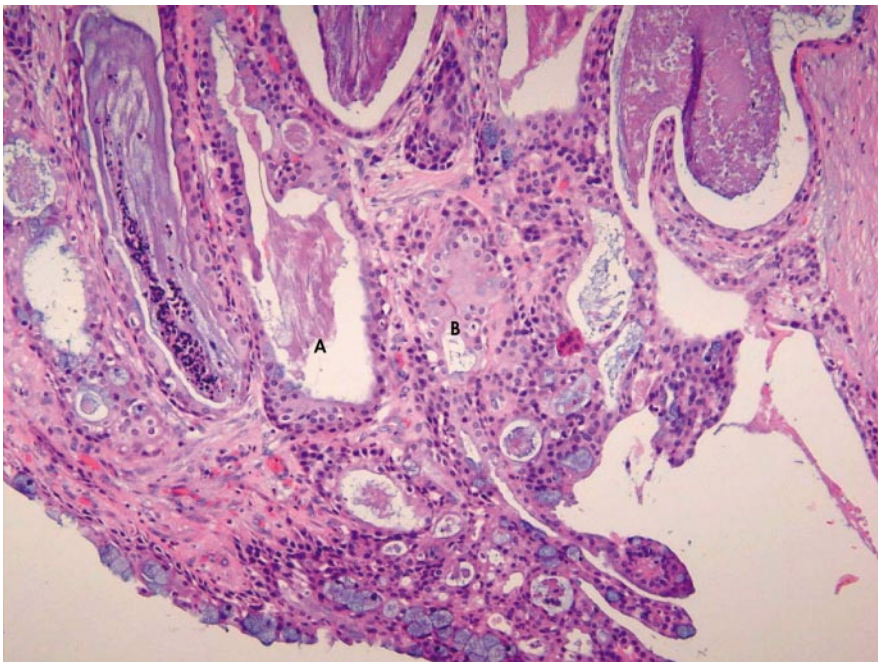


FIGURE 36–2. Photomicrograph of low-grade mucoepidermoid carcinoma showing cystic spaces filled with mucus (A) lined by mucus-producing cells and central epithelioid cells (B). (Hematoxylin-eosin stain, ×20 original magnification.) Courtesy of Michael D. Feldman, MD.

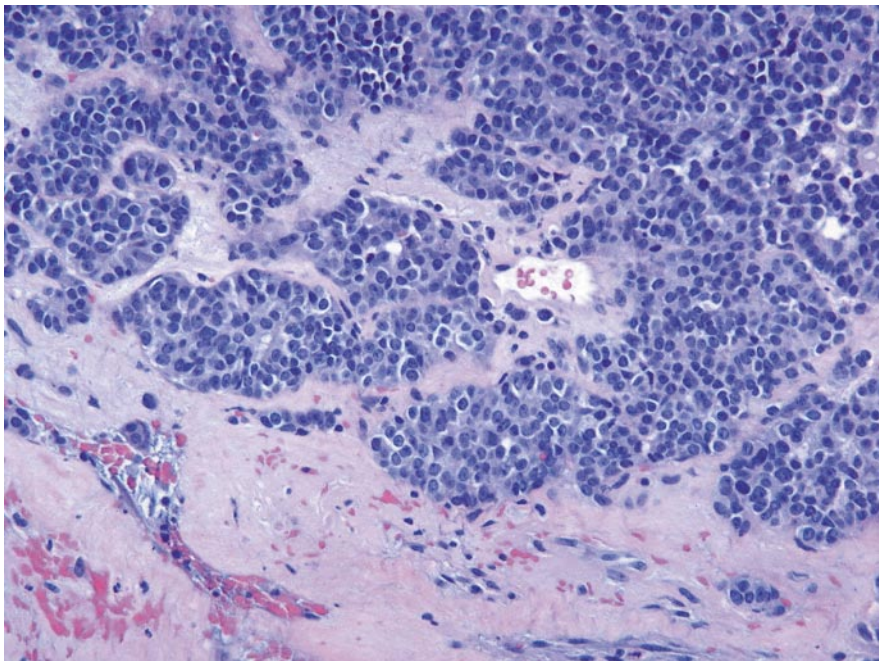


FIGURE 36–3. Photomicrograph of olfactory neuroblastoma showing small round tumor cells with fine nuclear chromatin growing in sheets. (Hematoxylin-eosin stain, ×20 original magnification.) Courtesy of Michael D. Feldman, MD.

Lymphoma makes up approximately 18% of all nonsquamous malignancies of the paranasal sinuses.[13] Most commonly, this consists of diffuse large B-cell lymphoma; however, natural killer/T-cell lymphoma is quite common.[15]

A myriad of neoplasms, both benign and malignant, may be found in the nose and paranasal sinuses. Even benign processes can have devastating consequences owing to local invasion or destruction in the presence of the numerous vital structures located in the area. Malignant tumors have a worse prognosis, but whether one histologic subtype is worse than another is subject to debate.

EPIDEMIOLOGY

The greatest obstacle in understanding nasal and paranasal sinus tumors is the rarity with which they occur. This rarity frustrates clinical studies exploring treatment options. Similarly, the low incidence, combined with their nonspecific symptoms, often leads to a critical delay in diagnosis. Commonly cited data indicate that these neoplasms account for 0.2 to 0.8% of all carcinomas and for 3% of those occurring in the upper aerodigestive tract.[1] Demographically, these tumors occur predominantly in the age range from 50 to 90 years. Their frequency in the young is low, being 0.1 to 0.3 per 100,000 in the first decade of life.[1] However, by the eighth decade, the frequency has risen to 7 per 100,000.[1] These tumors occur in all races, and there is no gender predilection.

The histologic differential diagnosis of paranasal sinus tumors is enormous. However, squamous cell carcinoma is by far the most frequently encountered malignancy, accounting for up to 80% of all neoplasms according to some studies.[1] This is followed by the salivary gland tumors (4 to 15%) and sarcomas (4 to 6%).[1] Other pathologic entities such as lymphomas, esthesioneuroblastomas, and melanomas are usually represented in most case series but rarely make up a substantial percentage in any one study.

Of the different sites, the maxillary sinus is commonly identified as the primary site, representing 55 to 80%.[16] It is followed by the nasal cavity (27 to 33%), the ethmoid sinuses (9 to 10%), and the frontal and sphenoid sinuses (1 to 2%).[6,17,18]

Environmental factors play a large role in the development of paranasal sinus malignancy. Up to 44% of these tumors are attributed to occupational exposures. The most well known is the association between wood dust and adenocarcinoma. Exposure to wood dust increases the relative risk of adenocarcinoma by 540 times.[10] However, it should be noted that squamous cell carcinoma also shares this relationship to wood dust, with prolonged exposure resulting in a 21 times increase in relative risk. Smoking has also been suggested to play a role in

nasal and paranasal sinus cancers. Additionally, numerous other industrial substances have been shown to be associated with the development of paranasal sinus malignancy (Table 36–2). [6,19]

Interestingly, there is also evidence that environmental pollution may be associated with an increasing incidence of paranasal sinus malignancy.[19] In a recent retrospective study, Calderon-Garciduenas et al found substantial increases in the rates of paranasal malignancy when comparing the years 1976 to 1986 and 1987 to 1998.[19] They postulated that this might be related to increasing exposure to complex chemical pollutants in their patient population during these time periods.

TABLE 36–2. Environmental Factors Associated with Nasal and Paranasal Sinus Malignancy

Wood dust
Nickel
Isopropyl oil
Organic fibers
Chromium
Volatile hydrocarbons

TABLE 36–3. Symptoms of Nasal and Paranasal Sinus Tumors

Early	*Late*
Nasal obstruction	Epistaxis
Rhinorrhea	Cranial nerve dysfunction
	Proptosis
	Facial pain
	Facial swelling
	Trismus

PATIENT EVALUATION

In contrast to many other areas in the head and neck, paranasal sinus tumors are not characterized by an early presentation. The typical delay in diagnosis is commonly cited as 8 months or more.[6,10] This is secondary to the fact that the early symptoms of paranasal sinus tumors are among the most common symptoms seen by primary care physicians. Complaints such as nasal obstruction, rhinorrhea, and epistaxis rarely receive thorough evaluations outside of offices of otorhinolaryngologists until they have become excessively prolonged or other symptoms develop. Late symptoms include epistaxis, ocular dysfunction secondary to extraocular muscle invasion or cranial nerve involvement, proptosis, facial pain, facial swelling, facial numbness, loosening of teeth, epiphoria, visual loss, anosmia, trismus, and even facial weakness (Table 36–3). The presence of any early symptom lasting over 4 weeks or any late symptom should prompt one to obtain further diagnostic studies.

Once the patient is suspected of having a paranasal sinus neoplasm, it is essential that a complete and organized evaluation be undertaken. This includes a thorough head and neck examination, radiologic evaluation, biopsy, and treatment planning.

PHYSICAL EXAMINATION

Although every patient warrants a complete and thorough examination, important areas the practitioner should focus on when examining a patient with a suspected paranasal sinus tumor include the orbital, oral, and neurologic examinations.

Proptosis, chemosis, or evidence of extraocular muscle impairment signifies orbital invasion and provides the practitioner with a readily available means of determining tumor extent. Evaluation by an ophthalmologist may be helpful for patients with subtle physical findings.

Similarly, the oral examination may help to delineate tumor spread. Key findings include mass effect in the palate and gingivobuccal sulcus. Trismus indicates invasion of the pterygoid muscles. The extent of tumor invasion is the primary focus, but other practical considerations for the surgeon include planning for operative intubation in the patient who cannot open his or her mouth.

The clinician must be sure to palpate the parotid glands and neck thoroughly in search of clinically positive lymph nodes. Such findings are prognostically significant, and their presence may significantly alter patient management.

A thorough neurologic examination must also be performed. Cranial nerve I dysfunction may be

related to either nasal obstruction or nerve involvement. A gross assessment of visual acuity and oculomotor function should be undertaken. If an abnormality is noted, a more thorough examination by an ophthalmologist is warranted for reasons of treatment planning. Finally, facial/palatal anesthesia should be noted to assess involvement of the trigeminal nerves.

Fiber-optic assessment of the nasal cavity may be performed with either a flexible or a rigid endoscope. This portion of the examination provides information on the origin of the tumor, where it extends, and its size. Occasionally, owing to the late presentation and advanced stage of disease, this portion of the examination is limited because of the difficulty in passing the endoscope posterior to the tumor. This is especially true of nasal cavity and ethmoid sinus tumors that tend to obstruct the nasal cavity. Decongestants and topical anesthesia are helpful in this endeavor.

Imaging

Various imaging options are available to the practitioner evaluating the patient with paranasal sinus cancer. However, the two most commonly employed today are computed tomography (CT) and magnetic resonance imaging (MRI). Both provide information useful in treatment planning. In practice, the majority of patients undergo both types of imaging prior to receiving treatment.

On CT, malignant tumors of the paranasal sinuses and nasal cavity appear as dense, homogeneous, destructive masses. The primary benefit of CT scans is the detailed demonstration of bony structure.[20] Finely cut coronal scans provide information regarding the integrity of both the orbital walls and skull base (Figure 36–4). Posterior extension of the tumor toward the pterygopalatine fossa, pterygomaxillary fissure, and infratemporal fossa may also be appreciated.[20] Although not diagnostic, CT scans may also provide evidence of advanced perineural invasion through the demonstration of enlarged neural foramina. The primary disadvantage of CT is the inability to distinguish between soft tissue and retained secretions. Another disadvantage is that although CT can give one an idea of the integrity of the orbital wall and skull base, this often does not correlate with orbital or dural invasion.

Magnetic resonance imaging is more sensitive in determining the true extent of tumor expansion. On T_1-weighted images, most tumors exhibit a low signal intensity that enhances after the administration of gadolinium. On T_2-weighted images, these tumors show a higher signal intensity than muscle but are not as bright as secretions[20] (Figure 36–5). The primary advantage of MRI is the ability to distinguish nasal secretions from tumor mass. Additionally, MRI allows accurate prediction of both dural involvement and venous sinus invasion. Specifically, the presence of pial enhancement, focal dural nodules, or dural thickening of more than 5

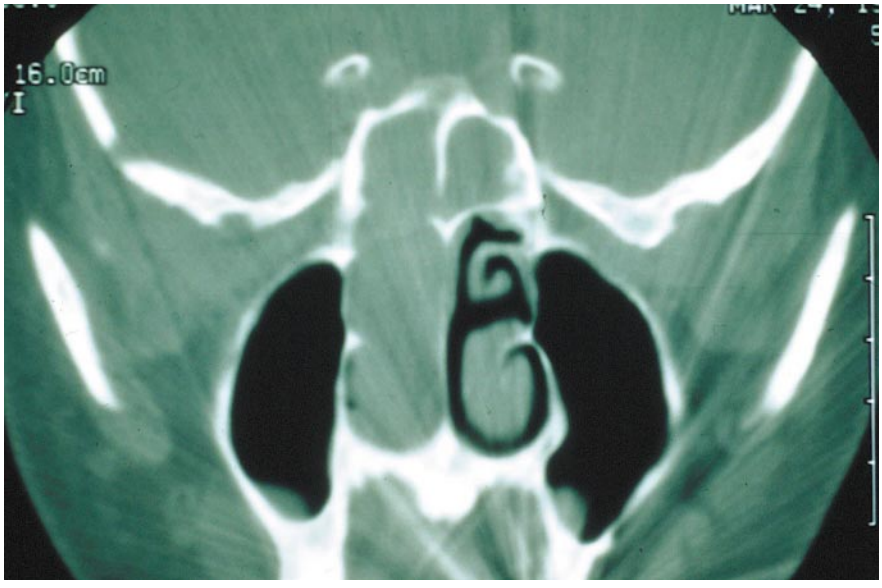


FIGURE 36–4. Coronal computed tomographic scan of the posterior part of the nasal cavity and the skull base showing tumor eroding the face of the right sphenoid sinus. One cannot differentiate retained mucus from tumor.

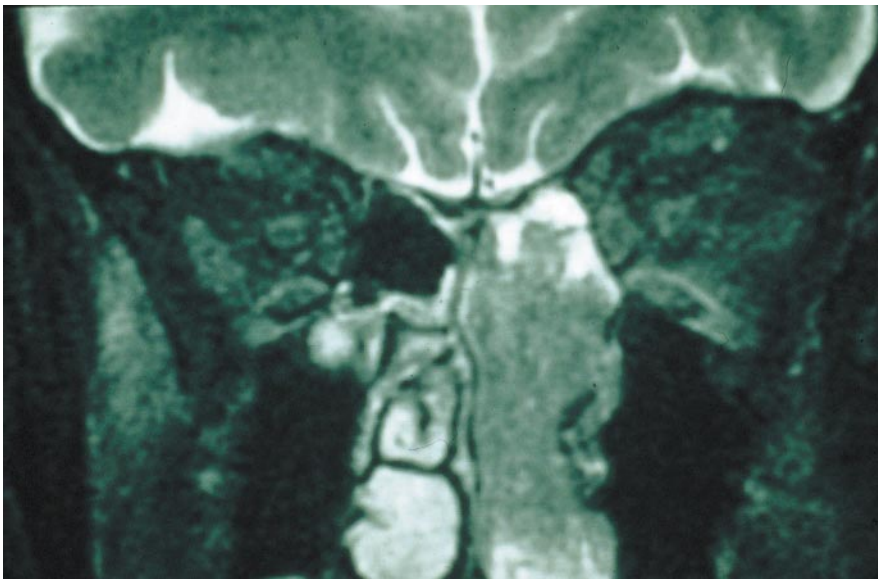


FIGURE 36–5. T_2-weighted coronal MRI showing a tumor in the left nasal cavity and ethmoid sinus with retained mucus (high signal intensity) above the tumor.

mm is highly sensitive and accurate in predicting the presence of dural invasion.[18] Disadvantages of MRI are its inability to provide information regarding the invasion of compact bone.[20] Additionally, like CT, MRI is not accurate in predicting perineural or orbital involvement.[18]

Angiography should be reserved specifically for patients in whom there is suspected internal carotid artery involvement. When carotid artery involvement is suspected, magnetic resonance angiography will often suffice to confirm or refute it.[21,22] Indications for conventional angiography are (1) preparation for embolization of a vascular lesion preoperatively, (2) to define intracranial arterial anatomy in a patient who is to undergo arterial bypass, and (3) balloon occlusion testing in a patient who will likely have the carotid artery sacrificed.

Most patients benefit from multiple imaging modalities during the workup of their paranasal neoplasm. Key issues, such as orbital involvement, skull base invasion, and vascular involvement, play significant roles in determining the proper treatment modality for the individual patient.

Staging

In an analysis of the various staging systems employed for maxillary cancer, Willatt et al listed four goals that an adequate staging system should meet.[23] They include (1) a balanced distribution of staging, (2) the ability to stage the majority of patients, (3) a correlation between T stage and treatment, and (4) a correlation between T stage and survival. Unfortunately, such a system has not been developed for paranasal sinus tumors. The lack of uniformity in staging is a major impediment to comparing treatment modalities from various studies. Since paranasal sinus tumors are so rare and no one clinic can hope to have a large number of patients to study, such comparisons are mandatory to form conclusions concerning the proper treatment of patients with paranasal sinus malignancy.

In the United States, the favored staging system is that proposed by the American Joint Committee on Cancer (Table 36–4). It is largely based on Ohngren's line.[23] This classification divides the facial

TABLE 36–4. American Joint Committee on Cancer Staging System for Malignant Neoplasms of the Maxillary Sinus

T1	Limited to antral mucosa without bony erosion
T2	Bony erosion of the infrastructure (hard palate, middle nasal meatus, or inferior nasal meatus)
T3	Invasion of cheek skin, posterior wall of sinus, or anterior ethmoid cells
T4	Invasion of orbital contents, cribriform plate, posterior ethmoid cells, sphenoid sinus, nasopharynx, soft palate, pterygopalatine fossa, infratemporal fossa, or skull base

skeleton into two parts by a theoretical plane drawn from the medial canthus of the eye to the angle of the mandible. Tumors of the maxillary sinus posterior to this plane have a poorer prognosis. However, this staging system fails to classify tumors that do not originate in the maxillary sinus. Obviously, this is a major limitation in discussing the wide variety of tumors that may occur in the nasal cavity and ethmoid sinuses.

One method to overcome this difficulty is to describe individually the site of involvement for each patient. This is cumbersome and presents difficulty in outcome analysis. As an alternative, many authors now use the staging system proposed by the International Union Against Cancer (UICC) (Table 36–5). Although the UICC staging system accounts for the most common sites of paranasal sinus malignancy, its variation from the AJCC staging of maxillary cancer makes analysis of studies using the different staging systems difficult. However, for the present time, the UICC staging system is the most complete system currently in use.

TABLE 36–5. International Union Against Cancer (UICC) Staging System for Malignant Neoplasms

Maxillary sinus	
T1	Tumor is limited to the antral mucosa with no erosion or destruction of bone
T2	Tumor causes bone erosion or destruction, except for the posterior antral wall, including extension into the hard palate and/or the middle nasal meatus
T3	Tumor invades any of the following: bone of the posterior wall of maxillary sinus, subcutaneous tissues, skin of cheek, floor or medial wall of orbit, infratemporal fossa, pterygoid plates, ethmoid sinuses
T4	Tumor invades orbital contents beyond the floor or medial wall including any of the following: the orbital apex, cribriform plate, base of skull, nasopharynx, sphenoid, frontal sinuses
Ethmoid sinus	
T1	Tumor is confined to the ethmoid with or without bone erosion
T2	Tumor extends into the nasal cavity
T3	Tumor extends to the anterior orbit and/or maxillary sinus
T4	Tumor with intracranial extension, orbital extension including apex, involving sphenoid and/or frontal sinus and/or skin of external nose

PROGNOSIS

Because of the previously discussed difficulty in comparing various clinical studies, it is difficult to state the extent that various clinical factors play in prognosis. Clearly, an advanced stage is associated with a worse prognosis. In fact, it has been asserted that the most accurate predictor of poor prognosis is the T category.[10,13,24] For example, although the overall total 5-year survival for patients with paranasal sinus malignancy may range from 40 to 50%, patients categorized as T1 and T2 may have a 5-year survival as high as 70%.[7–10,13,16] Conversely, patients with a tumor category of T3 or T4 commonly have 5-year survivals as low as 30%.[10] Specifically, tumors with anterior skull base involvement are known to have the worst prognosis.[10,18] Because of this, many authors have argued that a shorter delay in diagnosis would result in lower T categories at the time of presentation and improvement in overall survival.[25]

Lymph node involvement at the time of diagnosis is known to be associated with a uniformly dismal prognosis.[8,10,26] This manifestation of disease is rare in patients with paranasal sinus tumors and approximates 10% in most series.[7,8,16,25] It is because of the poor prognosis associated with lymph node disease that some clinicians now recommend for patients with advanced stages of paranasal sinus malignancy elective treatment of the N0 neck.

Interestingly, many studies have shown that histologic subtype does not play a role in the overall prognosis.[7,9,10,13,27] However, this is contradicted by other reports demonstrating improved survival in tumors such as adenoid cystic carcinoma and lymphoma and poor survival rates in undifferentiated carcinoma as well as melanoma.[8,9] Additionally, some investigators feel that tumor differentiation plays a role in prognosis, with well to moderately differentiated tumors having significantly better prognosis than poorly differentiated tumors.[8] The true extent that histology and differentiation play in patient prognosis remains to be clarified.

TREATMENT

A variety of treatment options are available to the clinician dealing with paranasal sinus tumors. As with tumors in other areas of the head and neck, these include surgery, radiation, and chemotherapy. Additionally, within each modality, a variety of combinations and techniques may be used. Regardless of the order of treatment, multimodality therapy is widely advocated.[16,24,26,27]

Surgery

There has been a long history of treating patients with nasal and paranasal sinus tumors with surgical resection. These attempts predate the modern era of computer imaging, fiber-optic endoscopy, and advanced surgical techniques. Advancement in technology has paved the way for advances in surgical care. Developments in imaging and the ability to perform an endoscopic examination allow one to assess preoperatively tumor size, location, and the integrity of surrounding structures.

As technological advances expand our ability to resect more advanced lesions, the surgeon is forced to address the question of what is resectable in the context of what is curable. Unfortunately, the answers to these two questions may not coincide. Proposed criteria for unresectable lesions include (1) transdural extension, (2) invasion of the prevertebral fascia, (3) bilateral optic nerve involvement, and (4) gross cavernous sinus invasion.

Operative technique in paranasal sinus surgery includes both the resection and the approach through which the resection takes place. The choice of resection is based on the location of tumor and may consist of a variety of procedures ranging from endoscopic removal of gross tumor to a total maxillectomy with orbital exenteration. The focus is to achieve total resection of the tumor since local recurrence is by far the most common cause of failure. En bloc resection is ideal, although this is rarely possible. A number of approaches are available depending on the size and location of the tumor and the excision to be performed. They include the endoscopic approach, midfacial degloving, lateral rhinotomy, lateral facial (Fisch-Mattox) approach, craniotomy, and combined craniofacial approaches.

Endoscopic Sinus Surgery The endoscopic approach to paranasal sinus malignancy is an option for treatment for low-grade or benign tumors involving the lateral nasal wall, ethmoid sinuses, or the sphenoid sinus. The primary advantage of this approach is the lack of need for facial incisions. Other advantages include the magnification that modern endoscopy provides.[17]

There has been a great deal of controversy surrounding the use of the endoscopic approach versus the more traditional transfacial approaches, even in treatment for benign lesions such as IP. Those advocating an endoscopic approach include Waitz and Wigand, who reported a recurrence rate of 17% using endoscopic medial maxillectomy for the resection of IP.[28] Their technique involved piecemeal removal of all gross tumor followed by histologically guided resection of tumor margins. Underlying bone was drilled to ensure total tumor removal. The authors noted that not only was their recurrence rate comparable to what they experienced for extranasal procedures, but the surgery was effective even in large lesions involving the posterior ethmoidal cells, nasofrontal duct, and sphenoid sinus. They concluded that many "mutilating" transfacial procedures could be avoided through the judicious use of nasal endoscopy. Sukenik and Casiano reported a 94% cure rate using only endoscopic resection of IP.[3] The efficacy of this treatment modality has been confirmed by Stankiewicz and Girgis, who demonstrated that the endoscopic approach could be effectively used in patients with recurrent disease as long as the papilloma was in an area that could be visualized with endoscopic equipment.[29]

Proponents of the transfacial approach cite the importance of en bloc resection in preventing tumor recurrence. The average rate of recurrence of IP using en bloc resection has been noted to be 8%, which is less than the 17 to 20% recurrence noted when using the intranasal approach.[30] Additionally, when IP recurs, there is the potential for multicentricity and malignancy. However, many investigators have noted that recurrent tumors are generally detected early in routine follow-up and can also be managed through an endoscopic approach.[31]

Tumor types other than IP have also been treated through the endoscopic approach. Good results have been obtained using this approach for hemangioma, esthesioneuroblastoma, adenoid cystic carcinoma, adenocarcinoma, and even squamous

cell carcinoma.[32] In fact, the most important factor in deciding on an endoscopic approach is not the type of tumor but rather the extent. Early tumors limited to the lateral nasal wall are good candidates for endoscopic resection (Figure 36–6). This may range from endoscopic medial maxillectomy to total ethmoidectomy with mucosal stripping and the drilling of potentially involved bone.[31] A contraindication to an endoscopic approach is invasion of inaccessible areas, including the lateral part of the maxillary sinus, periorbita, lacrimal sac, supraorbital ethmoid cells, frontal sinus, and skull base.[31] Certainly, the current standard dictates that the vast majority of patients with malignant lesions should be treated with an open en bloc resection. Additionally, even when one is dealing with a low-grade tumor, many authors caution that patients should be prepared preoperatively for the possibility of conversion to an open procedure.[29,32]

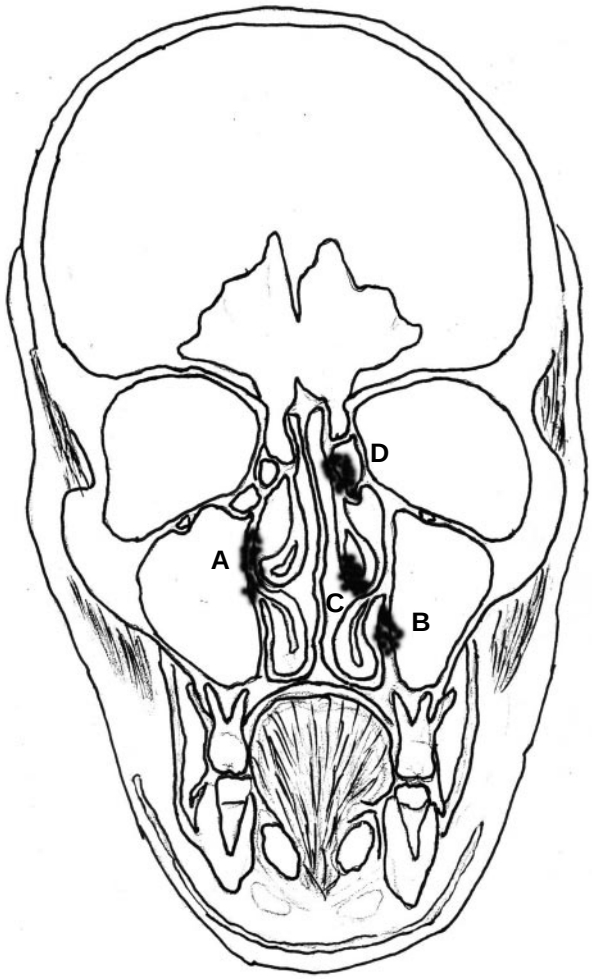


FIGURE 36–6. Diagram of intranasal tumor sites that can be considered for endoscopic intranasal resection: A, middle meatus; B, inferior meatus; C, middle turbinate; D, anterior ethmoid cells.

Most surgeons would agree that in the hands of a skilled endoscopist, endoscopic resection of early nasal and paranasal sinus tumors is appropriate. However, care must be taken to ensure that negative surgical margins are obtained as this has been shown to improve significantly patient prognosis.[31] Additionally, because of the risk of recurrence, frequent follow-up with serial endoscopic examination is strongly encouraged. Regardless of treatment approach, the role of serial endoscopy is established for follow-up of the patient's nasal or surgical cavity. Complications include cerebrospinal fluid (CSF) leaks, visual/orbital injury, synechia, and changes in the sense of smell.

Midfacial Degloving Approach Midfacial degloving represents one of many open approaches that one may use to approach an endonasal tumor. This approach consists of (1) bilateral intercartilaginous incisions, (2) a septocolumellar complete transfixion incision, (3) bilateral sublabial incisions, and (4) bilateral piriform aperture incisions.[33] This is a versatile approach that can provide access to a wide variety of areas, including the nasal cavity, nasopharynx, maxillary antrum, orbital floor, and zygoma (Figure 36–7). It may be combined with a craniotomy for tumors involving the anterior part of the skull base.[33] Various types of excisions can be performed through this approach such as medial maxillectomy, total maxillectomy with or without orbital exeneration, ethmoidectomy, sphenoidectomy, and resections of tumors invading the anterior part of the skull base when done in association with a craniotomy.

The primary advantages of the procedure are its lack of extensive facial incisions and good access provided for inferiorly based tumors. Additionally, there is a fairly low complication rate associated with the procedure.[33] Disadvantages include difficulty with access to superiorly based tumors. Other transfacial procedures are preferred for superiorly based tumors unless otherwise contraindicated.[6] Complications specific to midfacial degloving include lip numbness, vestibular stenosis, and oroantral fistula.[33]

Lateral Rhinotomy Lateral rhinotomy serves as the basis for the majority of transfacial procedures.

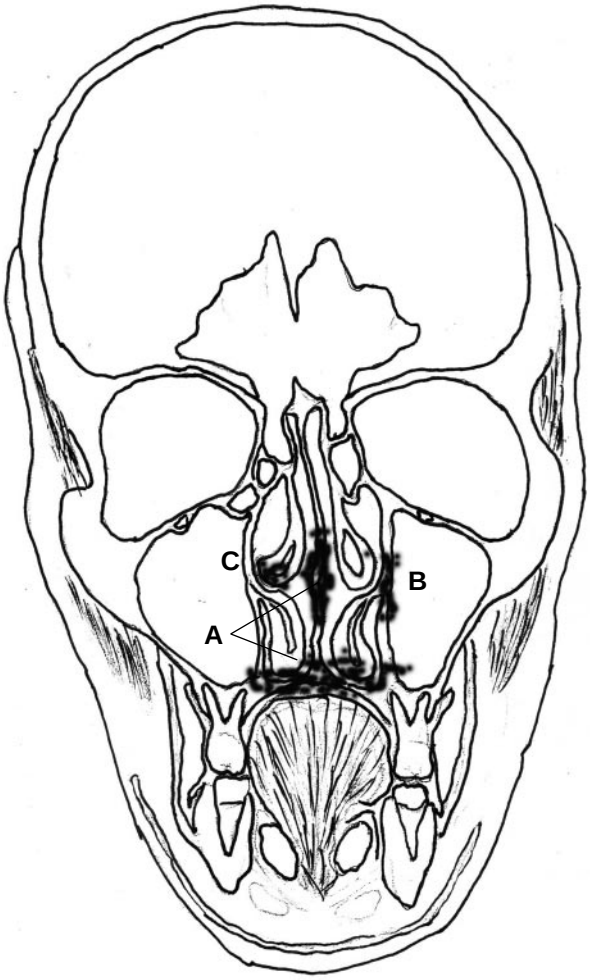


FIGURE 36–7. Diagram of tumor sites accessible through midfacial degloving: A, septum/nasal floor; B, lateral nasal wall, inferior meatus, middle meatus; C, middle turbinate; and nasopharynx (with or without Le Fort I osteotomy).

It consists of a curvilinear incision beginning at the inferomedial aspect of the brow and extending inferiorly midway between the nasal dorsum and medial canthus[34] (Figure 38–8). The distal portion of the incision should advance nasally from the nasofacial groove to allow for the optimal esthetic result. The incision can be extended through a variety of techniques, making it a versatile approach. Such extensions include a lip-splitting incision for retraction of the cheek (Figure 36–9), incision beneath the ipsilateral eyelid for further retraction (Weber-Ferguson incision), incision above the ipsilateral eyelid for orbital exenteration, or a superior nasal incision for access to the contralateral frontoethmoidal region.[35] The approach provides fairly wide access to tumors involving the supraorbital ethmoidal area, frontal duct, lacrimal fossa, orbit, or cribriform plate.[6] Through this approach, one may perform a variety of excisions including medial maxillectomy, total maxillectomy, radical maxillectomy, bilateral ethmoidectomy, and resection of most superiorly based tumors that are not directly invading the cribriform plate or fovea ethmoidalis[35] (Figure 36–10). It may be used in conjunction with a craniotomy for tumors that have invaded the skull base (craniofacial resection).

The primary advantage of the procedure is the wide access it provides combined with the relative ease in which the incisions may be modified to achieve better access. Disadvantages of the procedure include orbital complications including blepharitis, epiphora, and intermittent dacryocystitis related to transection of the nasolacrimal duct.[35] Proper attention must be paid to providing adequate reconstruction to this system at the time of surgery. This includes precise reconstruction of the medial canthal tendon in conjunction with dacryocystorhinostomy. In this setting, canaliculi stenting is not required. Another disadvantage of this approach is the presence of facial incisions. The cosmetic result is usually acceptable, but the degree of scarring, as well as the patient's feelings about it, can be difficult to predict.

Approach to the Frontal Sinus Primary involvement of the frontal sinus is rare, but the advanced nature of paranasal sinus tumors at presentation results in frequent invasion by tumors of nasal or ethmoid sinus origin. In these cases, the frontal sinus may be approached through a "gullwing" extension of the lateral rhinotomy inferior to the eyebrow in the upper eyelid sulcus. This is especially useful if the skin has been breached by tumor, in which case, one should include a wide margin of surrounding skin.[6] If the anterior table is not involved by tumor, one may use a bifrontal craniotomy through a coronal incision.[6,34] This may be used by itself for isolated frontal lesions or in combination with a lateral rhinotomy for lesions extending superiorly from the ethmoid sinuses or nasal cavity.[34]

The frontal sinus defect may be "cranialized," obliterated, or reconstructed. A specific complication associated with the management of frontal sinus tumors is the formation of a mucocele. This is

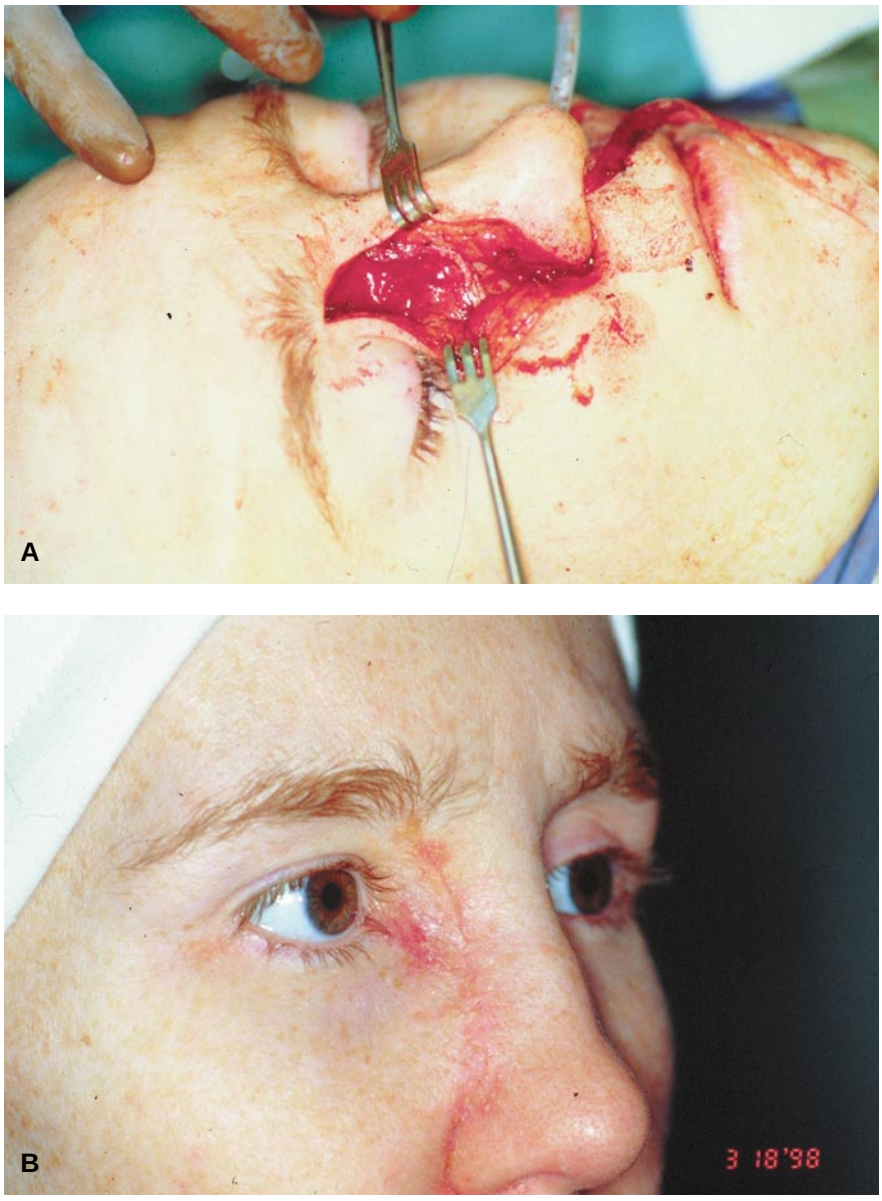
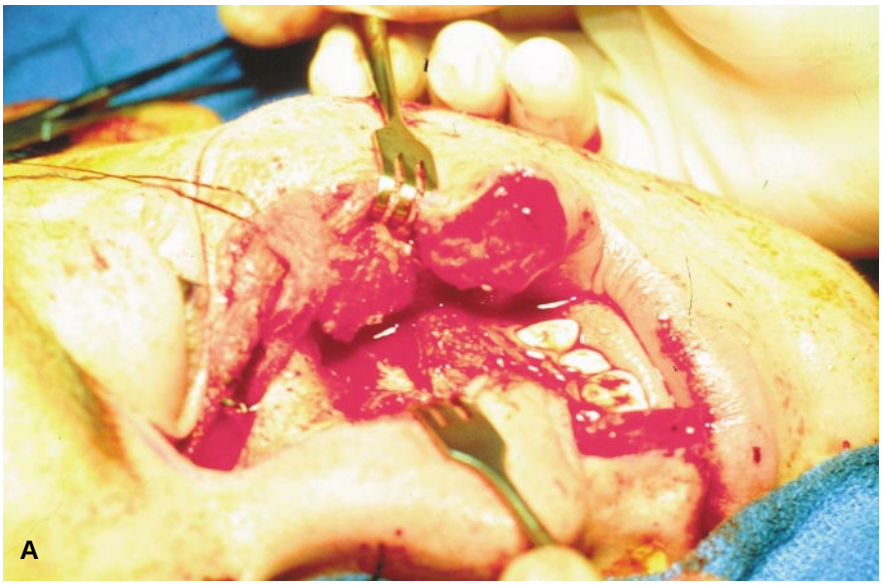


FIGURE 36–8. *A*, Lateral rhinotomy incision; intraoperative view of incision extending from the inferomedial eyebrow to the midpoint between the medial canthus and the nasal dorsum and to the junction of the alar base with the upper lip. *B*, Lateral rhinotomy incision appearance; early postoperative view. Note that the incision should ascend beyond the nasofacial groove in the middle third of the nose to the junction with the thinner nasal skin.

best avoided through careful removal of all frontal sinus and frontal recess mucosa followed by obliteration of the frontal recess with fascia and bone chips[34] in patients undergoing "cranialization" or obliteration. Reconstructed or preserved frontal sinuses will require radiographic follow-up. Additional complications include the poor cosmetic outcomes that may result when the anterior table must be resected. This defect is best reconstructed with a split calvarial bone graft.[34]

Anterior Lateral Skull Base Approach Although standard transfacial approaches offer good access to the sinonasal region, tumors that extend to the pterygopalatine fossa, sphenoid sinus, or nasopharynx are best approached through the lateral face/skull base[6] (Figure 36–11). The most well-described approach is the lateral facial split or the infratemporal fossa type C procedure of Fisch and Mattox.[6] The approach consists of a preauricular or modified Blair incision. The temporalis muscle is elevated and swung over the zygoma. Additionally, the zygoma may be removed, the pes anserinus of the facial nerve divided, retracted, or skeletonized, and the mandible disarticulated and advanced anteriorly. This provides exposure of the internal



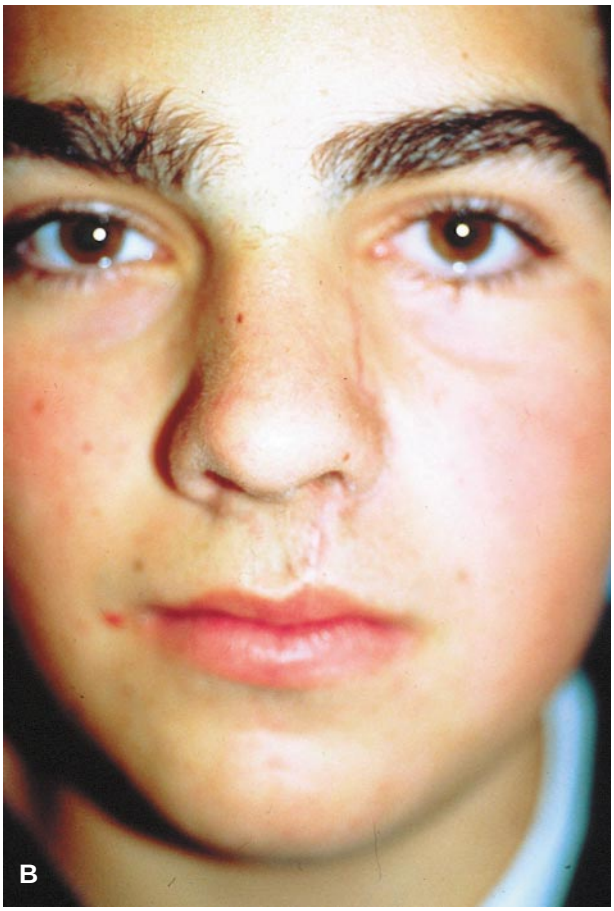


FIGURE 36–9. Intraoperative (*A*) and postoperative (*B*) views of lip-splitting extension of lateral rhinotomy incision. The extension of the incision to the lip can be along the philtral margin (preferred) or in the midline.

carotid artery from the bifurcation to the cavernous sinus.

The benefit of this approach is the access it provides to the central part of the skull base. Wide exposure of the internal carotid artery not only allows the opportunity to assess potential involvement with the tumor but also provides the option of bypass. In extreme cases, this exposure may necessitate division of the facial nerve or result in facial nerve weakness from traction on the nerve. In patients requiring facial nerve division, reconstructive neurorrhaphy is recommended.

Craniotomy/Craniofacial Approach An alternative approach to paranasal sinus tumors is the transcranial resection. The transfacial approaches described above are difficult and most often inappropriate to use in tumors adherent to the anterior part of the skull base. There is the added risk of CSF leak when attempting extracranial excisions in this area. Transcranial access is the approach of choice in tumors involving the cribriform plate or fovea ethmoidalis, especially when the tumor is of meningeal or cranial origin. Patients in whom there is known skull base involvement or intracranial invasion from the sinonasal cavity and possible adherence to or invasion of the dura require a combined craniofacial approach[36] (Figure 36–12). Most of these exposures will be with a transfacial (lateral rhinotomy) or midfacial degloving approach to guide the nasal and paranasal sinus components of

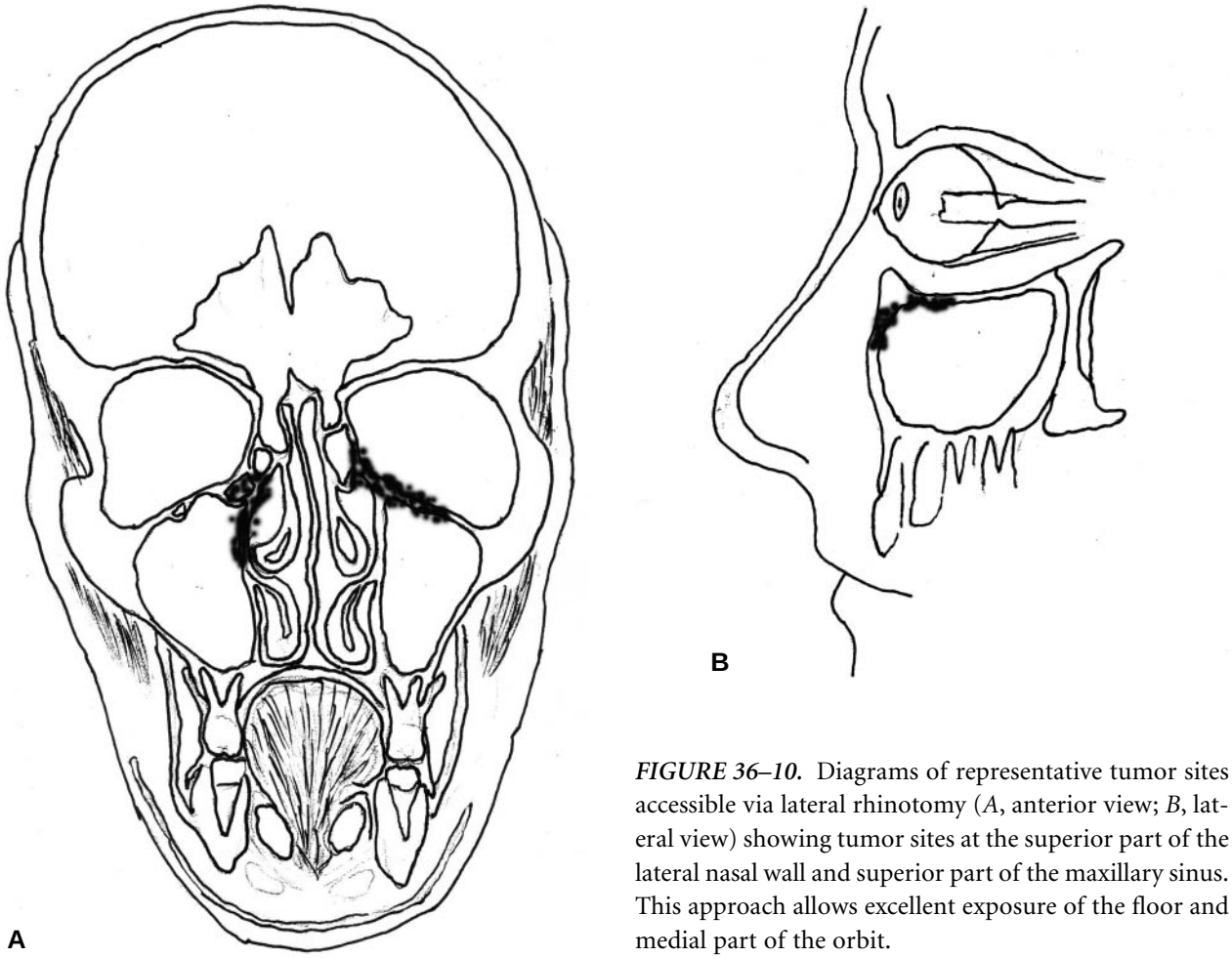


FIGURE 36–10. Diagrams of representative tumor sites accessible via lateral rhinotomy (*A*, anterior view; *B*, lateral view) showing tumor sites at the superior part of the lateral nasal wall and superior part of the maxillary sinus. This approach allows excellent exposure of the floor and medial part of the orbit.

the resection. Endoscopic guidance has added a minimally invasive component to guide the intranasal cuts for small tumors of the anterior part of the skull base. In the more uncommon scenario, the approach has the additional advantage of having no facial incisions. The intracranial exposure is recommended as the first step to ensure that the tumor is not unresectable owing to transdural or brain invasion.

The procedure involves a bicoronal incision with bifrontal craniotomy. An anteriorly based pericranial flap is created to provide an anatomic division between the dura and sinuses at the end of the procedure. An anterior osteotomy is placed above the supraorbital rims. The frontal lobes are gently retracted to provide access to the anterior cranial floor. The dura is then elevated except where it is directly involved in tumor and in the cribriform plate region. Small dural defects may be closed primarily, but larger defects that prohibit primary closure are grafted with cadaveric dura. At the end of the procedure, the pericranial flap is sutured to the dura, bone flaps are replaced, and the skin is closed. A lumbar drain, CSF-penetrating antibiotics, and high-dose corticosteroids are recommended in the perioperative period.[6]

In addition to the inherent morbidity of a craniotomy, a significant disadvantage of the approach is the limited access to the inferior and medial parts of the orbit as well as the maxillary sinus. For this reason, the procedure should be combined with a transfacial approach when there is involvement of the floor of the maxilla, the base of the nasal septum, or the soft tissues of the face or when one suspects that the excision will include orbital exenteration (Figure 36–13). The com-

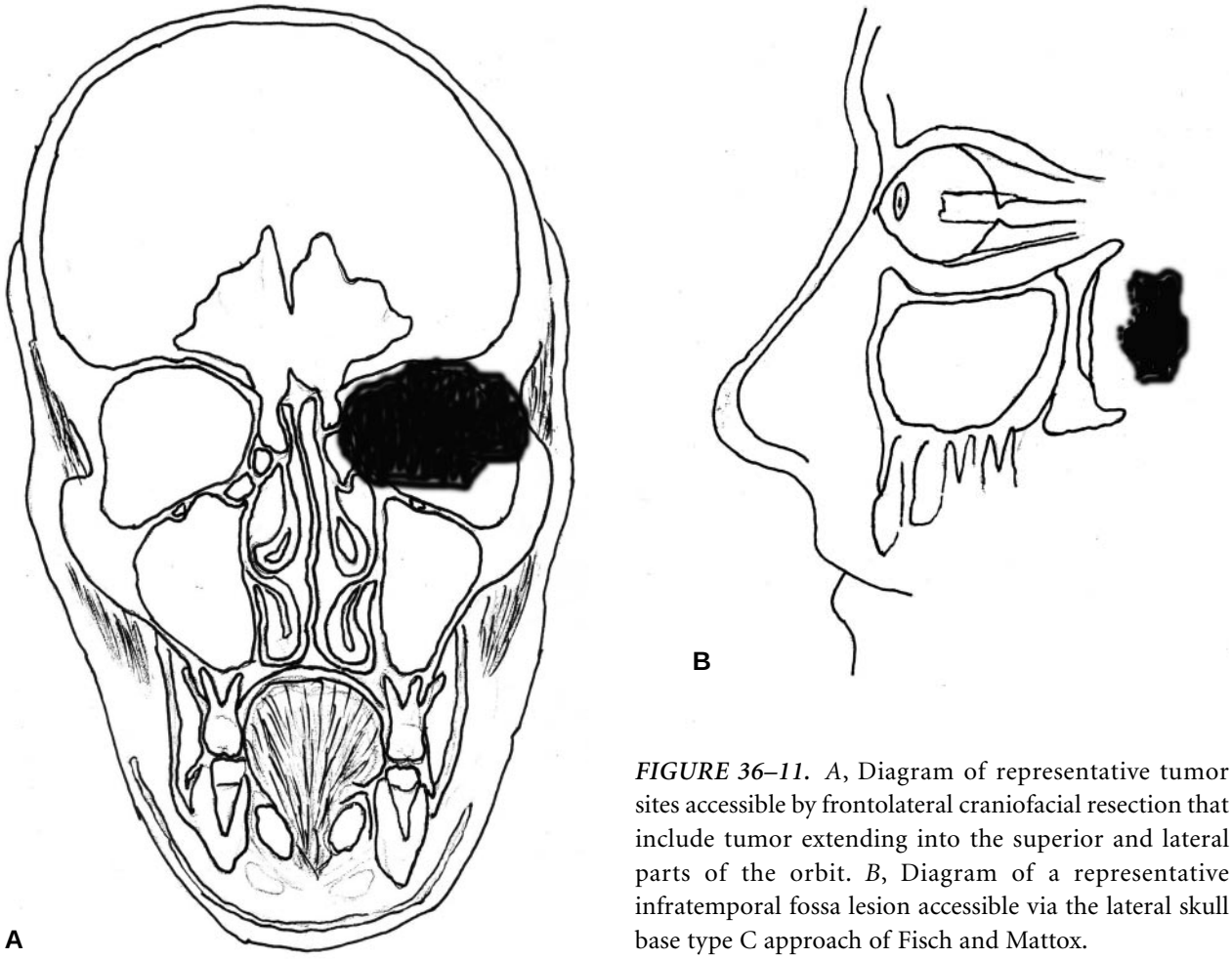


FIGURE 36–11. A, Diagram of representative tumor sites accessible by frontolateral craniofacial resection that include tumor extending into the superior and lateral parts of the orbit. *B*, Diagram of a representative infratemporal fossa lesion accessible via the lateral skull base type C approach of Fisch and Mattox.

bined craniofacial resection has allowed better access to the inferior and lateral nasal walls. This reduces operative morbidity by decreasing the amount of retraction needed on the frontal lobes and providing greater ease in obtaining adequate hemostasis.[36]

Specific complications related to the cranial or craniofacial approach include CSF leak, meningitis, and pneumocephalus. Our practice has been to use a tracheostomy to decrease the incidence of pneumocephalus. However, should it occur, it may be easily treated with aspiration. The other complications are rare when a pericranial flap has been used in conjunction with adequate perioperative antibiotics and a lumbar drain.[36]

The subfrontal craniofacial approach has been used for anterior skull base resections and offers the advantage of minimal brain retraction.[37] This technique may allow resection of one side of the cribriform plate with preservation of contralateral olfactory fibers. The craniofacial approach allows definite confirmation of tumor resectability from the dura/brain. If tumor is grossly through dura or into brain, these patients are at higher risk for carcinomatous meningitis. For our team, this has been a contraindication for craniofacial resection on patients with high-grade tumors. The lesions in these patients are typically not resected, and the patients are referred for multimodality therapy.

Radiation Therapy

Radiation therapy plays a large role in the treatment of patients with paranasal sinus tumors. Because of the late stage in which most of the patients present,

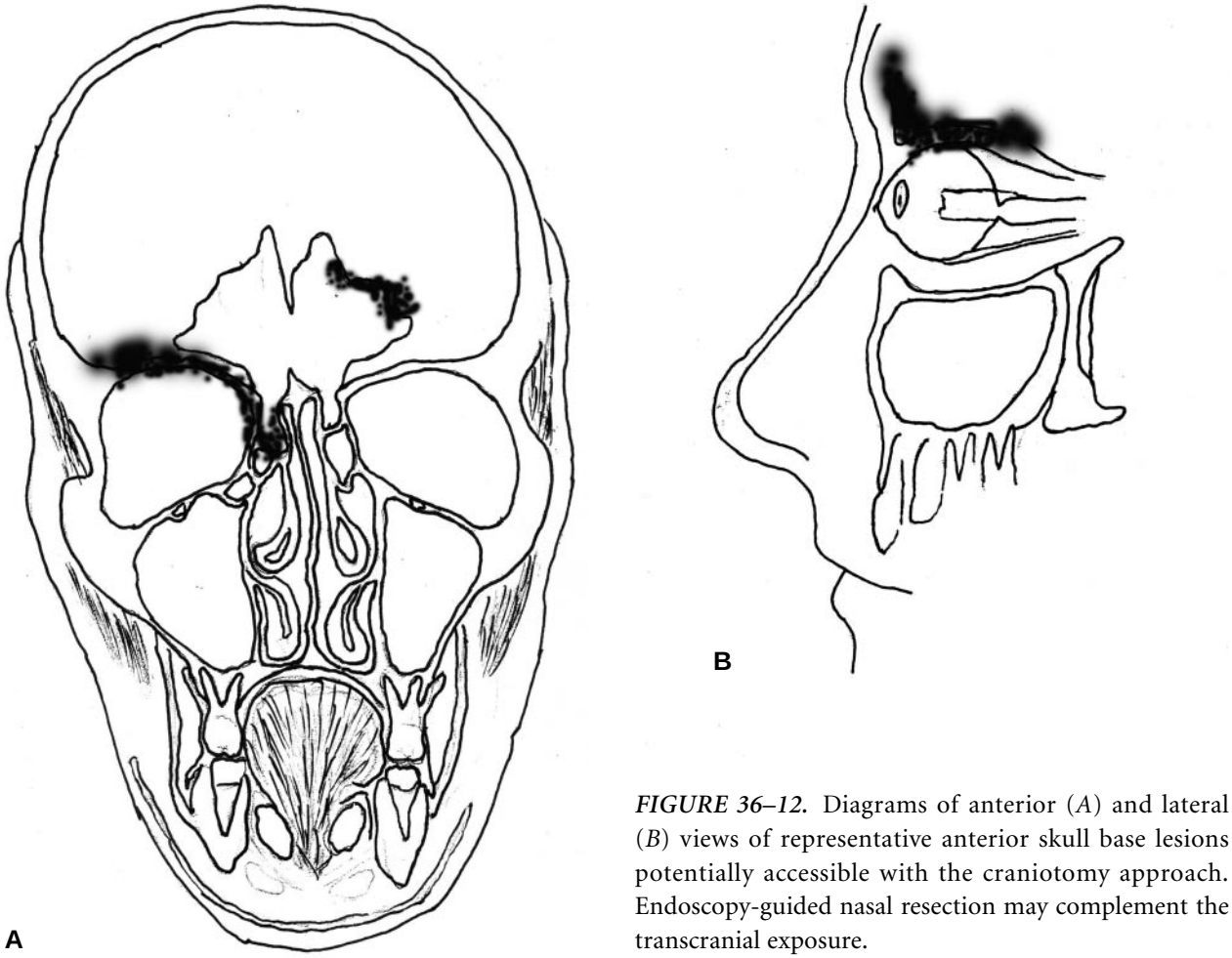


FIGURE 36–12. Diagrams of anterior (*A*) and lateral (*B*) views of representative anterior skull base lesions potentially accessible with the craniotomy approach. Endoscopy-guided nasal resection may complement the transcranial exposure.

multimodality therapy is nearly universal.[6,26,38] Additionally, in rare patients with early-stage cancers, radiation therapy alone has been shown to be as effective as surgical excision.[20,38,39] Some neoplasms such as lymphomas, plasmacytomas, and esthesioneuroblastomas appear to be especially radiosensitive.[6,10,38]

There is no clear evidence supporting preoperative radiation versus postoperative radiation.[6,27] Each has its distinct advantages and disadvantages. Postoperative radiation currently appears to be the most popular choice for several reasons. The dose of radiation that may be given is lower when using radiation to control microscopic residual disease, resulting in less morbidity.[38] Postoperative doses may range between 60 and 70 Gy. Patients who are not operative candidates may receive doses as high as 74 to 79 Gy.[40] Another reason for postoperative irradiation is the difficulty in determining the true borders of the radiated tumor and the potential for increased surgical morbidity in the irradiated field.[6] The pathologic results from the surgical specimen may be used to guide the radiation treatment. This is extremely important when dealing with areas such as the orbit, where intraoperative examination and subsequent pathology may be the only way to determine definitively the presence of tumor and help define required doses yet minimize morbidity.[38]

Notwithstanding, there are persuasive arguments for the use of preoperative radiation therapy as well. In common situations in which there is involvement of vital structures (orbit, dura, etc), preoperative irradiation can be used to "sterilize" these areas. This strategy is frequently used to preserve the eye when radiographic evaluation demonstrates that the structure is at risk. Debate continues as to whether this serves to preserve

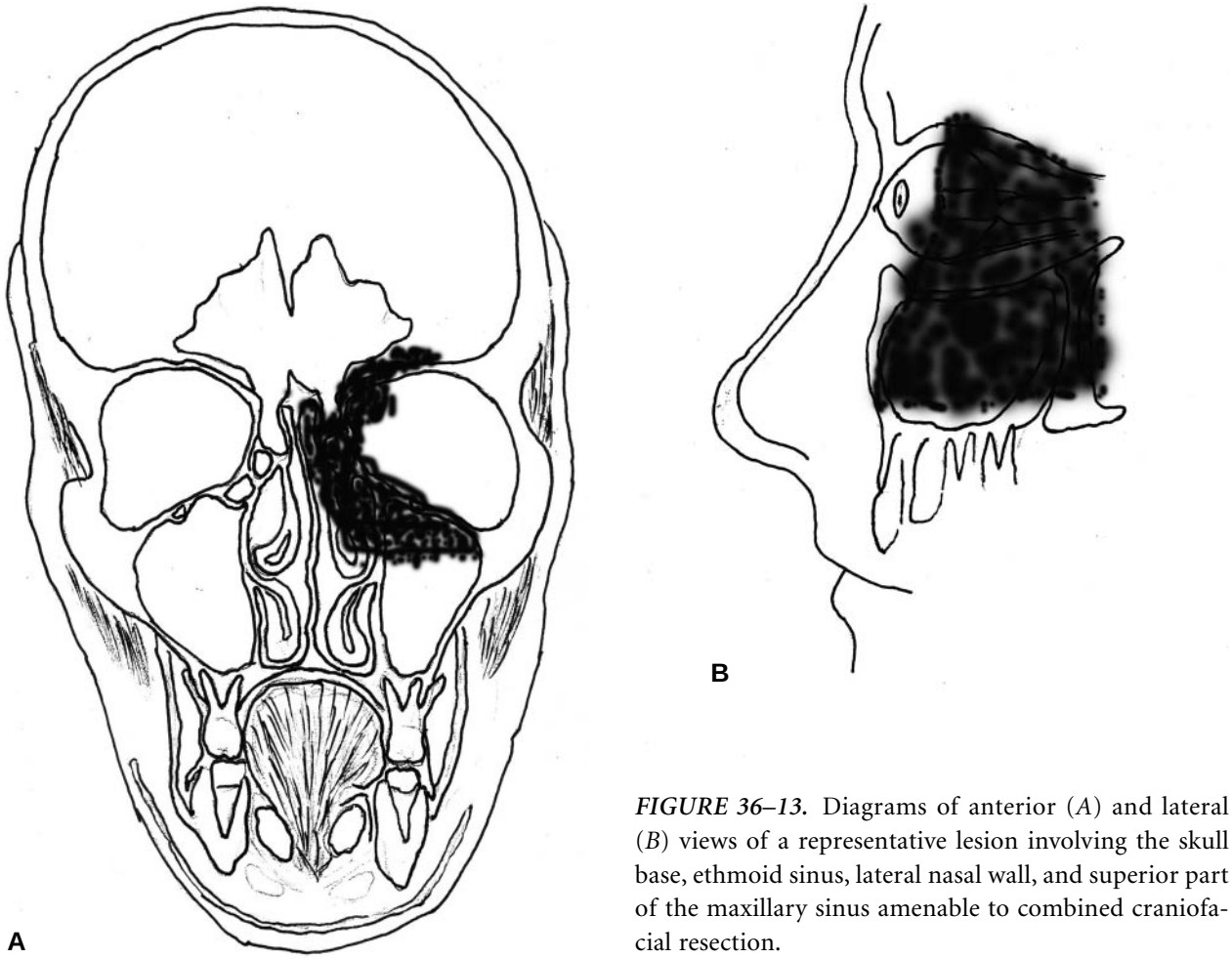


FIGURE 36–13. Diagrams of anterior (*A*) and lateral (*B*) views of a representative lesion involving the skull base, ethmoid sinus, lateral nasal wall, and superior part of the maxillary sinus amenable to combined craniofacial resection.

anatomic structures and improve survival. Some argue that preradiation surgery devascularizes the tumor bed, making the residual tumor less radiosensitive secondary to hypoxia. When there are grossly positive margins after surgery, higher doses of radiation are required to eradicate the disease.[38] It has been shown by Polin et al that presurgical response to radiation therapy is predictive of patient prognosis.[14] In their study of patients with esthesioneuroblastomas, patients with a favorable response to radiation demonstrated a significantly higher rate of disease-free survival.

As with surgery, there are numerous risks to radiation treatment. The majority relate to damage to the orbits and brain. Reported complications include blindness, central nervous system sequelae, otitis media, nasolacrimal duct obstruction, sinusitis, nasal bone destruction, chronic orbital pain, osteoradionecrosis, retinopathy, and medial canthus fistula.[16,40,41] The rates of these complications increase as higher doses of radiation are used. For instance, the retina and optic nerve exhibit a tolerance of approximately 50 Gy. Parsons et al found the rate of unilateral blindness in these patients to be as high as 33%.[39] Bilateral blindness occurred in 10% of the patients. Structures surrounding the orbit are also important to consider. The risk of injury to the lacrimal gland with resulting corneal ulceration, opacification, or vascularization increases substantially with doses greater than 30 Gy.[39] Neurologic sequelae are less common (10%). The rates of these complications have diminished markedly with the use of CT for treatment planning.[42] Computed tomographic scans allow for precise localization of both the tumor and surrounding critical structures. This localization provides for

treatment planning that minimizes radiation given to surrounding structures.

Chemotherapy

Several trials have evaluated the effects of chemotherapy in treating patients with paranasal sinus tumors. Although results have been promising, there is no definitive evidence that chemotherapy improves survival. However, several studies have shown evidence of tumor regression with the addition of chemotherapy.[38,43] For instance, Brasnu and colleagues have used cisplatin as a neoadjuvant chemotherapy prior to surgical resection in the treatment of adenocarcinoma reaching the skull base.[41] In this setting, the investigators achieved a 22.7% complete clinical response and a 13.6% complete histologic response. As with radiation therapy, a response to chemotherapy was found to be predictive of a better prognosis.[14,41] The morbidity of this treatment was minimal and was not increased by the addition of postoperative radiation therapy. Platinum-based chemotherapy has also been shown to be effective against esthesioneuroblastomas and other paranasal sinus tumors as well.[14,38,43] Alternatively, Sakai et al noted 5-fluorouracil (5-FU) in combination with radiation and surgical curettage to be relatively effective with minimal morbidity, although Tsujii and Tsuji found no difference in patients whether they were treated with 5-FU or not.[42,44] Although chemotherapy has not shown an improvement in overall survival, it must be remembered that it is difficult to prove an improvement in survival with *any* therapy for this disease. Further research is required and should be encouraged.

CONTROVERSIAL TOPICS IN PARANASAL SINUS TUMORS

Obviously, there is still much to be learned concerning the optimal management of paranasal sinus malignancy. However, two topics in particular warrant consideration.

Orbital Preservation

In the past, radical excision with orbital exenteration was performed if the tumor approached the orbit.[45] With the emphasis on multimodality therapy, there is a trend toward using preoperative radiotherapy and/or chemotherapy with the goal of preserving structures such as the orbit. Proponents of orbital preservation argue that the orbit is surrounded by a layer of periosteum that is resistant to tumor invasion.[46,47] This argument has been furthered by clinicoanatomic studies demonstrating the presence of a distinct fascial layer separating orbital fat from periorbita as well as an infrequent incidence of tumor invasion of orbital contents.[45,48,49] Proponents argue that for this reason, there is rarely a reason to remove the orbit at the time of surgical resection.

Surgeons favoring radical excision of the orbit point out that there is no reliable radiographic means to determine the extent of orbital involvement. Additionally, with the use of neoadjuvant radiation treatment, it can be difficult to determine the extent of orbital involvement at the time of surgical excision. The function of the eye must also be considered when making the decision to pursue preservation. Stern et al have shown that ocular function tends to be poor when the orbital floor is removed during excision.[50] There is also the morbidity of radiation therapy to consider. Postradiation complications involving the orbit include keratitis, cataracts, retinopathy, glaucoma, and chronic pain. As has been previously discussed, the rates of these complications can be quite high.

At the present time, the decision to preserve the eye is made on an individual basis. Clinical studies demonstrate that as long as the periorbita is not involved, it is oncologically sound to leave the orbit intact. However, attention must be paid to the chances of obtaining useful function from the preserved organ. This analysis should be based on both the amount of resected orbital floor as well as the final amount of radiation the orbit will receive. As Stern et al pointed out, "Strong consideration should be given to orbital exenteration at the time of surgery when the orbital floor is resected—especially if postoperative radiation fields will include the eye."[50]

Elective Treatment of the Neck

Traditionally, elective treatment of the N0 neck has not been advocated secondary to the low rate of occult lymph nodal disease.[6,40] Additionally, many argue that the primary lymphatic watershed of the sinonasal region is the retropharynx. Therefore, the

areas most likely to contain regional cancer reside in this region and not in the neck. This dogma has recently been challenged based on the following reasoning: the rate of lymph nodal recurrence ranges from 12 to 28%.[27] Patients with neck relapse are known to have inferior long-term survival and a higher risk of distant metastases compared with those who did not have locoregional failure. This work has been corroborated by Paulino et al, who also found that patients with T3 and T4 tumors have a higher rate of nodal recurrence and those patients with nodal recurrence had poorer survival rates.[51] In addition to this, Le et al demonstrated that elective neck irradiation effectively prevented nodal recurrence in the 25 patients in whom it was used.[52] For this reason, it is now considered reasonable to irradiate the N0 neck in the presence of a primary tumor of advanced stage (T3 and T4).

CONCLUSION

What we know concerning nasal and paranasal sinus tumors stands in stark contrast to the questions that need to be answered. Small cohorts and isolated case presentations contribute relatively little toward global understanding. As consensus is reached regarding staging, cross-study comparisons will be facilitated. This will undoubtedly lead to a greater understanding of these tumors in terms of pathogenesis, prognosis, and optimal treatment options.

REFERENCES

1. Osguthorpe JD. Preface. In: Osguthorpe JD, editor. The Otolaryngologic Clinics of North America—paranasal sinus tumors. Philadelphia: WB Saunders; 1995. p. xi–ii.
2. Mortignoni G, Huvos AG. Assessment of patients with suspected sinonasal neoplasm. In: Kraus DH, Levine HL, editors. Nasal neoplasia. New York: Thieme; 1997. p. 21–42.
3. Sukenik MA, Casiano R. Endoscopic medial maxillectomy for inverted papillomas of the paranasal sinuses: value of the intraoperative endoscopic examination. Laryngoscope 2000;110:39–42.
4. Blaugrund S, Parisier SC. Benign tumors of the nasal sinuses. Otolaryngol Clin North Am 1971;4: 143–57.
5. Goepfert H, Luna M, Lindberg RD, White AK. Malignant salivary gland tumors of the paranasal sinuses and nasal cavity. Arch Otolaryngol 1983; 109:662–8.
6. Osguthorpe JD. Sinus neoplasia. Arch Otolaryngol Head Neck Surg 1994;120:19–25.
7. Haraguchi H, Ebihara S, Saikawa M, et al. Malignant tumors of the nasal cavity: review of a 60-case series. Jpn J Clin Oncol 1995;25:188–94.
8. Harbo G, Grau C, Bundgaard T, et al. Cancer of the nasal cavity and paranasal sinuses. A clinico-pathological study of 277 patients. Acta Oncol 1997;36:45–50.
9. Svane-Knudsen V, Jorgensen KE, Hansen O, et al. Cancer of the nasal cavity and paranasal sinuses: a series of 115 patients. Rhinology 1998;36:12–4.
10. Alvarez I, Suarez C, Rodrigo JP, et al. Prognostic factors in paranasal sinus cancer. Am J Otolaryngol 1995;16:109–14.
11. Naficy S, Disher MJ, Esclamado RM. Adenoid cystic carcinoma of the paranasal sinuses. Am J Rhinol 1999;13:311–4.
12. Brandwein MS, Rothstein A, Lawson W, et al. Sinonasal melanoma. A clinicopathologic study of 25 cases and literature meta-analysis. Arch Otolaryngol Head Neck Surg 1997;123:290–6.
13. Kraus DH, Sterman BM, Levine HL, et al. Factors influencing survival in ethmoid sinus cancer. Arch Otolaryngol Head Neck Surg 1992;118:367–72.
14. Polin RS, Sheehan JP, Chenelle AG, et al. The role of preoperative adjuvant treatment in the management of esthesioneuroblastoma: the University of Virginia experience. Neurosurgery 1998;42:1029–37.
15. Cuadra-Garcia I, Proulxx GM, Wu CL, et al. Sinonasal lymphoma: a clinicopathologic analysis of 58 cases from the Massachusetts General Hospital. Am J Surg Pathol 1999;23:1356–69.
16. Paulino AC, Marks JE, Bricker P, et al. Results of treatment of patients with maxillary sinus carcinoma. Cancer 1998;83:457–65.
17. Sievers KW, Greess H, Baum U, et al. Paranasal sinuses and nasopharynx CT and MRI. Eur J Radiol 2000;33:185–202.
18. Eisen MD, Yousem DM, Montone KT, et al. Use of preoperative MR to predict dural, perineural, and venous sinus invasion of skull base tumors. AJNR Am J Neuroradiol 1996;17:1937–45.
19. Calderon-Garciduenas L, Delgado R, Calderon-Garciduenas A, et al. Malignant neoplasms of the nasal cavity and paranasal sinuses: a series of 256 patients in Mexico City and Monterrey. Is air pollution the missing link? Otolaryngol Head Neck Surg 2000;122:499–508.

20. Giri SP, Reddy EK, Gemer LS, et al. Management of advanced squamous cell carcinomas of the maxillary sinus. Cancer 1992;69:657–61.
21. Lund VL, Howard DJ. Tumors of the paranasal sinuses. In: Gershwin ME, Incaudo GA, editors. Diseases of the sinuses. Totowa (NY): Humana Press; 1996. p. 291–310.
22. Mosesson RE, Som PM. The radiographic evaluation of sinonasal tumors: an overview. Otolaryngol Clin North Am 1995;28:1097–115.
23. Willatt DJ, Morton RP, McCormick MS, Stell PM. Staging of maxillary cancer. Which classification? Ann Otol Rhinol Laryngol 1987;96:137–41.
24. Curran AJ, Gullane PJ, Waldron J, et al. Surgical salvage after failed radiation for paranasal sinus malignancy. Laryngoscope 1998;108:1618–22.
25. Tufano RP, Mokadam NA, Montone KT, et al. Malignant tumors of the nose and paranasal sinuses: Hospital of the University of Pennsylvania experience 1990-1997. Am J Rhinol 1999;13:117–23.
26. MacNab TI, Flores AD, Anderson DW. Treatment of paranasal sinus malignancy: the BCCA experience. J Otolaryngol 1992;21:244–8.
27. Flores AD, Anderson SW, Doyle PJ, et al. Paranasal sinus malignancy—a retrospective analysis of treatment methods. J Otolaryngol 1984;13:141–6.
28. Waitz G, Wigand ME. Results of endoscopic sinus surgery for the treatment of inverted papillomas. Laryngoscope 1992;102:917–22.
29. Stankiewicz JA, Girgis SJ. Endoscopic surgical treatment of nasal and paranasal sinus inverted papilloma. Otolaryngol Head Neck Surg 1993;109:988–95.
30. Myers EN, Petruzzelli GJ. Endoscopic sinus survey for inverted papilloma. Laryngoscope 1993;103: 711.
31. Rice DH. Endonasal surgery for nasal wall tumors. Otolaryngol Clin North Am 1995;28:1117–25.
32. Levine HL, Timen SM, Papsidero MJ. Endoscopic management of nasal and sinus neoplasms. In: Kraus DH, Levine HL, editors. Nasal neoplasia. New York: Thieme; 1997. p. 21–42.
33. Maniglia AJ, Phillips DA. Midfacial degloving for the management of nasal sinus and skull base neoplasms. Otolaryngol Clin North Am 1995;28: 1127–43.
34. Catalano PJ, Sen C. Management of anterior ethmoid and frontal sinus tumors. Otolaryngol Clin North Am 1995;28:1157–74.
35. Weisman R. Lateral rhinotomy and medial maxillectomy. Otolaryngol Clin North Am 1995;28: 1145–56.
36. McCutcheon IE, Blacklock JB, Weber RS, et al. Anterior transcranial (craniofacial) resection of tumors of the paranasal sinuses: surgical technique and results. Neurosurgery 1996;38:471–9.
37. Marenette L, Moore CE. Subcranial approach to tumors of the anterior cranial base. Otolaryngol Head Neck Surg 2000;122:466–7.
38. Raben A, Pfister D, Harrison LB. Radiation therapy and chemotherapy in the management of cancers of the nasal cavity and paranasal sinuses. In: Kraus DH, Levine HL, editors. Nasal neoplasia. New York: Thieme; 1997.
39. Parsons JT, Bova FJ, Fitzgerald CR, et al. Severe dry eye syndrome following external beam irradiation. Int J Radiat Oncol Biol Phys 1994;30:775–80.
40. Parsons JT, Kimsey FC, Mendenhall WM, et al. Radiation therapy for sinus malignancies. Otolaryngol Clin North Am 1995;28:1259–68.
41. Brasnu D, Laccourreye O, Bassot V, et al. Cisplatin-based neoadjuvant chemotherapy and combined resection for ethmoid sinus adenocarcinoma reaching and/or invading the skull base. Arch Otolaryngol Head Neck Surg 1996;122:765–8.
42. Tsujii H, Tsuji H. [Several factors for improving results of radiation therapy]. Gan No Rinsho 1990;36:2393–7.
43. McElroy EA, Buckner JC, Lewis JE. Chemotherapy for advanced esthesioneuroblastoma: the Mayo Clinic experience. Neurosurgery 1998;42:1023–7.
44. Sakai S, Hohki A, Fuchihata H, Tanaka Y. Multidisciplinary treatment of maxillary sinus carcinoma. Cancer 1983;52:1360–4.
45. Perry C, Levine PA, Williamson BR, Cantrell RW. Preservation of the eye in paranasal sinus cancer surgery. Arch Otolaryngol Head Neck Surg 1988;114: 632–4.
46. McCary WS, Levine PA. Management of the eye in the treatment of sinonasal cancers. Otolaryngol Clin North Am 1995;28:1231–8.
47. Stern SJ, Callender DL. Management of the orbit. In: Kraus DH, Levine HL, editors. Nasal neoplasia. New York: Thieme; 1997.
48. Larson DL, Christ JE, Jesse RH. Preservation of the orbital contents in cancer of the maxillary sinus. Arch Otolaryngol 1982;108:370–2.

49. Tiwari R, van der Wal J, van der Wal I, Snow G. Studies of the anatomy and pathology of the orbit in carcinoma of the maxillary sinus and their impact on preservation of the eye in maxillectomy. Head Neck 1998;20:193–6.
50. Stern SJ, Goepfert H, Clayman G, et al. Orbital preservation in maxillectomy. Otolaryngol Head Neck Surg 1993;109:111–5.
51. Paulino AC, Fisher SG, Marks J. Is prophylactic neck irradiation indicated in patients with squamous cell carcinoma of the maxillary sinus? Int J Radiat Oncol Biol Phys 1997;39:283–9.
52. Le Q, Fu KK, Kaplan MJ, et al. Lymph node metastasis in maxillary sinus carcinoma. Int J Radiat Oncol Biol Phys 2000;46:541–9.

CHAPTER 37

Reconstruction of the Outstanding Ear

Daniel G. Becker, MD, Stephen Y. Lai, MD, PhD

The normal auricle has a well-recognized configuration of depressions and projections. Although there are many variations, significant deviations from "normal" are immediately evident. In particular, prominent ears are readily apparent and are a relatively frequent cause of patient concern.

Correction of the outstanding ear requires a careful understanding of the discrete elements that comprise the normal ear. Careful anatomic analysis to determine the precise cause allows appropriate preoperative planning for the correction of a protruding ear. In selecting appropriate techniques from the many that are described in the literature, the surgeon must place greatest emphasis on reliably achieving a natural-looking improvement. In general, techniques that reposition rather than resect cartilage may be safer and are therefore preferred.

EMBRYOLOGY

Development of the auricle or pinna is initially visible in the 39-day-old embryo.[1] The auricle is formed from six mesenchymal proliferations (hillocks of His) at the dorsal ends of the first and second branchial arches surrounding the first branchial groove. These elements begin in the lower neck and ascend to the level of the eyes at the side of the head during gestation. Cartilage formation is visible in the seventh week, and hillock fusion occurs in the twelfth week of gestation. The recognizable shape of the auricle is visible by the twentieth week of gestation. Although each of the six hillocks was previously correlated to a specific element of the auricle, current studies suggest that the first branchial arch elements contribute to the tragus and the second branchial arch elements form the remaining structures.[2–4]

Despite much speculation, the functional role of the auricle in humans remains unclear. Unlike the pinnae of other species, the auricles of humans and primates have poorly developed external musculature and mobility. As an immobile structure, the auricle appears to facilitate localization of high-frequency sounds.[1] In other species, the auricle can be manipulated to determine the direction of sound and to protect the internal auditory structures. Voluntary apposition of the tragus and antitragus protects the ear from water and insect entry.

ANATOMY

Surface Features

The vertical axis of the ear along its longest dimension is rotated approximately 20 degrees from the vertical axis of the skull. The width should be approximately 55% of the length. "Idealized" dimensions are 63.5 mm by 35.3 mm for males and 59.0 mm by 32.5 mm for females.[3] The auricle is 85% of adult size by 3 years of age and is 90 to 95% of full size by 5 to 6 years of age, although it may elongate an additional 1 to 1.5 cm during life.[5] The helical rim along its lateral edge is approximately 1 to 2 cm from the mastoid skin. The angle of protrusion of the auricle or the auriculomastoid angle is usually between 15 and 30 degrees.

Among important surface features is the helix, the prominent rim of the auricle (Figure 37–1). Parallel and anterior to the helix is another prominence known as the antihelix or antihelical fold. Superiorly, the antihelix divides into a superior and an inferior crus, which surround the fossa triangularis. The depression between the helix and antihelix is known as the scapha or scaphoid fossa. The antihelical fold surrounds the concha, a deep cavity posterior to the external auditory meatus. The crus helicis, which represents the beginning of the helix, divides the concha into a superior portion, the cymba conchae, and an inferior portion, the cavum conchae. The cavity

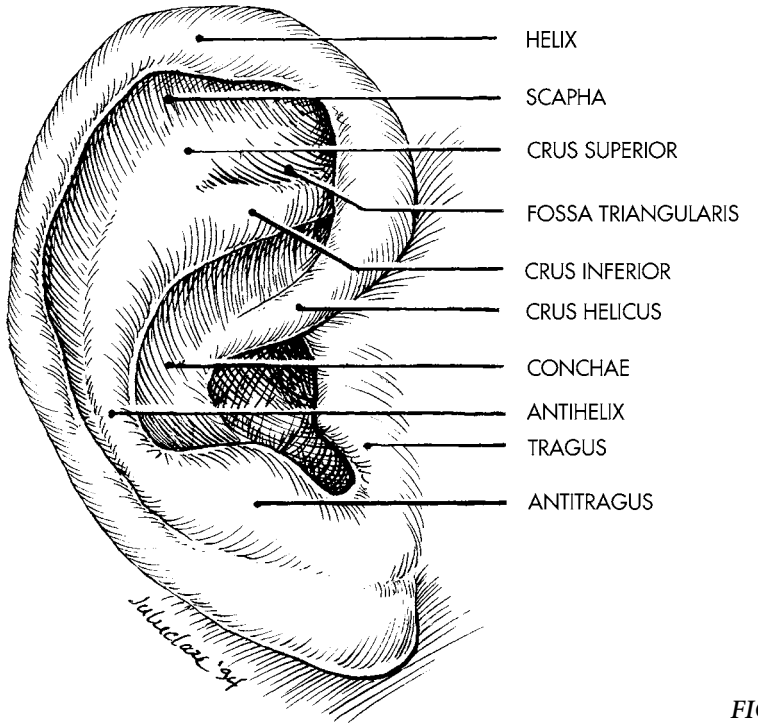


FIGURE 37–1. Auricular landmarks.

formed by the concha on the anterior (lateral) surface of the ear corresponds to a bulge or convexity on the posterior (medial) surface of the ear that is known as the eminence of the concha.

Anterior to the concha and partially covering the external auditory meatus is the tragus. The antitragus is posteroinferior to the tragus and is separated from it by the intertragic notch. Below the antitragus is the lobule that is composed of areolar tissue and fat. Anatomic variations such as preauricular tags or Darwin's tubercle, a small projection along the helix, may be present on the ear and should be recognized and documented in the preoperative assessment.[6]

Structural Features

Except for the lobule, the auricle is supported by thin, flexible elastic fibrocartilage. This cartilaginous framework is 0.5 to 1.0 mm thick and covered by a minimum of subcutaneous tissue.[4] The skin is loosely adherent to the posterior surface and helix of the auricular cartilage. The close approximation of the skin to the anterior surface of the cartilage provides the auricle with its unique topographic features (Figure 37–2).

The auricle has two groups of ligaments and musculature. The extrinsic ligaments connect the auricle with the side of the head and the intrinsic ligaments connect various parts of the cartilage to itself and to the external auditory meatus. The intrinsic group of muscles is rudimentary and serves no recognizable functional purpose. These include the major helix, minor helix, tragus, antitragus, transverse, and oblique muscles. The extrinsic muscles include the anterior auricularis, superior auricularis, and posterior auricularis. Certain individuals may have limited control of these muscles to "wiggle" the auricle.

The arterial supply of the auricle is derived from the anterior auricular branch of the superficial temporal artery, the posterior auricular branch of the external carotid artery, and a branch of the occipital artery. Venous drainage is from the posterior auricular vein and the superficial temporal, external jugular, and retromandibular veins. Lymphatic drainage from the ear moves anteriorly to the parotid lymph nodes and posteriorly to the cervical lymphatics.

Innervation to the extrinsic muscles is supplied by the facial nerve (cranial nerve [CN] VII). The posterior auricularis is innervated by the posterior auricular branch of the facial nerve proximal to the pes anserinus.[1] The temporal branch of the facial nerve innervates both the anterior and superior auricularis. Many nerve branches contribute to the sensory innervation of the auricle. The greater auricular nerve (C3) and the mastoid branch of the lesser

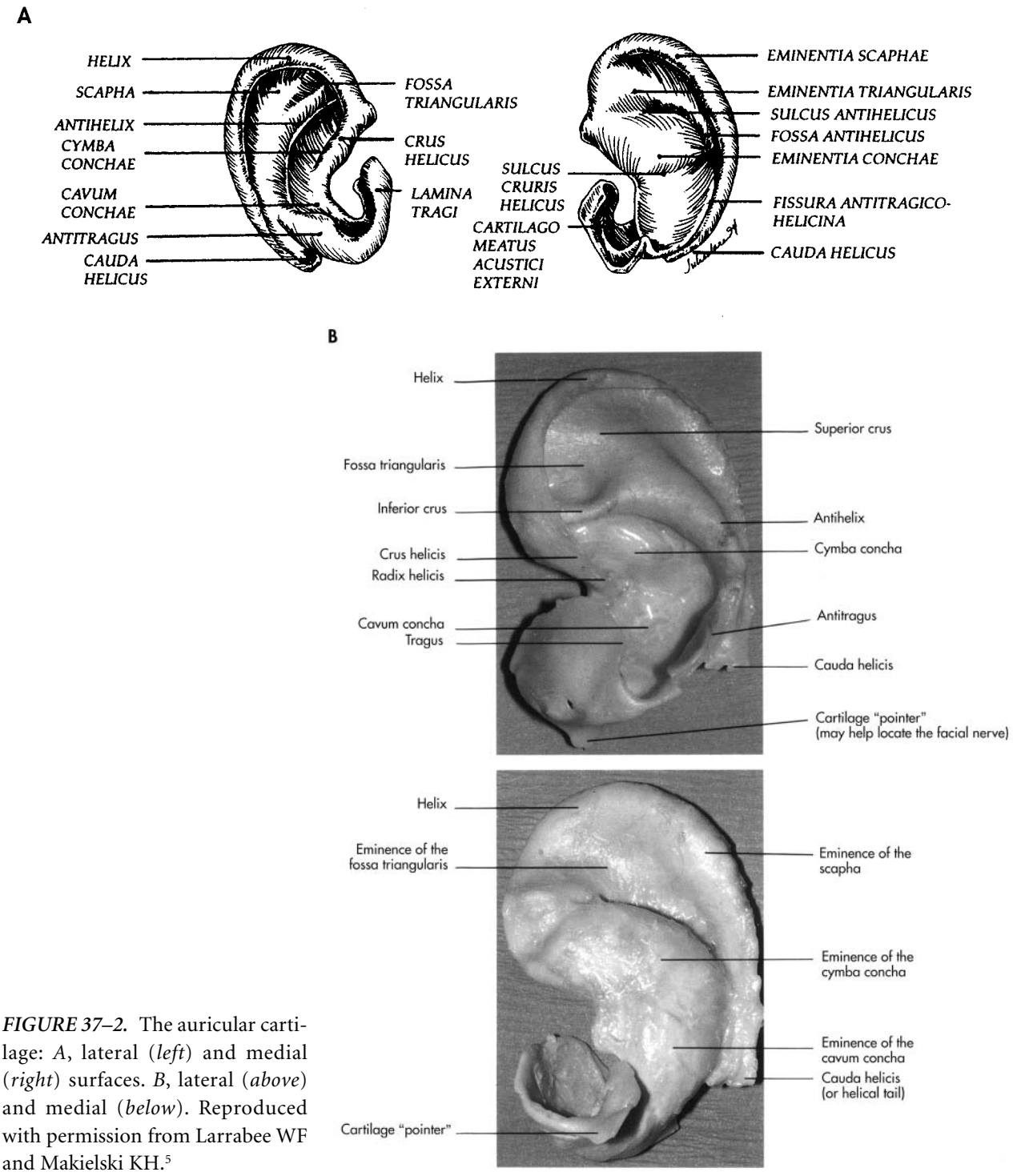


FIGURE 37–2. The auricular cartilage: *A*, lateral (*left*) and medial (*right*) surfaces. *B*, lateral (*above*) and medial (*below*). Reproduced with permission from Larrabee WF and Makielski KH.[5]

occipital nerve (C2, C3) carry sensation from the posterior surface of the auricle. The auriculotemporal nerve (CN V_3) provides sensory innervation to the anterosuperior portion of the auricle. Branches from the facial nerve (CN VII) and Arnold's nerve (CN X) innervate the concha.[4–6]

PATHOGENESIS OF THE OUTSTANDING EAR

Malformations of the auricle are not unusual and range from complete absence to macrotia. The incidence of abnormally protruding ears is approxi-

mately 5% in Caucasians.[7] As the auricle assumes a recognizable form by the end of the twelfth gestational week, the greatest number of malformations occur. Although they may be a dominant or recessive trait, most ear deformations are inherited as an autosomal dominant trait with incomplete penetrance.[8] Understanding the pathogenesis of these deformities aids the plastic surgeon in developing an operative plan.

Davis and Kittowski pointed out that during development, the ear protrudes from the head because the crura of the antihelix are not formed.[9] The margins of the auricle curl in the sixth fetal month to form the helix, followed by the folding of the antihelix and the development of the superior and inferior crura. The formation of the antihelix and its crura brings the auricle closer to the head.

The most common cause of outstanding ears is the lack of development of the antihelical fold (Figure 37–3). This malformation of the antihelix is present in approximately two-thirds of all cases of protruding ears. However, other pathologic features may also contribute to the outstanding ear. A wide, protruding conchal wall is present in approximately one-third of all cases. Additionally, the prominent concha is often accompanied by a thickened antitragus.[10]

The outstanding ear is a single entity within a wide spectrum of auricular malformation. Depending on the degree of severity, the protruding ear may also have structural abnormalities seen in the classically described "lop ear" and/or "cup ear." The term "lop ear" is used to describe a deformity of the helix characterized by a thin, flat ear that is acutely folded downward at the superior pole. In the "cup ear"

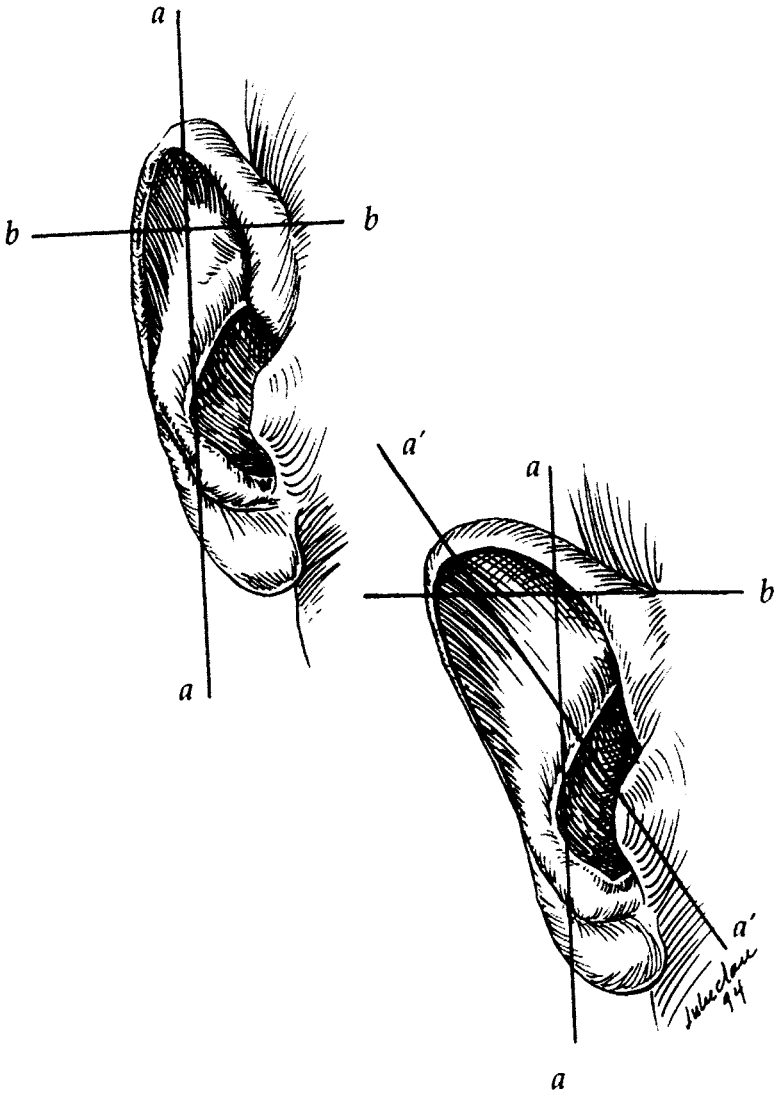


FIGURE 37–3. The most common pathology of the outstanding ear is underdevelopment of the antihelical fold. In the *upper figure*, the customary relationship of the ear to the head is seen. The plane of the scapha (represented by *a-a*) makes a right angle with the plane of the concha (represented by *b-b*). In the *lower figure*, *a′-a′* makes a much broader angle with *b-b*.

deformity, weak cartilage with resulting limpness of the auricle results in cupping or deepening of the conchal bowl. The "cup ear" is often smaller than normal and folded on itself. Poor development of the superior portion of the ear results in a short, thickened helix and a deformed antihelix (Figure 37–4). The surgical techniques used to correct the outstanding ear may be applied to the "lop ear" and to the "cup ear," but the correction of these malformations extends beyond the scope of this chapter.[7–10]

Preoperative Evaluation

Patients with outstanding ears typically present early in childhood, although some present in adulthood. The optimal age for surgical correction is between 4 and 6 years of age. At this age, the auricle is near or at full adult size, and the child is capable of participating in the postoperative care of the ear. Also, the child is typically about to enter school, and, unfortunately, children with protruding ears are commonly subjected to severe ridicule by their young peers.[11–13]

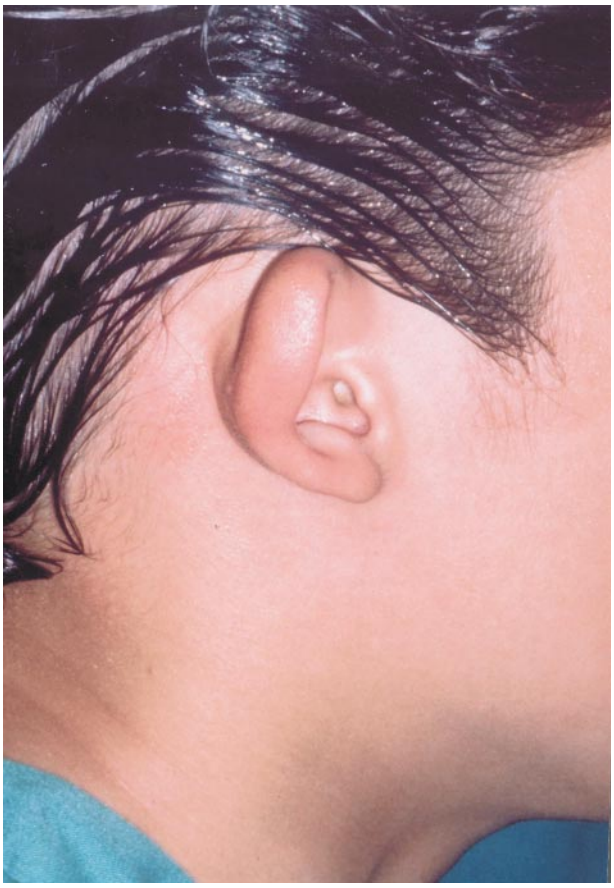


FIGURE 37–4. "Cup ear" deformity.

A comprehensive and quantitative approach for complete evaluation of the patient's ears is essential.[12] The auricles are compared with each other, both in overall symmetry and projection from the head. The proportion of the auricles to facial features and the head must be appreciated. The appearance of the auricles is judged by the symmetry from the front along the lateral helical rim. The superior aspect of the auricles should be level with the eyebrows. Additionally, development of the surface landmarks should be noted, along with additional features such as preauricular tags. The individual features within the auricle should be assessed in relation to other surface landmarks. For example, there should be a balance between the size and prominence of the helix and antihelix or between the tragus and antitragus. Finally, the redundancy of the postauricular skin should be noted, and the thickness and stiffness of the cartilage should be assessed and compared between the ears.

Precise measurements can be made to document the height, width, and axis of the auricle. Additionally, the angular relationships of the auricle and concha to the mastoid can be documented. Symmetry between the two ears can also be compared by a standard set of measurements, which can provide standards for assessment preoperatively and intraoperatively.[13]

Although the classic description of the outstanding ear attributes this deformity to the absence of the antihelical fold, two other attributes must be carefully assessed. Overprojection of the concha and/or the lobule will also contribute to the appearance of the protruding ear. Consideration and correction of these elements will contribute to the ultimate goal of a normal-appearing auricle.

As with any cosmetic procedure, preoperative and postoperative photographs are absolutely critical for careful planning of the surgical procedure and to document changes to the auricle. Uniform lighting and views of the auricle should be used before and after the surgery. The photographs should include a full-face anterior and full-head posterior view, an oblique/lateral view of both sides of the head, and close-up views of the ears. For patients with long hair, a hair clip or headband can be useful to prevent the hair from obstructing accurate photodocumentation.

GOALS OF SURGERY FOR THE OUTSTANDING EAR

The primary goal of otoplasty is to re-establish a "natural" appearance to the auricle and the relationship of the auricle with the head. Careful assessment of the outstanding ear, as described above, will reveal those individual elements of the auricle that contribute to its abnormal appearance. McDowell provided guidelines that should be considered when undertaking correction of the protruding ear[14]:

1. Symmetry of shape and protrusion of the ears should vary no more than 3 mm. Correction will often require bilateral alterations.
2. Maintain the normal appearance and curvature of the auricular components. The helix should arch backward naturally from its crus. It should be furled at its superior aspect and lead smoothly to the lobule. The antihelix should similarly curve forward into the superior crus.
3. The distance of the helical rim from the mastoid skin should be 10 to 12 mm at the superior pole, 16 to 18 mm at the middle one-third, and 20 to 22 mm at the level of the cauda helix. The proper auriculomastoid angle is 15 to 25 degrees. Achieving these distances may require reduction of an overly large conchal bowl.
4. The helical rim should not be seen beyond the antihelix from the frontal view, at least down to the mid-ear.
5. The postauricular sulcus should be preserved.
6. Protrusion of the upper one-third of the ear must be corrected. Protrusion of the lower ear may be tolerable, but only if the superior portion of the auricle has been corrected.
7. All visible surfaces should be smooth, without buckles, puckering, scars, and ridges that would reveal operative manipulation.

Although the physical dimensions and structural features are essential in the evaluation of the outstanding ear, the subjective assessment of the patient and his/her parents is also important. The surgeon should understand precisely what a patient dislikes about his/her ears and what he/she hopes the operation will achieve. This will help the surgeon determine if surgery can achieve the patient's desires and whether the patient's goals are realistic. Thus, the surgeon can explain the specific goals and limitations of otoplasty to the patient and his/her parents. Also, at times, unrealistic expectations for the surgery may exist. If this is the case, surgery should be deferred, and referral for counseling may be appropriate.

Coexisting anomalies, bleeding disorders, and any history or tendency toward hypertrophic scarring or keloid formation should be noted and addressed preoperatively.

SURGICAL TECHNIQUES: A HISTORICAL PERSPECTIVE

In 1845, Dieffenbach described correction of the outstanding ear through removal of skin from the back of the ear and suturing of the auricular cartilage to the periosteum of the mastoid bone.[15] In 1881, Ely and others described similar procedures with the additional excision of the posterior portion of the auricular cartilage.[16] More complicated procedures were later described that involved complex cartilage manipulations and even the use of fascia lata. However, these procedures are now primarily of historical interest because they resulted in ears that were abnormally flattened against the head and that often had sharp visible edges, excess skin wrinkling, and other problems.

In 1910, Luckett identified the absence or underdevelopment of the antihelical fold as the cause of the classic protruding ear.[17] His technique for the correction of this deformity involved the excision of a crescent-shaped segment of posterior skin and cartilage from the intended site of the antihelical fold and re-creation of the antihelix by everting and suturing together the cartilage edges. Unfortunately, this approach resulted in less than optimal results and fell out of favor. Nevertheless, Luckett's identification of the primary cause of protruding ears led to the development of currently employed techniques.

Otoplasty techniques have evolved toward the goal of a natural-appearing ear that is not too close to the head and does not have sharp ridges along the anterior surface of the auricle. Davis and Kittowski described a technique in which the new antihelix fold was placed in line with the inferior crus. In their approach, skin was removed from the posterior auricular surfaces of the concha and mastoid.[9] Cartilage was then removed from the region of the new antihelix, and the remaining edges were everted and sutured together.

Young would later modify this approach by creating the new antihelix along the superior crus.[18] Becker advocated multiple incisions along natural

lines of the cartilage to shape a new antihelix and to conceal surgical ridges along normal folds of the auricle.[19] However, this method still required excision of cartilage. Although the results from this approach were gratifying, like other cartilage-splitting procedures, the excision of cartilage from the antihelix occasionally left a sharpened ridge. This ridge was reported to be acceptable to most patients at the time but was disconcerting to the surgeon seeking a natural-looking result.[6]

In the late 1950s, Gibson and Davis demonstrated techniques to modify the shape of cartilage.[20] Disruption of perichondrium from one side of the cartilage led to bowing of the cartilage toward the surface with the intact perichondrium. Nachlas and associates and Stenstrom and Heftner described procedures based on this cartilage-sculpting principle.[11,21] Access to the anterior surface of the auricular cartilage involves a posterior auricular incision and either dissection over the helical rim or an incision through the auricular cartilage. Various instruments such as a wire brush, rasp, or diamond bur are used to disrupt the anterior perichondrium to permit bowing of the cartilage to form a new antihelix. These techniques are technically complex, requiring a great deal of surgical experience.

Mattress Suture Otoplasty

In the mid-1960s, Mustarde described a corrective otoplasty technique that gained quick and ready acceptance and wide popularity as it was seen as a marked improvement over existing techniques.[22] Horizontal mattress sutures placed in the auricular cartilage along the scapha re-create the natural curve of the antihelix, blending gently into the scaphoid fossa. Dimensions of the horizontal mattress sutures have been described with outer cartilage bites of 1 cm separated by 2 mm. The distance between the outer and inner cartilage bites is 16 mm (Figure 37–5).[13]

Advantages of this approach included that no through-and-through cartilage incisions are necessary, so the potential sharp edges of other techniques are avoided. Also, transperichondrial sutures may be positioned, test-tied, and then maintained or replaced as necessary to develop a natural antihelix. This was in contrast to the cartilage-splitting approach, in which the cartilage incision was irreversible and uncorrectable. Furthermore, the procedure has satisfactory long-term results and requires less dissection of the ear and less surgical trauma than other approaches. Surgeons also found this approach relatively easy to learn and to teach.

Whereas Mustarde's technique addressed the most common deformity of the protruding ear, the absent antihelix, Furnas described a suture fixation method to address the deep conchal bowl.[23] Furnas described the placement of a permanent suture to adjust the apposition of the conchal bowl to the mastoid periosteum, decreasing the angle between the concha and the mastoid. Additionally, a suture from the fossa triangularis to the temporalis fascia may further correct conchal height or contour. Care must be taken when placing the suture to avoid rotation of the auricle anteriorly with resultant external auditory canal narrowing.

Combination Techniques

Subsequent to the description of mattress suture techniques, cartilage-cutting approaches have remained useful in combination with suture tech-

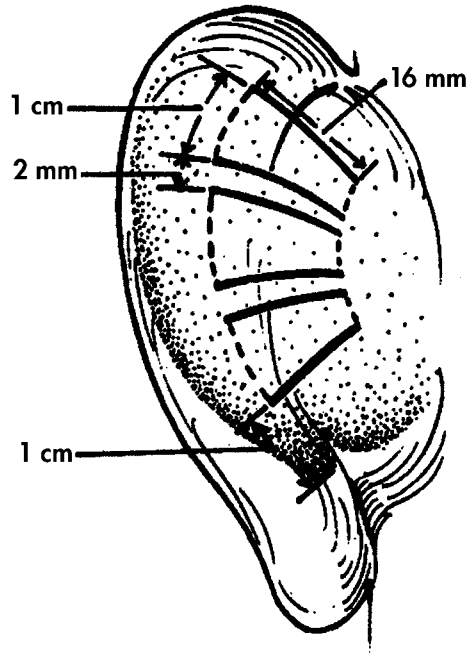


FIGURE 37–5. Posterior view demonstrating placement of sutures in a Mustarde suture otoplasty technique. The sutures must be placed through cartilage and anterior perichondrium but not through anterior skin. They should be placed close enough together so that the ear does not buckle when the sutures are secured. Reproduced with permission from Nachlas NE et al.[13]

niques. Pitanguy and Flemming devised a method to create an island flap of cartilage that was secured anteriorly to the rest of the conchal cartilage by placement of permanent horizontal mattress sutures.[24] Converse and Wood-Smith described a technique of thinning portions of the cartilage between incisions that followed natural lines of the cartilage. This technique facilitated the creation of an island of cartilage that sets anteriorly to the rest of the conchal cartilage to refashion the antihelix.[25] The technique of Farrior requires the removal of multiple longitudinal wedges from the level of the superior crus and the location of the new antihelical fold.[26] The auricular cartilage is then incised to permit shaping by the placement of horizontal mattress sutures.

All of these techniques have been employed by experienced surgeons to create a gently curved appearance to the antihelix and a pleasing appearance to the postoperative ear. The cartilage-cutting techniques may be especially useful when large anatomic deformities must be corrected or when the auricular cartilage is especially inflexible or thick. However, the caveat must be repeated that these are technically difficult procedures that require substantial surgical experience.

Mattress Suture Otoplasty: Method of Tardy

Mattress suture techniques with modifications are widely used today. One approach will be reviewed in detail here (Figure 37–6).[6] Suture fixation techniques are relatively easy to perform and do not require incisions or excisions of the cartilage that permanently alter cartilage characteristics or leave permanent postoperative stigmata.

Although the horizontal mattress suture is the primary mode of repair in this technique, it is important to emphasize the importance of addressing thick and inflexible cartilage. The mattress suture procedure is frequently augmented by thinning, weakening, and occasionally limited incision of the cartilage to achieve natural and symmetric results. Thinning the cartilage by shave excision or with a bur and incisions through the cartilage to facilitate folding will reduce the tension on the horizontal

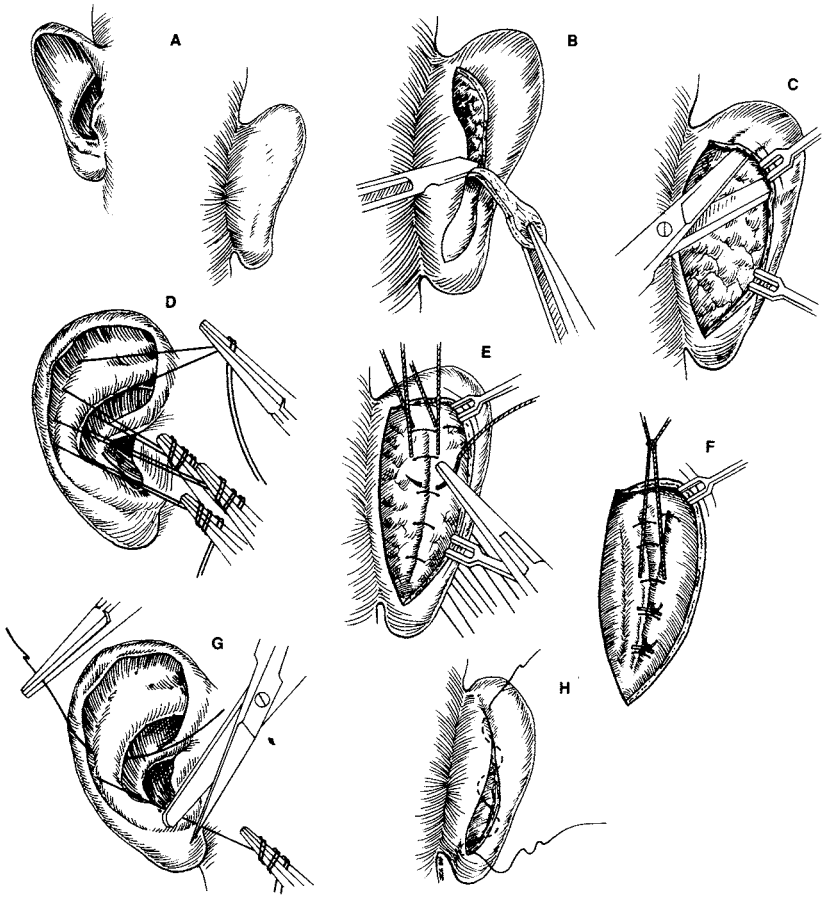


FIGURE 37–6. Modified mattress suture otoplasty technique (method of Tardy). *A*, Anterior and posterior outstanding ear deformity. *B*, Conservative excision of postauricular skin for surgical access. *C*, Conservative undermining of the skin. *D*, Temporary silk sutures placed as guides to accurate placement of permanent buried transperichondrial mattress sutures. *E*, Temporary sutures indicate sites for mattress suture placement. *F*, Transperichondrial mattress sutures positioned and test-tightened to determine the precise effect in creating a new antihelical fold. *G*, Temporary guide sutures removed. *H*, Postauricular incision closure shown here with a running intradermal absorbable suture. Alternative methods of incision closure are acceptable.

mattress sutures. Thus, every surgeon performing otoplasty must be comfortable addressing the protruding ear with more than one technique. Knowledge of one technique only is inadequate.

In the operating room, the ears are reassessed with regard to the causes of protrusion. Special attention is directed to the depth of the conchal bowl, the position of the lobule, and the strength and flexibility of the auricular cartilage. The periauricular areas are prepared with a sterile cleansing solution (hexachlorophene or povidone-iodine) and draped with sterile towels. The postauricular skin and subcutaneous tissue are infiltrated with local anesthetic (eg, 1% lidocaine with 1/100,000 epinephrine) for analgesia and hemostasis. The head is draped in a manner that permits comparison of both ears intraoperatively.

A fusiform or "elliptical dumbbell"-shaped incision is made posteriorly, exposing the portion of auricular cartilage in the area of the soon-to-be-formed antihelix. Care is taken to avoid removal of skin in the postauricular sulcus, which would cause flattening of the ear against the head. The skin is excised, leaving the posterior deep soft tissue and perichondrium that facilitates later scar formation, which is the strength of the repair. The remaining skin is undermined to the postauricular sulcus and to the helical rim. Meticulous hemostasis should be maintained at this juncture and throughout the procedure.

A deep conchal bowl, when it exists, may be addressed initially. Undermining along the posterior aspect of the cartilage reveals the posterior eminence of auricular cartilage underlying the conchal bowl. Excess cartilage in the posterior eminence frequently causes this area to impinge on the mastoid process, preventing the ear from resting closer to the head. Using a scalpel, small disks of cartilage can be shaved from this region to allow retropositioning of the auricle. This cartilage sculpturing is often sufficient to retroposition the ear and makes conchal setback sutures unnecessary. Excision of cartilage in this area will weaken the cartilage, reducing overall tension on the mattress sutures that will be placed in the antihelix region. *Great care is taken to achieve partial-thickness excision of cartilage, and through-and-through excision is avoided*. Nevertheless, on occasion, the auricle with a very deep cavum conchae may require conchal setback sutures or, rarely, the excision of a semilunar segment of cartilage within the cavum conchae to reconstruct the neoantihelix properly.

The new antihelix is created by manipulating the auricular cartilage and blending this fold into the inferior crus. Temporary 4-0 silk marking sutures may be placed from anterior to posterior to mark the location of the horizontal mattress sutures and thereby precisely guide their placement. This method avoids the use of ink or sharp needles to guide placement of the permanent sutures.

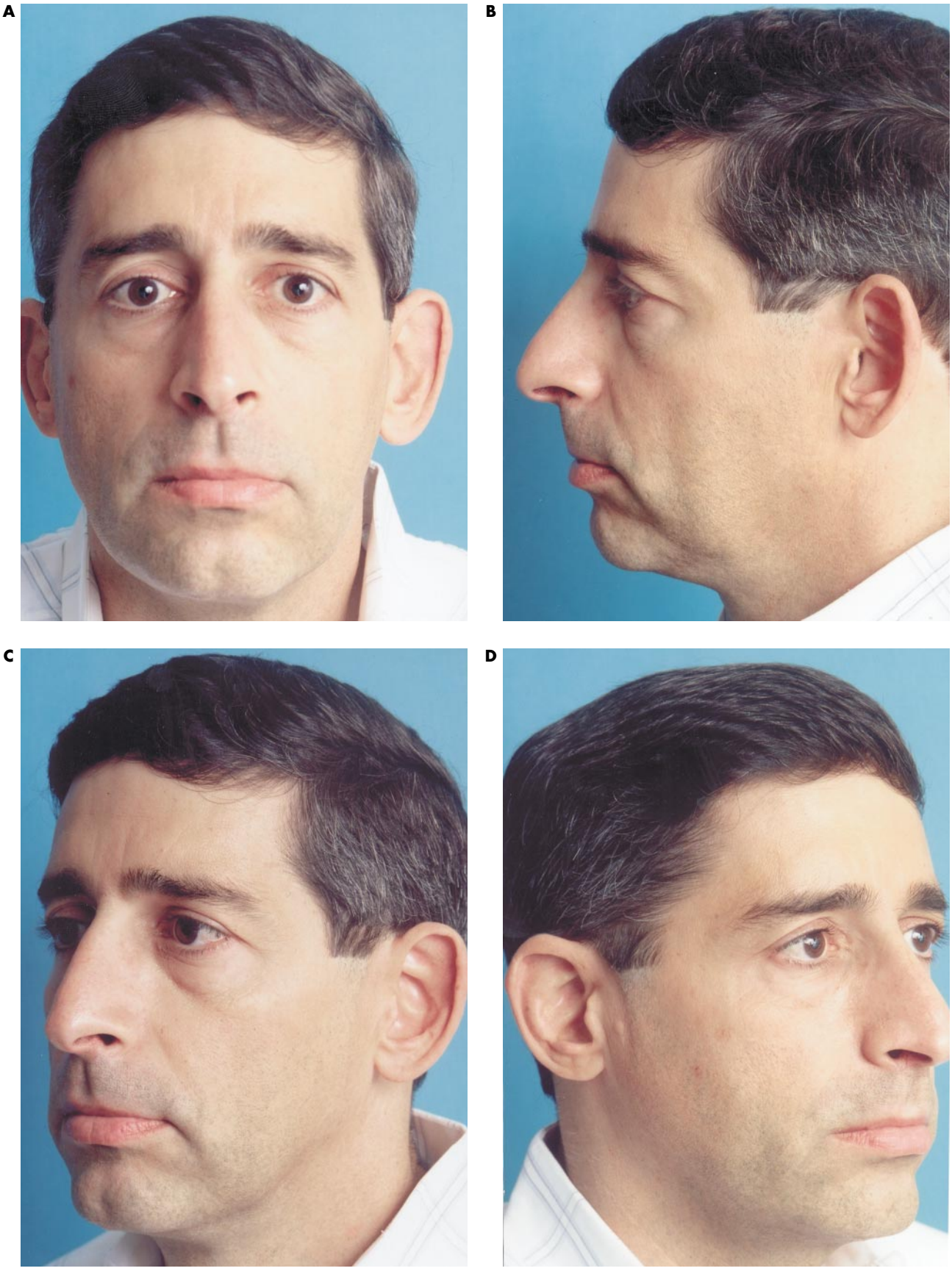
Once the new antihelix has been marked, 4-0 braided nylon (Tevdek) horizontal mattress sutures are placed sequentially, from caudal to cephalad along the neoantihelical fold. These horizontal mattress sutures are placed through the posterior perichondrium, auricular cartilage, and anterior perichondrium. Careful palpation with the free hand along the anterior surface of the auricle ensures that the needle does not pass through the anterior skin. Incorporation of the anterior perichondrium in the horizontal mattress suture is necessary to prevent the suture from tearing through the cartilage when it is tied down. Additionally, the sutures are not placed near the incision site to prevent future suture extrusion.

The horizontal mattress sutures are placed from caudal to cephalad and test-tied. Sutures are removed and replaced as necessary to achieve the desired fold on the auricular cartilage and then held with a hemostat. The sutures are tied securely once the antihelix has been completely formed. Typically, four or more mattress sutures are necessary to distribute the tension evenly and to hold the repair until sufficient scar tissue forms, usually in 2 to 3 months. If the sutures are adequately placed, it is unnecessary to overcorrect the repositioning of the auricle since the sutures will maintain their position without slippage. The postauricular skin is closed with a fast-absorbing 5-0 chromic gut suture in a continuous fashion.

Following creation of a neoantihelical fold, the position of the lobule is assessed. Ideally, the helix and antihelix should be in the same plane as the lobule. Commonly, simple skin excision and reattachment are sufficient to position the lobule in the appropriate plane, although more extensive intervention may be required at times.

The procedure is completed on the opposite ear. Frequent comparison between both ears ensures as symmetric a repair as possible. Given the nature of auricular deformities, complete symmetry between both ears is nearly impossible (Figure 37–7).

At the conclusion of the surgery, a conforming dressing is applied followed by a bulky head dress-


A
B
C
D

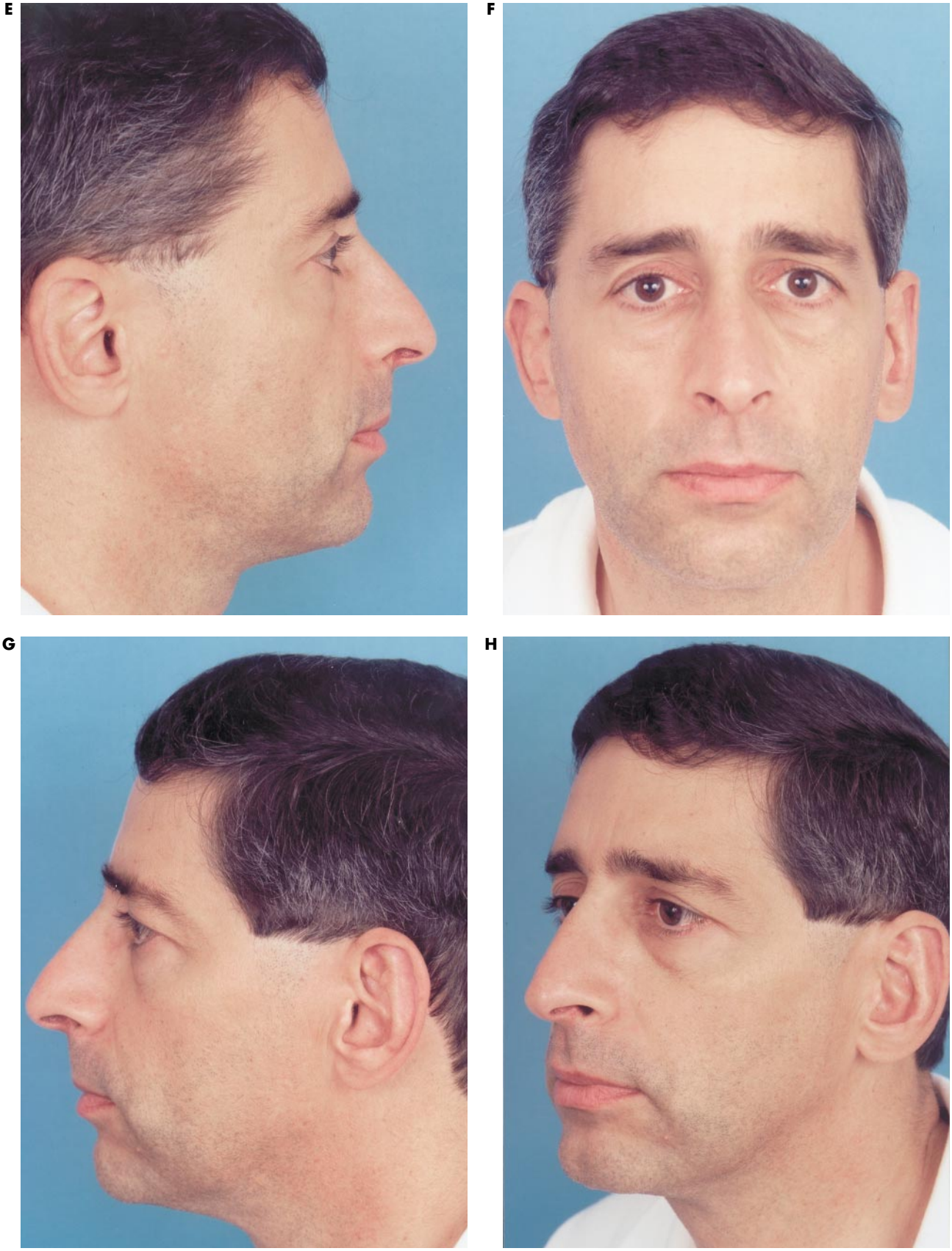


FIGURE 37–7. Preoperative (*A* to *E*) and postoperative (*F* to *H*) photographs of a patient who underwent left otoplasty with the technique described here. This patient had a relatively minor deformity of the right ear and was not disturbed by it, so only a left otoplasty was performed.

ing, which is removed and replaced with a smaller dressing that the patient wears for an additional 36 to 72 hours.

Outcomes and Complications

Surveys have demonstrated high rates of patient satisfaction (82 to 95%) following otoplasty.[21,27,28] Unsatisfactory cosmetic results were reported in 4.8 to 13.6% of cases; the rate was often higher for cartilage-cutting techniques.[29] Rates of relapse of auricular protrusion range from 2 to 13%.[27,30,31] Interestingly, more than almost half of the patients in this group reported an injury in the postoperative period.[30]

Regardless of the surgical techniques employed, complications (Table 37–1) are reported to occur at rates ranging from 7.1 to 11.4%.[13,27,28,30] In the early postoperative period, patient concerns regarding postoperative pain or periauricular "tightness" should be taken seriously as they may herald the presence of a hematoma. Physical examination might reveal fluctuance under a tense, dusky area of skin. The incidence of hematoma formation is about 3% and occurs slightly more often in cartilage-cutting techniques.[28]

TABLE 37–1. Otoplasty Complications

Early complications
Hematoma
Infection
Perichondritis/chondritis and permanent auricular deformity
Late complications
Paresthesia/hypersensitivity
Suture extrusion
Suture granuloma formation
Skin necrosis
Hypertrophic scarring
Keloid fomation
Anatomic/esthetic complications
Inadequate correction
Antihelix overcorrection and hidden helix
Telephone ear
Prominent concha
Malposition of the lobule
Sharp cartilage edges
Distortion of the external auditory canal (stenosis)

The cartilaginous framework is avascular, and incisions of the cartilage and manipulation of the thin skin–subcutaneous envelope may compromise the structural integrity of the auricle. Meticulous hemostasis during otoplasty and a postoperative pressure dressing are valuable measures for preventing postoperative complications. Immediate and late complications, if unrecognized, may lead to undesirable anatomic sequelae. Chondritis and necrosis of the auricle are among the most feared and potentially devastating complications. Early manifestations of these situations must be quickly identified and managed to protect the integrity and appearance of the auricle. Intraoperatively, great care must be taken during dissection, and hemostasis must be diligently maintained. In the event that a hematoma forms, the wound must be opened and the hematoma evacuated rapidly. Otherwise, infection leading to perichondritis, chondritis, and permanent cosmetic deformity may ensue.

Infection may occur secondary to an unevacuated hematoma but may also result from lapses in sterile technique or the presence of contaminated suture. Evolution to perichondritis occurs in approximately 1% of all cases.[28] Treatment must be aggressively implemented to prevent chondritis and resulting cartilage necrosis and auricular deformity. Antibiotic treatment must empirically cover *Pseudomonas aeruginosa*. Persistent infection may also necessitate the removal of permanent sutures and débridement locally. Preoperatively, a single dose of a broad-spectrum antibiotic can minimize the risk of infection.

In the late postoperative period, patients may complain of paresthesias and hypersensitivity, especially to cold temperatures in the weeks to months following otoplasty. These symptoms may indicate that injury may have occurred that is related to aggressive traction on the sutures intraoperatively or as a result of nerve transection. The complaints usually resolve or become tolerable over the next 4 to 6 months without active intervention.

Over the long term, permanent sutures may induce granuloma formation or be extruded. Suture extrusion may occur in as many as 10% of patients, but, typically, the area heals without adverse sequelae. In placing sutures, care must be taken in distributing cartilage tension over an adequate number of permanent cartilage-shaping sutures. Additionally, techniques to weaken the cartilage may decrease the

overall tension applied to the sutures and decrease the risk of suture extrusion.

As the wound heals, there is some risk of hypertrophic scarring and keloid formation, especially on the posterior surface. This is more common in younger patients, patients with a history of scarring, and in African American and Asian American patients. Conservative treatment includes triamcinolone acetonide injection weekly. Scar excision and closure without tension are also a consideration.

Esthetic complications arise from either incomplete or overly aggressive treatment of the original deformity. Although not considered a deformity in the nonoperated ear, a helical rim that is retropositioned relative to the antihelix is undesirable and is usually secondary to overcorrection of the neoantihelix. This deformity can be seen if Mustarde sutures are drawn too tightly or if overaggressive skin excision is undertaken.

In undertaking conchal repositioning, care must be taken in suture placement. Sutures placed too far posteriorly on the concha or too far anteriorly on the mastoid can result in stenosis of the external auditory canal.

Overcorrection of the middle portion of the ear leads to a "telephone ear" deformity owing to the relative prominence of the superior and inferior poles. One commonly described cause of the telephone ear deformity is over-reduction of the hypertrophic concha. Alternatively, overcorrection of the upper and lower poles may result in the "reverse telephone ear" deformity.

As discussed in this chapter, otoplasty methods that incise rather than reposition cartilage run the risk of visible sharp edges or prominent creases.

With significant attention focused on the antihelical fold and addressing the cavum concha, the position of the lobule may be overlooked. The lobule and helix should lie in approximately the same plane. Although amputation of the cauda helices or cauda repositioning is sometimes necessary, most commonly, excision of a soft tissue triangular segment from the posterior lobular surface is effective.

REFERENCES

1. Gulya AJ. Developmental anatomy of the ear. In: Glascock ME, Shambaugh GE, editors. Surgery of the ear. 4th ed. Philadelphia: WB Saunders; 1990. p. 5–33.
2. Wood-Jones F, I-Chuan W. Development of the external ear. J Anat 1934;68:255.
3. Farkas LG. Anthropometry of normal and anomalous ears. Clin Plast Surg 1978;5:401–12.
4. Weerda H. Embryology and structural anatomy of the external ear. Facial Plast Surg 1985;2:85–91.
5. Larrabee WF, Makielski KH. Surgical anatomy of the face. New York: Raven Press; 1983.
6. Kotler HS, Tardy ME. Reconstruction of the outstanding ear (otoplasty). In: Ballenger JJ, Snow JB, editors. Otorhinolaryngology head and neck surgery. 15th ed. Baltimore: Williams and Wilkins; 1996. p. 989–1002.
7. Bardach J. Congenital ear deformities. Surgery for congenital and acquired malformations of the auricle. In: Cummings CW, Frederickson JM, Harker LA, Krause CJ, et al, editors. Otolaryngology—head and neck surgery. St. Louis: Mosby; 1986. p. 2861–98.
8. Potter E. A hereditary ear malformation transmitted through five generations. J Hered 1937;28:255–8.
9. Davis JD, Kittowski EA. Abnormal prominence of the ears: a method of readjustment. Surgery 1937;2: 835–48.
10. Tanzer RC. Congenital deformities. Deformities of the auricle. In: Converse JM, editor. Reconstructive plastic surgery. 2nd ed. Philadelphia: WB Saunders; 1977. p. 1671–1719.
11. Nachlas NE, Duncan D, Trail M. Otoplasty. Arch Otolaryngol 1970;91:44–9.
12. Aguilar EA. Congenital auricular malformation. In: Bailey BJ, Calhoun KH, Deskin RW, Johnson JT, et al, editors. Head and neck surgery—otolaryngology. 2nd ed. Philadelphia: Lippincott-Raven; 1998. p. 2753–67.
13. Nachlas NE, Smith HW, Keen MS. Otoplasty. In: Papel ID, Nachlas NE, editors. Facial plastic and reconstructive surgery. St. Louis: Mosby-Year Book; 1992. p. 256–69.
14. McDowell AJ. Goals in otoplasty for protruding ears. Plast Reconstr Surg 1968;41:17–27.
15. Dieffenbach JF. Die Operative Chirurgie. Leipzig (Germany): FA Bruckhaus; 1845.
16. Ely E. An operation for prominence of the auricles. Arch Otolaryngol 1881;10:97.
17. Luckett WH. A new operation for prominent ears based on the anatomy of the deformity. Surg Gynecol Obstet 1910;10:635–47.
18. Young F. The correction of abnormally prominent ears. Surg Gynecol Obstet 1944;78:541–50.
19. Becker OJ. Surgical correction of the abnormally protruding ear. Arch Otolaryngol 1949;50:541–60.
20. Gibson T, Davis WB. The distortion of autogenous cartilage grafts: its cause and prevention. Br J Plast Surg 1958;10:257.

21. Stenstrom SJ, Heftner J. The Stenstrom otoplasty. Clin Plast Surg 1978;5:465–70.
22. Mustarde JC. Correction of prominent ears using simple mattress sutures. Br J Plast Surg 1963;16: 170–6.
23. Furnas DW. Correction of prominent ears by conchamastoid sutures. Plast Reconstr Surg 1968;42: 189–93.
24. Pitanguy I, Flemming I. Plastic operations of the auricle. In: Naumann HH, editor. Head and neck surgery. Philadelphia: WB Saunders; 1982. p. 1–18.
25. Converse JM, Wood-Smith D. Technical details in the surgical correction of the lop ear deformity. Plast Reconstr Surg 1963;31:118–28.
26. Farrior RT. A method of otoplasty. Arch Otolaryngol 1959;69:400–8.
27. Minderjhan A, Huttl WR, Hildmann H. Mustarde's otoplasty: evaluation of correlation between clinical and statistical findings. J Maxillofac Surg 1980;8: 241–50.
28. Goode RL, Profitt SD, Rafaty FM. Complications of otoplasty. Arch Otolaryngol 1970;91:352–5.
29. Davis JS. Prominent ears. Clin Plast Surg 1978;5:471–7.
30. Adamson PA, McGraw BL, Tropper GJ. Otoplasty: critical review of clinical results. Laryngoscope 1991; 101:883–8.
31. Maniglia AJ, Maniglia JJ, Witten BR. Otoplasty: an eclectic technique. Laryngoscope 1977;87:1359–68.

CHAPTER 38

Nasal Reconstruction and Rhinoplasty

M. Eugene Tardy Jr, MD

Incisions and Excisions

External nasal incisions should be sited in inconspicuous areas, leading to minimal distortion of nasal features and symmetry.[1] *Junctions of facial landmarks* hide surgical scars well; therefore, incisions along the nasomaxillary groove, the alar-facial junction, and the columellar-labial (or nasolabial) junction heal inconspicuously. Lesion excision in these areas should be preplanned so that the ultimate suture line(s) will symmetrically re-create these natural landmark borders. Strict attention to maintaining or re-creating symmetry leads to superior esthetic results. Local pedicle flaps transposed into these areas should be similarly designed.

Natural folds created by the synergistic interaction of muscle groups at the root of the nose provide ideal sites for incision and excision camouflage. Horizontal, oblique, and vertical wrinkles apparent in this area, blending into the glabellar region, provide wide latitude in scar camouflage in the aging patient (Figure 38–1). Redundant nasal and glabellar skin allows considerable excisional license without sacrificing normal landmarks. It is usually possible and always preferable to reconstitute these natural folds during the course of repair.

Alternate, but less ideal, incision sites exist. Midcolumellar incisions generally heal with minimal scar evidence,[2] although the lateral columellar incision lends similar surgical access and creates less potential scar. Staggered or W-shaped incisions in the caudal midcolumellar area are acceptable for tumor excision and as an approach to external rhinoplasty procedures.

Congenital nasal tumors (dermoid cysts, lymphangiomas) in children require incisional approaches through the nasal dorsum, with wide exposure required for total excision. A precise midline dorsal incision or semilunar transverse incision creates a visible but symmetric scar that, if meticulously repaired, fades acceptably with time. Open rhinoplasty incision approaches may be preferable.

Mature nasal scars on exposed nasal epithelium (ie, not camouflaged in landmark junctions or natural folds) may be rendered less conspicuous with superficial mechanical dermabrasion. Minimal

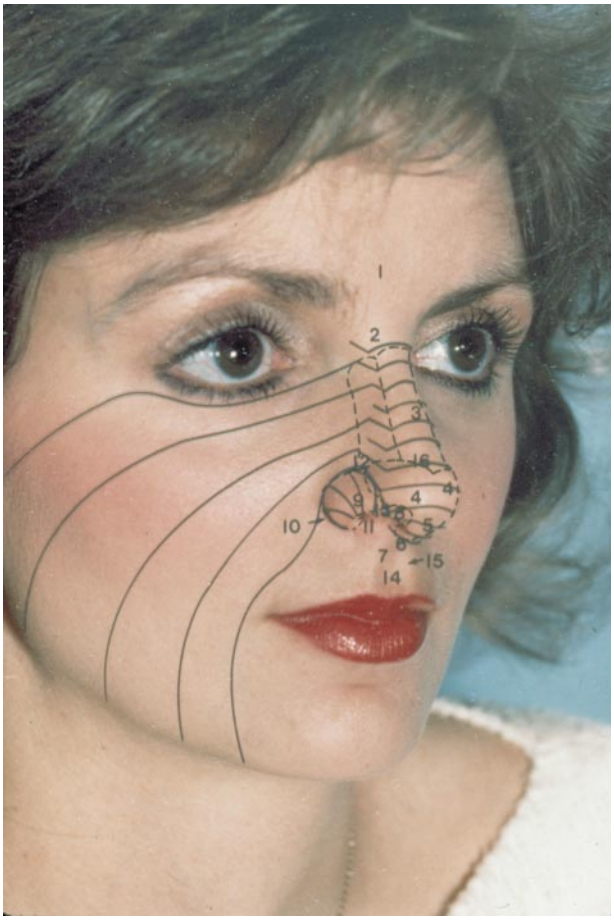


FIGURE 38–1. Favorable lines for incision and excision in the nasal and paranasal region.

scars resulting from laceration repair or pedicle flap reconstruction of the lower half of the nose respond particularly well to dermabrasion techniques (widened, depressed scars caused by inexact dermal healing require re-excision and meticulous repair; dermabrasion is ineffective in such cases).

A useful technique for diminishing the extent of a large nasal excisional defect, a cardinal principle in defect repair, is to remove a portion of an existing nasal hump (bone and cartilage), creating a relative excess of skin for repair, proportionally diminishing the defect, and favorably improving the patient's profile and nasal length.

Several areas of nasal anatomy should be considered inviolate because scars there almost always heal with unacceptable results. Incisions or lacerations cutting across the delicate alar rim remain conspicuous (notching is a common sequela), as may incisions cutting across the columella in transverse fashion. Because most incisions bridging a concavity contract and heal as a tight, tethered "bowstring" contracture, transverse incisions crossing the concave upper lateral nasal skin toward the inner canthus must be avoided or repaired with multiple Z-plasty techniques.

Finally, the nasal defect created by traumatic laceration or lesion excision provides the surgeon with an unparalleled opportunity to study exposed nasal anatomy and framework, aiding considerably in an understanding of the delicate, precise structural rearrangement necessarily less apparent during septorhinoplasty.

Epithelial Reconstruction

Epithelial nasal defects commonly result from minor to major excision of nasal lesions, benign or malignant. Occasionally, epithelial reconstitution is necessary for replacement of irradiated nasal skin of poor quality. Traumatic avulsion defects demand immediate repair if esthetics are to be properly served and unacceptable scarring and contracture avoided. Because it occupies the conspicuous central portion of the face, the nose draws immediate attention to itself, and nasal defects, however minor, accentuate that attention. Therefore, camouflage repair of nasal defects deserves high esthetic and functional priority, with strict attention to achieving reconstructive symmetry by repair that respects topographic subunit principles (Figure 38–2).

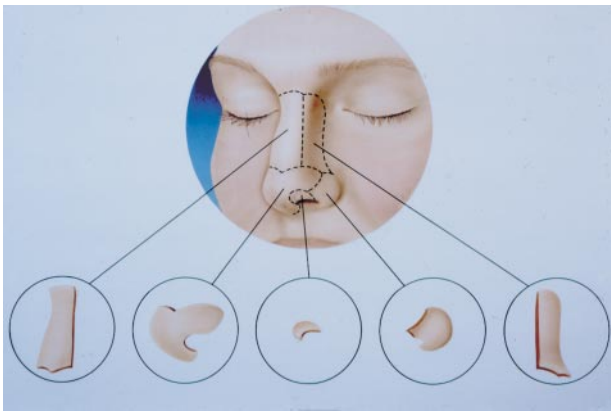


FIGURE 38–2. Topographic subunits of the nose (after Burget and Menich[1]).

Until recent decades, the split-thickness or full-thickness skin graft served the surgeon well as an effective tissue for immediate epithelial replacement. Skin is readily available around the head and neck, its rate of successful "take" is high, and one-stage repairs conserve the patient's time and the surgeon's time and skills.

Gradually, however, the multiple advantages of adjacent or regional pedicle flaps become apparent to head and neck surgeons interested and skilled in a higher degree of esthetic camouflage, effective one-stage (or, on occasion, two-stage) repair, and superior defect effacement and color match. Properly designed and executed flaps should replace missing tissue with like tissue, a fundamental concept in plastic surgery. Furthermore, flaps possess the following distinct advantages over skin grafts:

1. They provide their own blood supply.
2. They contract less than grafts.
3. Cosmetically and chromatically, they are superior to grafts.
4. They provide bulk and lining.
5. They create superior protection for bone and cartilage.
6. They resist infection.
7. They undergo minimal pigment change.
8. They may incorporate cartilage, bone, or skin in composite fashion.

In their simplest form, adjacent flaps may be classified as *advancement, rotation*, *transposition*, and *interposition* flaps, all designed and derived from tissue adjacent to the nasal defect. Regional flaps are best

used when more abundant tissue is required for repair or adjacent skin is inadequate or unsatisfactory. They use unipedicled flaps of similar texture, thickness, and color from adjacent regions (glabella, forehead, scalp, cheek, neck). A second stage of repair is required to transect the bridge of the flap, reconstructing both the defect and donor site to render them inconspicuous. Regional flaps are invariably designed and transposed from areas of relative epithelial excess and redundancy in the head and neck.

Advancement Flaps

Advancement flaps are created when tissue is undermined and advanced generally in a straight line, along the same axis as the defect.

Mobile cheek skin may be undermined and advanced to repair heminasal defects of varying proportions (Figure 38–3, A and B). Design is such that ultimate closure falls in the nasolabial fold and infraciliary area and at the junction of the nose to the face (Figure 38–4, A to C). Where possible, the nasolabial fold should be preserved or reconstituted to achieve bilateral facial symmetry, with incisions falling in the nasolabial crease. The inherent elasticity and redundancy of skin allow a variety of geometric designs in creating advancement flaps with ultimate suture lines lying in favorable areas for camouflage.

Advancement flaps along the dorsum of the nose may be created after transverse fusiform excision of lesions, diminishing the defect when possible by shortening the nose and accomplishing hump removal.

Rotation Flaps

A local flap in which the axis is created in a plane different from that of the defect is termed a rotation flap. Rotation flaps are extremely useful in nasal repair, particularly when designed so that the donor site may be closed in an inconspicuous straight line, natural fold, or landmark junction. Typical ideal donor sites are the nasolabial fold (Figure 38–5, A and B) and glabellar regions (Figure 38–6, A and B).

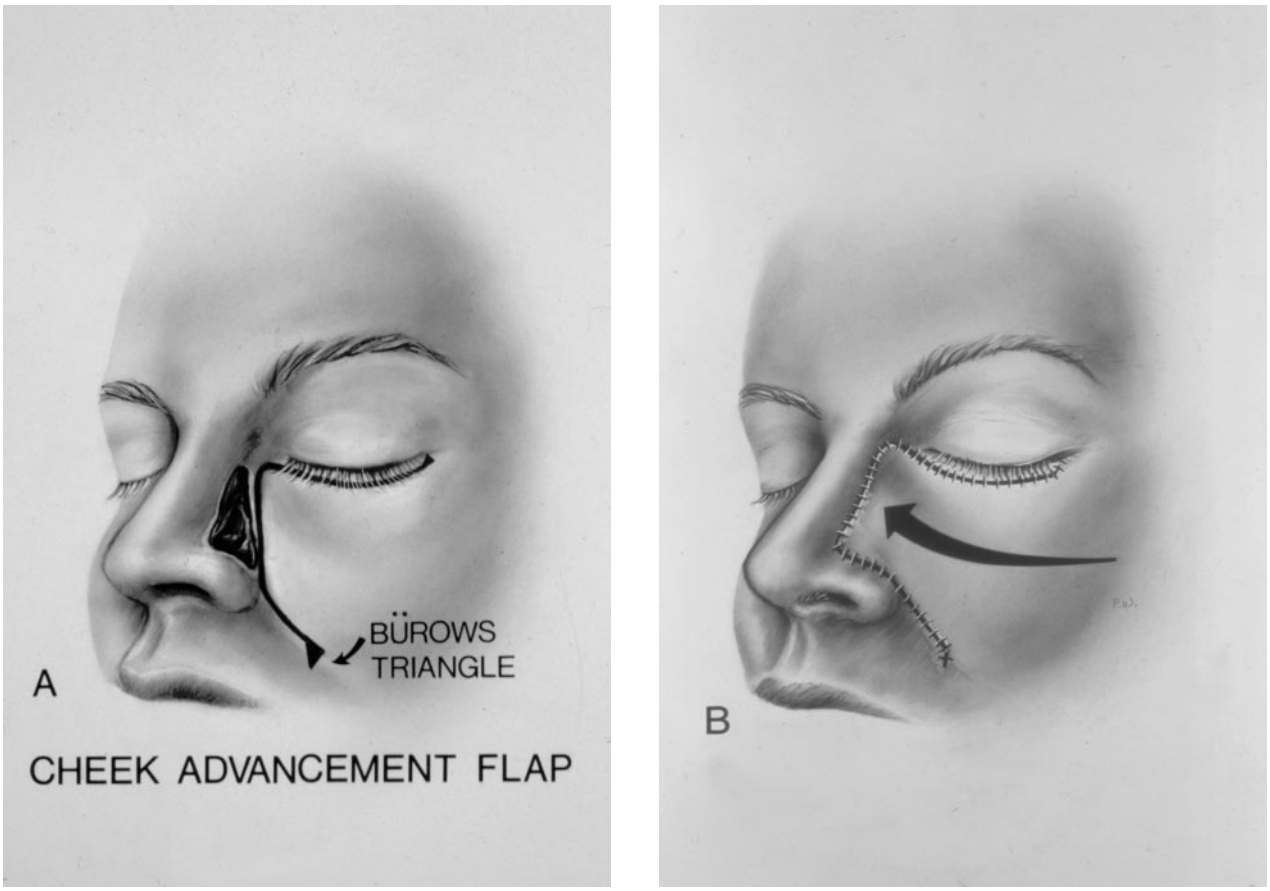


FIGURE 38–3. *A* and *B*, Cheek advancement flap for lateral nasal reconstruction.

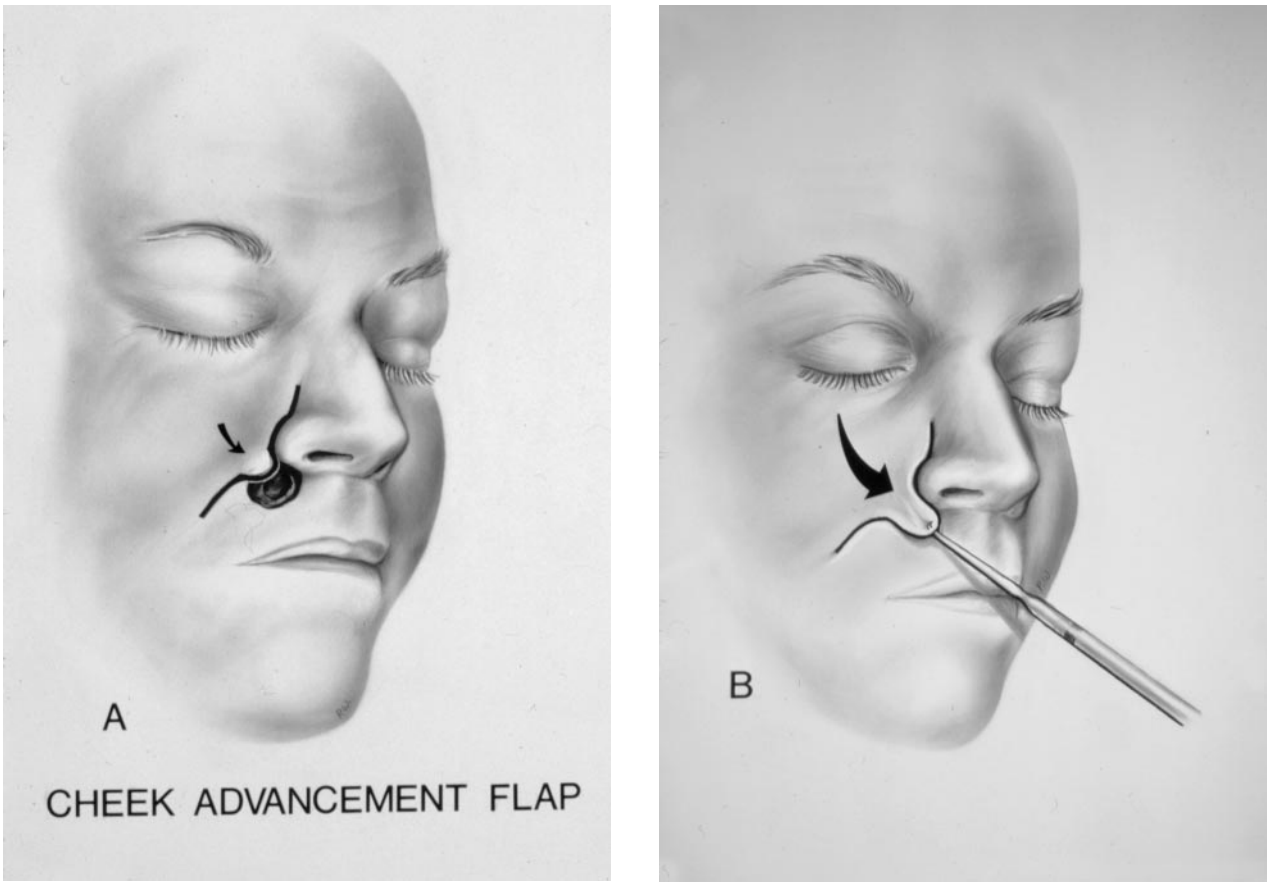


FIGURE 38–4. *A* to *C*, Modified cheek advancement flap to reconstruct the alar-facial junction and tissue void of the lateral upper lip.

The surgeon should approach any potential nasal defect with several alternative repair techniques in mind, ultimately choosing the flap or technique that possesses the most advantages and fewest risks for defect closure. The extent of almost all excisional defects may be diminished by circumferential undermining. The important concept of *reverse planning* is used, measuring the final defect size and planning the flap accordingly before creating final incisions. Small local flaps undergo minimal shrinkage during healing (less than 10%); therefore, almost exact flaps may be tailored to reconstitute the nasal defect precisely. Flap thickness should approximate the depth of the defect for full and complete effacement. Suture technique should include key interrupted, buried, absorbable sutures in the dermis and subcutaneous tissue to match the flap precisely to the defect, with nonabsorbable fine sutures used for 3 to 5 days to approximate epithelium. Tension-

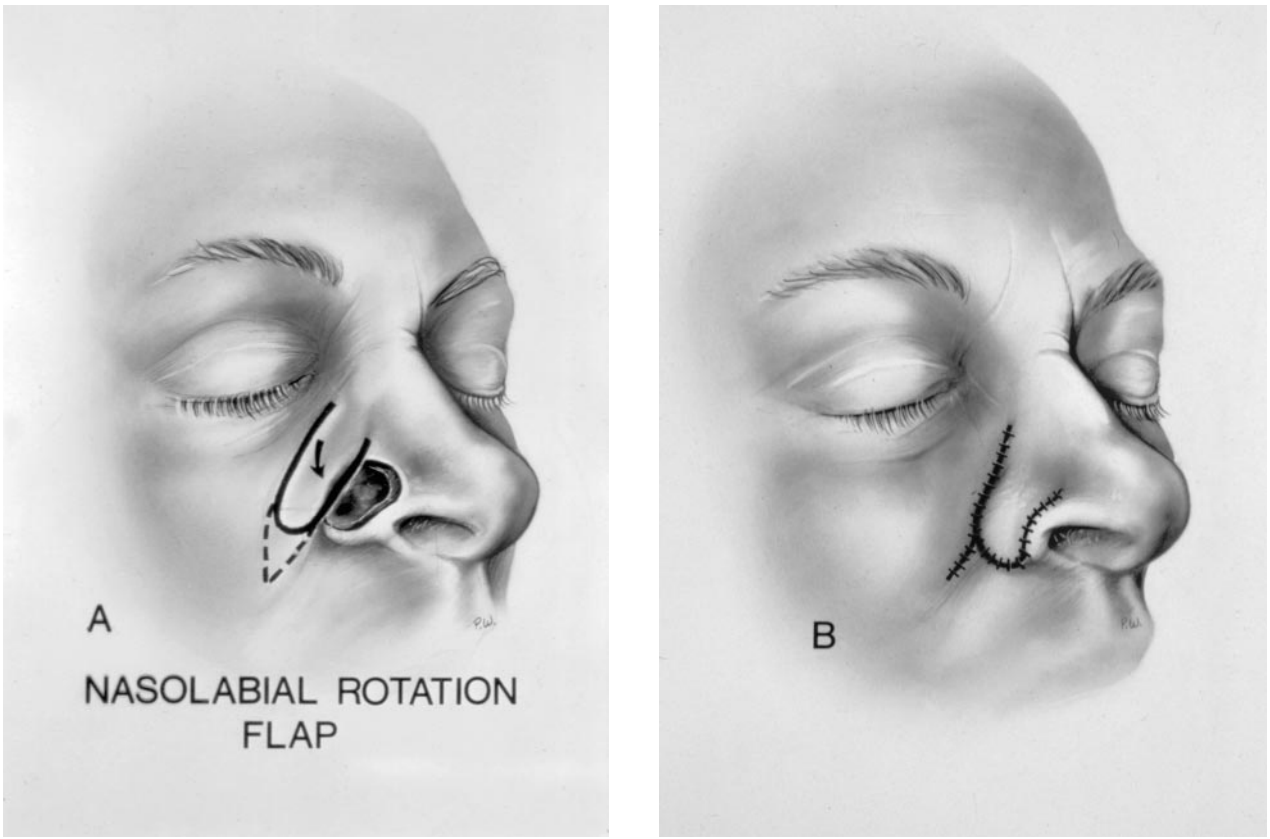


FIGURE 38–5. A and *B*, Nasolabial rotation flap for reconstruction of a nasal alar defect. The alar-facial junction is reconstituted.

reducing taping (Steri-strips) or tissue glues (Histoacryl blue, Dermabond) enhance wound approximation and allow earlier removal of sutures. Light compression bandages are used for 24 to 48 hours and then removed and replaced by Neodecadron ophthalmic ointment to keep the suture lines free of debris and clot, maximizing early healing.

Properly designed and executed, rotation flaps of adjacent nasal tissue may be expected to provide superior three-dimensional reconstruction of surgical and traumatic nasal defects within several weeks of repair. Defatting and "touch-up" procedures are rarely necessary. Light dermabrasion of the resulting fine nasal scars, 3 to 4 months later, can favorably enhance the blending effect.

Transposition Flaps

By far the most versatile of adjacent flaps is the transposition flap. Transposition flap designs available for nasal repair are varied and reliable. The variety of designs possible with transposed flaps provides the surgeon with many near-equal possibilities for nasal defect repair. Transposition flaps ordinarily allow primary closure of defects larger than those repaired with the simpler advancement and rotation flaps. When the flaps are properly designed, little tension is created on suture lines, and "dog ears" or "standing cones" are minimized.

In addition to the classic transposition flap design, *bilobed* transposition flaps, *rhomboid* transposition flaps, and the standard ubiquitous Z-plasty are superior methods of epithelial nasal reconstruction with adjacent tissue.

The classic transposition flap design is useful in the glabellar area, transposing redundant, lax tissue to upper lateral nasal defects. Proper design allows the ultimate incision lines to fall in or near natural folds, creating effective camouflage (Figure 38–7). Similar designs can take advantage of redundant nasolabial fold cheek tissue, sliding tissue medially to provide ample cover for lateral nasal tissue deficits; an undesirable side effect of this maneuver is the partial obliteration of the nasolabial groove, which is a highly desirable landmark.

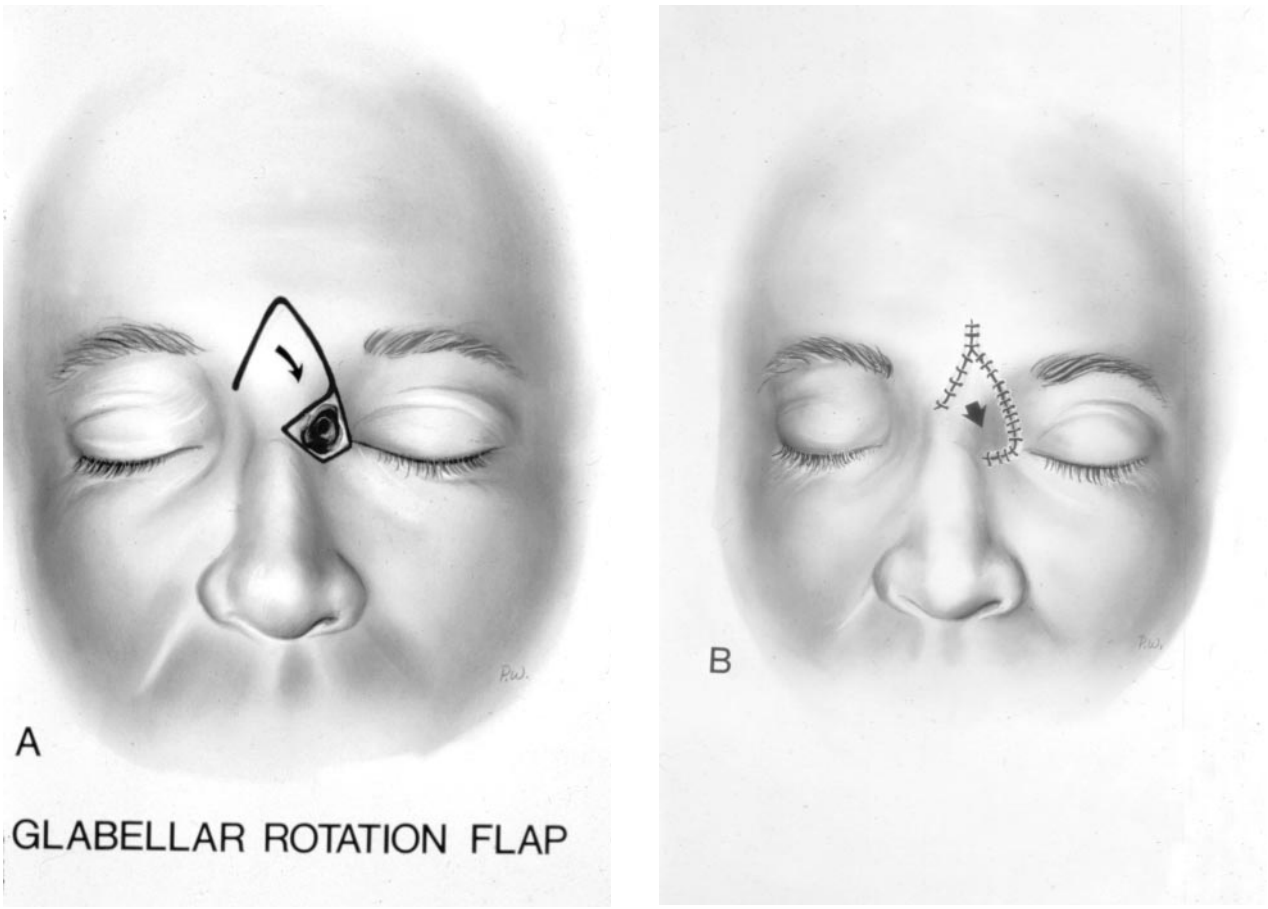


FIGURE 38–6. *A* and *B*, Glabellar rotation flap.

A tissue-abundant and versatile example of transposition flaps is the *bilobed flap* (Figure 38–8, A to C), consisting of two lobes separated by more or less of an angle and based on a common pedicle. This variety of flap design has several general applications in facial repair and two specific applications in nasal repair. Ideally, the dual flaps used should be similar (although not necessarily exact) in size and designed to rotate no more than 90 degrees for repair. In practice, this angle can vary from 45 to almost 120 degrees, depending on the location of the defect. In the upper half of the nose, in which the primary donor site of the bilobed flap may be easily reduced in size by undermining and advancement of its edges, the required size of the secondary donor flap may be reduced accordingly. The major advantage of the bilobed flap is derived from its use of the laxity (elasticity) and redundancy of tissue along two axes at approximately right angles to each other, compounding these tissues into one flap.

Abundant glabellar skin with vertical wrinkles, common in the aging face, combines favorably with the lax skin of the inner canthus area to supply tissue for the glabellar bilobed flap. The primary lobe of the flap is of ideal thickness and color match for dorsal nasal repair. The glabellar skin, which constitutes the secondary flap, is often thicker than inner canthal skin and must be carefully thinned after undermining to match evenly at the suture line. Heavy, thick eyebrows extending into the glabellar midline may impair the cosmetic effectiveness of this transposition flap.

Similar principles are involved in deriving donor skin for nasal resurfacing from redundant nasolabial fold skin (Figure 38–9, A and B). The secondary donor site is closed primarily in the fold, effectively camouflaging its presence. For large nasal defects especially, the cosmetic appearance of the healed flap is preferable in most situations to free skin or composite grafts. Tissue defects of the nonmobile

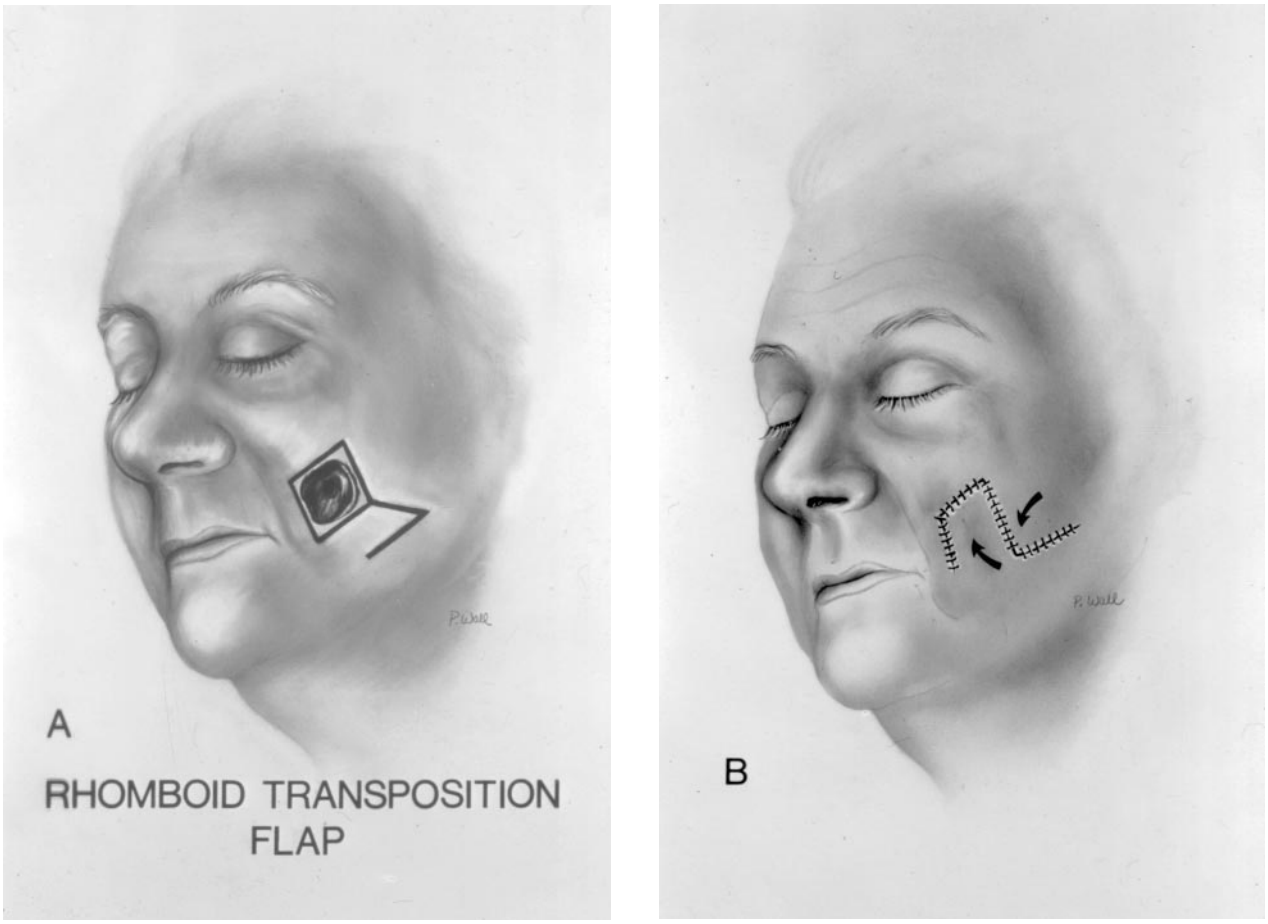


FIGURE 38–7. A and *B*, Rhomboid transposition flap.

skin of the nasal tip and infratip lobule are better repaired with full-thickness skin grafts harvested from the nasolabial fold or preauricular skin, however.

The principles involved in bilobed flap design can be used in a variety of reconstructive situations around the head and neck.

Regional Flaps

Pedicle skin flaps designed and transposed from head and neck regions other than those immediately adjacent to the nose are termed *regional flaps.* They are judiciously used when tissue in more abundance than that provided by adjacent flaps is required or when flap transport of buried skin, bone, or cartilage is required for framework reconstitution. In the latter incidence, the flap becomes a *composite* or *compound flap* and always requires delay and staging of 14 to 21 days before total elevation and transposition of the pedicle to the recipient site. Regional flaps are mainly designed around vigorous named vessels in the head and neck to allow nondelayed primary elevation and transposition. As circulatory efficiency in the flaps approaches its optimum level (14 to 21 days), the bridge of the flap is transected, final repair of the recipient site is effected, and all or a portion of the unused pedicle either is replaced in its previous anatomic bed or is discarded if donor site repair is unnecessary. Regional flaps from the midline forehead, scalp, and temple provide excellent tissue of appropriate bulk and near-ideal color match for nasal reconstitution. Furthermore, the ultimate donor sites in these areas are easily and effectively camouflaged, making them the site of choice for regional flap nasal repair. Tissue derived from cervical and more distant unexposed regions (shoulder, arm) lacks this ideal color match potential.

Midline Forehead Flap

Of all of the subtotal forehead flaps described for nasal reconstitution, the precise midline (or para-

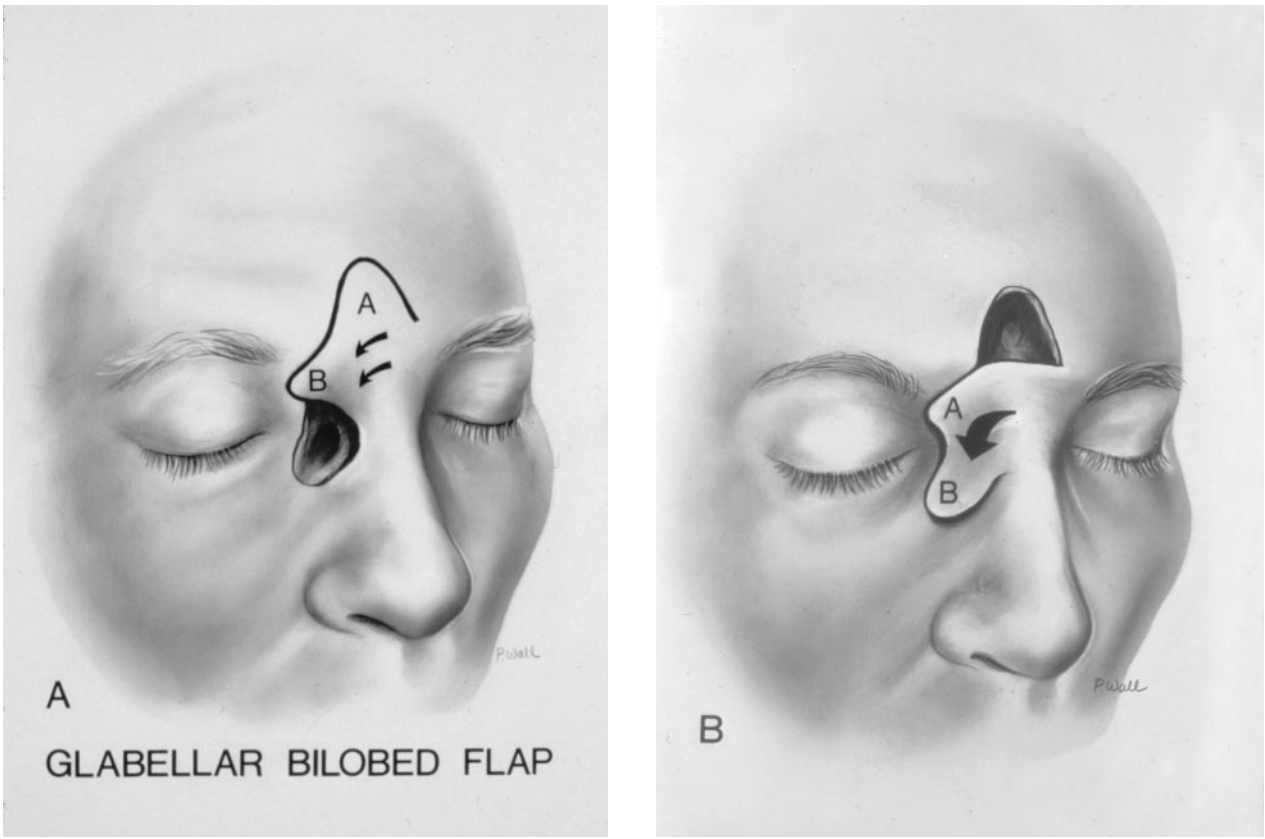


FIGURE 38–8. *A* to *C*, Bilobed transposition flap derived from redundant skin of the glabellar region.

median) vertical forehead flap is preferred for immediate transposition (Figure 38–10, A to C).[1] The color match is nearly perfect, the midline donor site defect may be closed immediately by advancing the lateral edges of the defect, and the resultant forehead scar is negligible if closure is meticulous and free of tension. Indeed, a midline vertical wrinkle already exists in many older patients, inviting scar camouflage. A flap width of 2.5 to 4 cm is available, and length is limited only by the extent of the hairline widow's peak. Abundantly nourished by the supratrochlear and dorsal nasal vessels, the midline forehead flap is suitable for primary restoration of defects of the lateral nasal epithelium, alar margin, and columella. A dressing of Adaptic gauze and antibiotic ointment substitutes for lining on the undersurface of the flap.

If desirable and appropriate, forehead tissue may be carried as a subcutaneous island pedicle flap beneath the tunnel of skin at the root of the nose; although this design provides one-stage repair, it

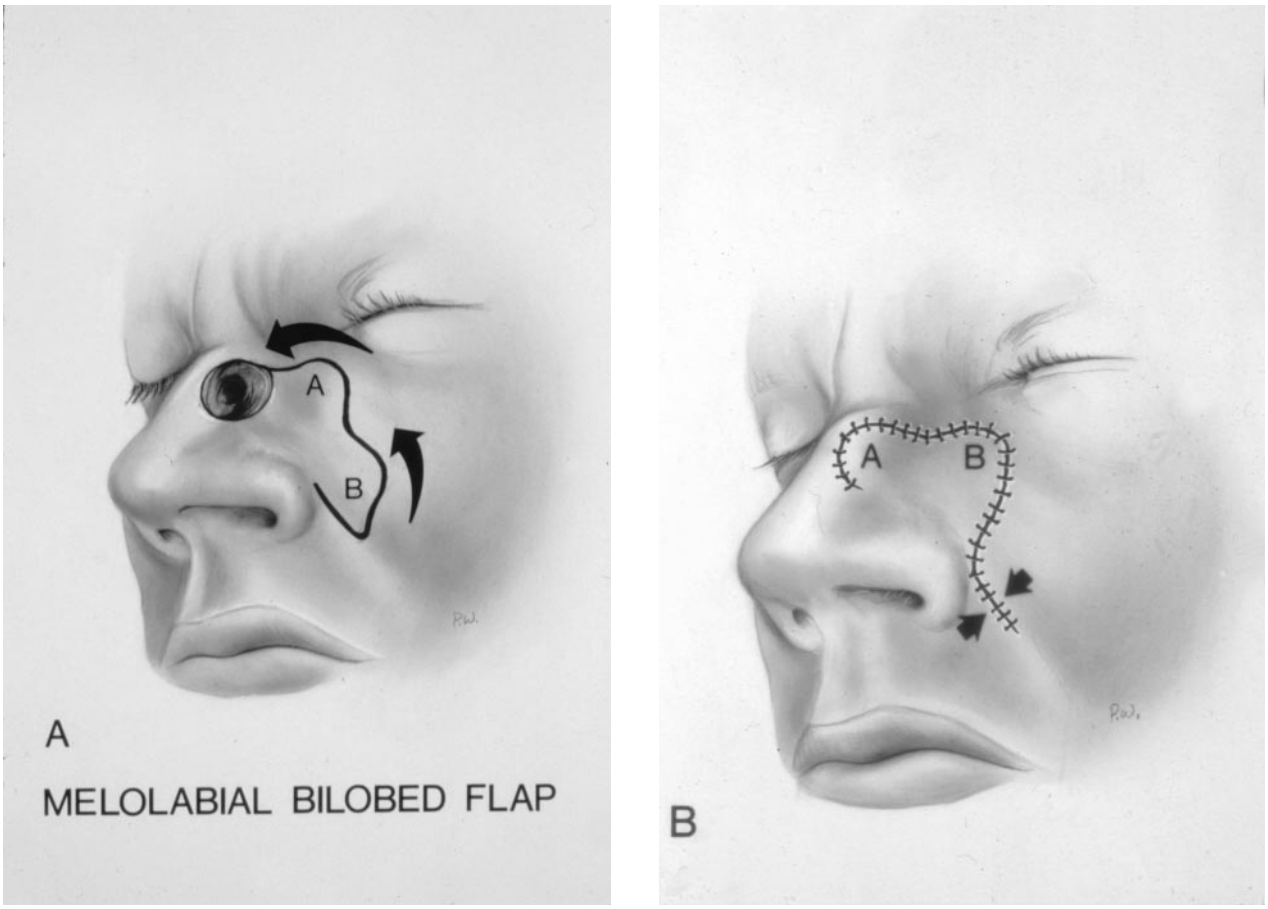


FIGURE 38–9. *A* and *B*, Paranasal bilobed transposition flap for reconstruction of lateral nasal defect.

adds unnecessary vascular hazard to the repair and often unsightly fullness in the tunnel region.

Fourteen to 18 days after primary flap elevation and transfer, the vascular viability of the transposed tissue is challenged by circumferential tourniquet compression; generally, division of the flap bridge is safely accomplished at this time. Most of the flap is discarded, restoring only sufficient tissue to reconstitute the oblique wrinkle lines at the root of the nose. Camouflage is excellent.

The esthetic advantage of the midline flap is apparent to all surgeons who use it regularly. Oblique, transverse, and "off-center" forehead flaps impose a higher penalty in ultimate scar legacy and therefore are less ideal esthetic choices. Midline flaps, carefully designed, preserve forehead symmetry and expression.

In near-total nasal reconstruction, midline forehead flaps can serve as the internal epithelial lining for the new nose, underlying a scalp flap designed to reconstitute nasal form. Redundant forehead skin created by tissue expansion is invariably stiff, less flexible and pliable; it is not considered ideal for nasal reconstruction.

Scalping flaps

Broad exposure of scalp and forehead tissue may be primarily elevated and transposed without fear of vascular embarrassment for near-total nasal reconstitution (Figure 38–11). In older men with balding tendencies, the donor site portion of the flap is best designed to lie on the scalp, ideally at the vertex of the skull. In younger patients and in women, the laterally placed skin superficial to the frontalis muscle is the donor tissue; a postauricular full-thickness skin graft covers the forehead defect over the intact frontalis muscle. Careful flap design with a reverse planning pattern technique provides ample tissue for alar simulation by enfolding skin edges for both

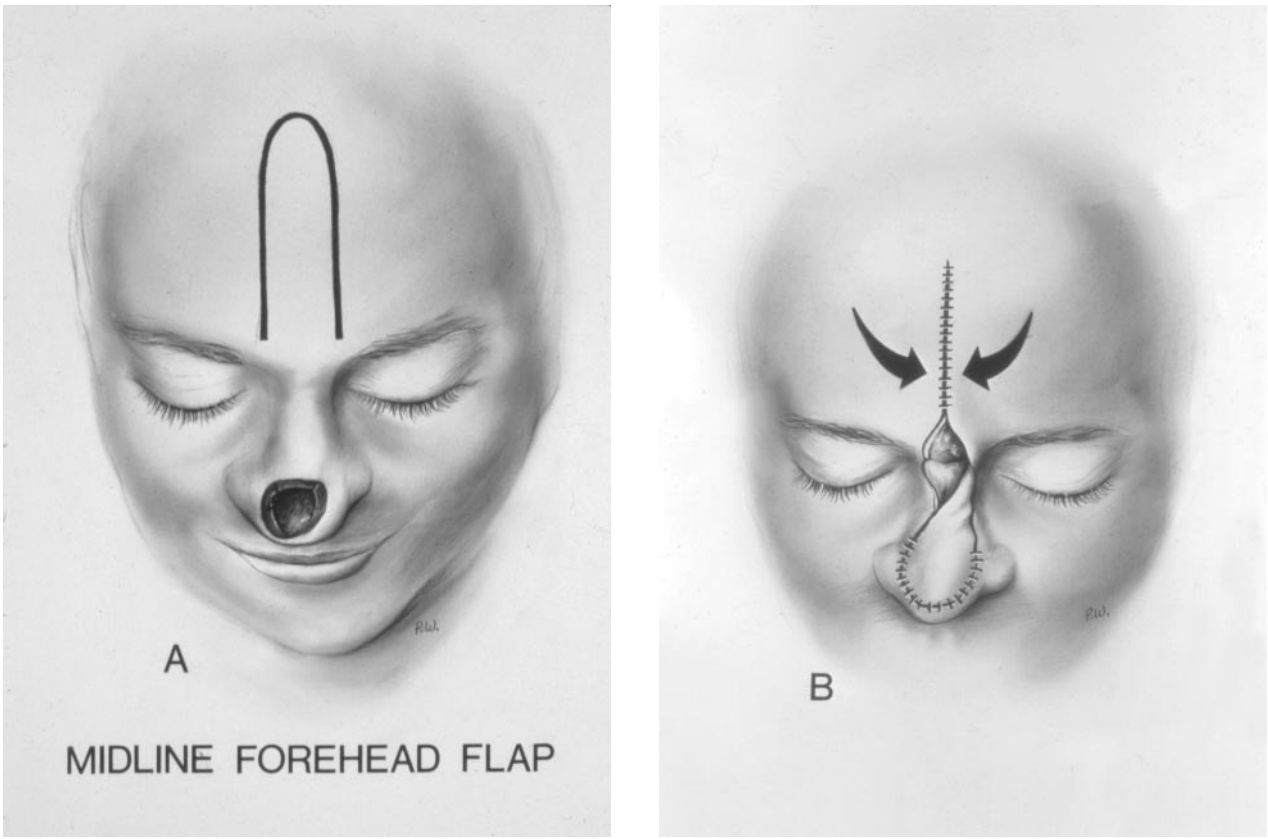


FIGURE 38–10. A to *C*, Midline forehead flap interposed into a large nasal defect. Flaps designed as paramedian, based on a single supratrochlear artery and vein, are preferable in certain reconstructive situations.

inner and outer nasal lining. Sufficient length must be incorporated into the design for adequate columellar length; failure to provide ample length disallows adequate tip projection. Small local turnover flaps should be developed when possible at the lateral edges of the nasal defect to provide an enhanced vascular bed for side-to-side rather than edge-to-edge suture repair.

Skin Grafts

Although infrequently used, skin grafts of split and full thickness are useful in nasal repair. The alternatives of primary closure or adjacent flap repair are generally preferable because of the inherent advantages previously detailed.

Skin grafts possess several distinct assets when employed for one-stage repair when no sacrifice of surrounding tissue is required: no further incisions

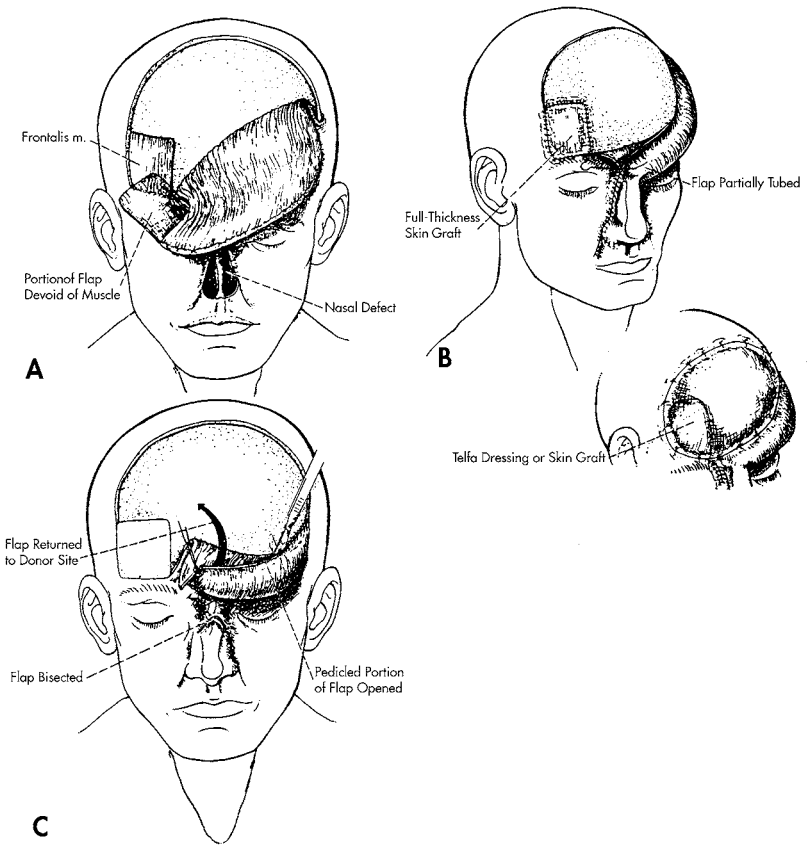


FIGURE 38–11. *A* to *C*, Scalping flap of Converse employs thin skin of the forehead for near-total nasal reconstruction.

are required adjacent to the defect, and planning is certainly less complex and sophisticated. Normal landmarks are generally undisturbed, and, should the graft fail for any reason, secondary grafting can proceed almost immediately after wound débridement, freshening, and infection control.

Superficial nasal epithelial defects of the lower half of the nose (trauma, burns, surgical excisions) are readily repaired with skin grafts derived from nasolabial, preauricular, or postauricular skin. The postauricular site has been shown by experience to be a less ideal color match than skin derived from the first two sites. An ideal but neglected donor site in older patients is *redundant skin of the nasolabial fold.* Full-thickness skin donated from this site is abundant and of excellent color match and allows camouflage of the donor defect in the nasolabial crease. Full-thickness defects of the nasal lobule and columella heal beautifully with nasolabial fold full-thickness skin grafts. Similarly, the preauricular and glabellar areas may harbor redundant skin with similar advantages. Seldom does the need arise to graft with skin derived from the more classic distant skin graft sites (inner arm, abdomen, thigh, buttocks).

Deep shave excision repair of the nose afflicted with rhinophyma may dictate the need for near-total skin graft repair of the nose. Shave excision and dermabrasion with preservation of epithelial islands, however, uniformly lead to re-epithelialization of remarkable esthetic appearance, thereby obviating grafting.

Grafts applied to the nose require delicate handling and special attention to detail in the operative and postoperative periods. Human noses are by no means inanimate; consequently, immobilization of skin grafts by means of stents or sutured-bolus dressings is necessary for the critical early period of 4 to 5 days. Tie-over bolus dressings compressing the graft dressed with Adaptic or Telfa gauze provide light but reliable immobilization and protection of the graft. In contradistinction, adjacent flaps require no compression dressing and thus provide convenience to the patient and early return to normal activities.

Split-thickness skin grafts seldom satisfy the need for three-dimensional augmentation or effacement of nasal defects; they may become depressed and thinner with time, seldom provide appropriate covering for implants of bone and cartilage, and

carry none of their own blood supply. For these reasons and those previously outlined, adjacent flaps remain the tissue of choice for immediate nasal repair in all areas except the nasal tip and infratip lobule. Specifically, grafts are generally inappropriate in areas of dense scarring and/or irradiation with consequently poor recipient site blood supply.

ESTHETIC SEPTORHINOPLASTY

Patients today seek functional and esthetic improvements in nasal appearance more than ever before. Rhinoplasty refinements currently allow subtle as well as major modifications with little discomfort, early healing, and predictable results. Perhaps no other surgical procedure blends artistic and technical skills to the degree required in esthetic rhinoplasty. Although it is one of the more common operations performed, only a few surgeons ever master its subtleties and nuances.

An artistic ability to visualize in advance the ultimate result is a critical skill necessary to surgical excellence. Successful rhinoplasty is initially preceded by careful analytic assessment of the nasal configuration, the deformity, and its relationship to the surrounding facial features. A realistic estimate of surgical correction, based on the possibilities and limitations imposed by the characteristics of the nasal tissues, is formulated—the preoperative "game plan." The goal of the surgery is to fashion a natural nose that is in harmony with its surrounding facial features and does not draw attention to itself.

The master surgeon is separated from the novice by the ability to *predict* the favorable and *compensate for* the unfavorable healing factors that influence ultimate nasal appearance evolving over many years. Only with continued experience and study, coupled with a continued impartial analysis of one's own long-term results, can the latter capability be developed and refined.

The truly capable surgeon must have wide knowledge of predictable procedures to be implemented in the unlimited variety of nasal configurations encountered. Strict adherence to basic principles does not necessarily always produce the ideal result. It is essential that an understanding of dynamic nasal structure transcend the components of static bone and cartilage; the relationship of shape and form to muscle tension and skin texture, the relationship of bone and cartilage to surrounding structures, the degree of postoperative thickening and/or relaxation of tissues, and the role of interrelated structures in the production of the deformity must be realized and evaluated.

Variations in rhinoplasty are manifold, ranging from minor corrections to complete reconstruction of the nose. Esthetic rhinoplasty aims at the creation of a nose that can be considered ideally proportioned, but proper physiologic function is essential. In post-traumatic deformities and many developmental and congenital deformities, the correction of respiratory derangements is paramount, and esthetics, although significant, is secondary. Some congenital deformities, such as the cleft lip or bifid nose, encompass both functional and cosmetic requirements of a special nature.

ESTHETIC RHINOPLASTY

The objective of esthetic nasal plastic surgery is the creation of a nose that is in harmony with the other features of the face. This simple statement represents a complex problem. The explanation of what is "beautiful" or "ideal" has been an age-old question, and the answer involves a multiplicity of emotional reactions and prejudices. In addition, values and assessments of beauty vary within different age groups and social structures. To evaluate what is beautiful entails a study of physical and cultural anthropology, ethnology, psychology, and esthetics, the ramifications of which may be endless.

Anthropometric Factors The anthropometrist uses basic measurements to divide and subdivide the face (Figures 38–12 and 38–13).[3] Designated points in the midline of the face are labeled as follows: the *trichion,* located in the center of the forehead at the hairline; the *nasion,* at the frontonasal suture line; the *subnasale,* a point at the root of the nasal spine; and the *gnathion,* the most anterior point of the symphysis of the mandible. If horizontal lines are drawn through these points, the face is normally divided into three equal parts. If, in profile view, another line is projected from the upper rim of the external auditory canal (the auricular point) to the lower rim of the orbit or infraorbital point, a line known as the *Frankfort horizontal* is formed. This line divides the face into two equal parts from the trichion to the gnathion (see Figure 38–12).

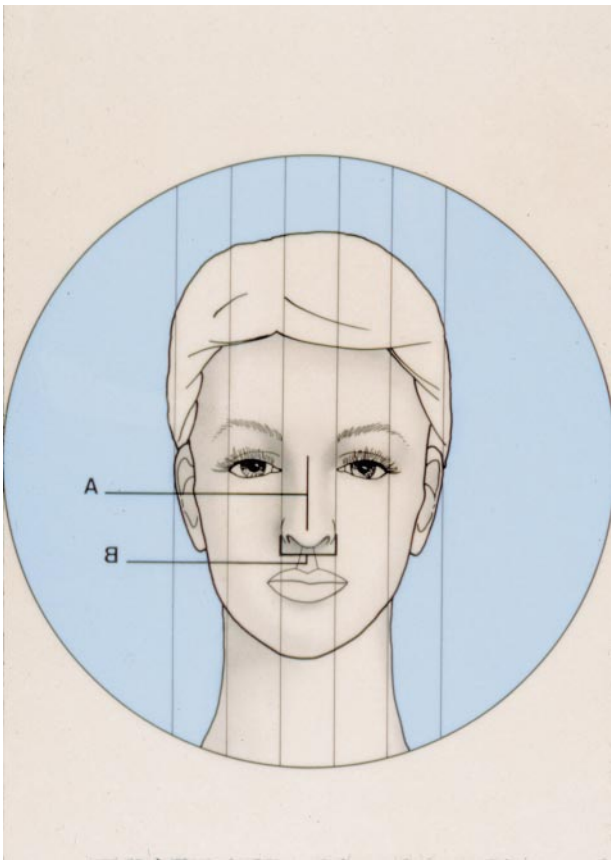


FIGURE 38–12. Anthropometric measurements of the "normal" face. The ideal face is equally divided into fifths.

Many of the basic observations on the divisions of the face were originally described in the fifteenth century by Leonardo da Vinci in the section of his notebooks entitled "Of the Parts of the Face." We use his accepted statement, "The space from the chin to the beginning of the bottom of the nose is the third part of the face, and equal to the nose and to the forehead," and have set up our blocks accordingly. After dividing these three groups into equal spaces, we are immediately led to certain observations. Although this spacing is equal, it is not monotonous for the forms lying next to each other occupy different numbers of spaces. Nevertheless, the proportions are equal, indicating a certain regularity in size, shape, and position of forms.[4] For example, the forehead occupies one entire group or one-third of the height of the face. The eye from the brow to lower lid fills one-third of the height of the nose, or, again, one can observe that the upper lip is to the lower lip and the chin as the height of the nostril is to the remainder of the nose. In this manner, one can divide and subdivide and compare indefinitely. It is interesting to notice that the eye tends to group a large area with a small one and to compare this grouping with other similar groupings of large and small areas.

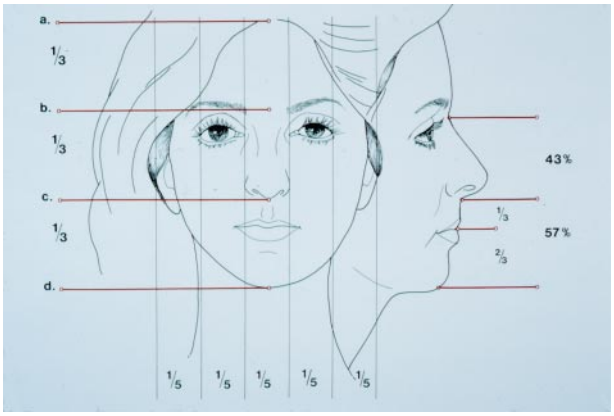


FIGURE 38–13. Relatively ideal facial dimensions.

The midline is the most obvious division of the front view. In the midline of the face, the measurement used is the width of one eye for "the space between the eyes is equal to the size of one eye" (da Vinci). Furthermore, by the size of the eye, the face is divided into five equal vertical sections. These may again be equally divided, and mathematical comparisons may be made, as before, to prove pleasing points of interest. An example is that the alae nasi occupy the space between the two inner canthi or one eye's breadth.

Psychological Factors The importance of beauty and the concept of our own body image, that is, the attitude regarding our physiognomy and body structure, are potent factors in the emotional and intellectual development of the individual.[4]

The nose, being a conspicuous organ, has definite secondary sexual characteristics and is frequently referred to as masculine or feminine. The patient with a poor body image, especially one who has an unconscious sexual or libidinous misidentification, may suffer a severe conflict because of a masculine or feminine appearance or the erroneous belief of such an appearance. Some of the most remarkable psychological changes in patients are in these individuals, often out of proportion to the surgical result. Regardless of the physical change from a masculine to a feminine nose, if the patient believes that the change has wrought a new identification

with his or her sex, the psychic change is spectacular. This psychic change is usually based on unconscious factors, the patient never being aware of the reasons for the dramatic change in attitude.

The psychological factors basic to an evaluation of one's appearance are instinctive and conditioned; therefore, much of the individual's reaction depends on his or her concept of the *body image.* To a great extent, this is influenced or conditioned in early life, usually by the attitude of the parents. The individual with a distorted and unrealistic self-image is usually an emotionally disturbed person. A *realistic* appraisal of one's own image, within reasonable limits, is usually indicative of a mature individual.[5]

Patients seeking rhinoplastic surgery for a gross deformity usually have a realistic view of themselves and a valid reason for correction of the deformity. The individual with a minor defect may harbor a poor self-image and sometimes exaggerates the deformity. The surgeon's ability to understand the motives and needs of patients seeking esthetic surgery is summarized by certain exhibited clues and characteristics. Diagnostic clues tending to identify the patient with a good prognosis include (1) obvious disfigurement, (2) occupational reason for seeking surgery to improve appearance, (3) a realistic wish to appear younger, and (4) a statement that the patient has "wanted to do this for a long time." A less ideal (or guarded) prognosis might be expected, according to Schulman, in patients who (1) state an unrealistic motive ("so I can look more masculine"), (2) request surgery on a sudden whim, (3) expect surgery to be the solution to all of their problems (eg, "to save my marriage"), (4) exhibit a history of hospitalization or treatment of recurrent psychiatric illness, (5) relentlessly go "surgeon shopping," (6) have undergone repeated surgery with consistent dissatisfaction, (7) are unable or unwilling to follow important instructions provided by surgeon or staff, (8) exhibit obsessive-compulsive tendencies, and (9) interact poorly with office personnel.

In summary, we can state that the motivations for and the psychological effects of rhinoplastic surgery are a complex subject. A few factors are reviewed, and an evaluation and a definition of beauty from perceptual and conceptual views are attempted.

Beauty of the human face is neither abstract nor absolute; it varies among different ethnic groups and is subject to the projection of the individual. This attitude is based on a multiplicity of factors varying according to the body image and cultural values of individual conditioning, particularly during the formative years, when it often becomes part of the unconscious mind.

In any ethnic group, beauty is the approximation of the so-called norm; for Western civilization, we have depicted this norm by mathematical proportions. It is therefore a form of esthetic ethnocentrism influenced by extraneous factors of culture. A great latitude in the variety and balance of features exists, so there are no absolute rules by which beauty may be determined.

Anatomic Description The nose on frontal view is in the shape of a pyramid, of which approximately the upper two-fifths comprise the bony vault and the lower three-fifths the cartilaginous vault (Figure 38–14). The upper narrow end joins the forehead at the glabella and is called the *radix nasi* or root of the nose, and its free angle at the lower point is termed the tip or *apex nasi.* The two elliptical orifices, the *nares,* are separated from each other by a skin cartilage septum known as the *columella.* The lateral surfaces of the nose form the dorsum by their union in the midline. The lateral surface ends below in a rounded eminence, the *ala nasi.*

The bony vault is composed of the two nasal bones, which are set on the nasal process of the frontal bone above, the frontal process of the maxilla laterally, and the perpendicular plate of the ethmoid and the septal cartilage below.

The upper portion of the cartilaginous vault consists of the two *upper lateral cartilages,* which are

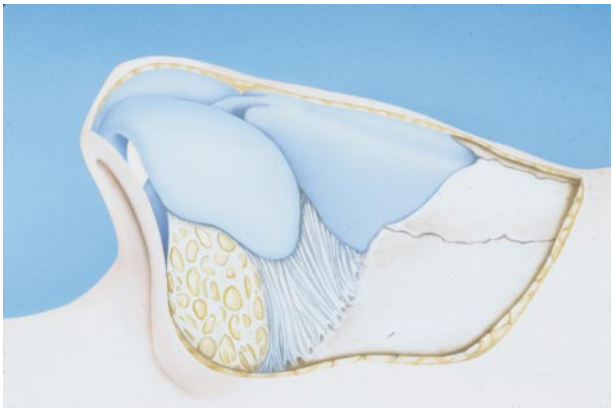


FIGURE 38–14. Normal anatomy of the lateral wall of the nose.

triangular and fuse with the septum in the midline. The upper margins underride the nasal bones and the frontal process of the maxillae and are attached to them by connective tissue. The lower cartilaginous vault is composed of two *lower lateral cartilages* (alar cartilages), which are of variable shape and more or less frame the nares and help in forming the alae. The two medial crura are attached to each other by fibrous tissue and to the lower end of the septum by skin, making up the columella and the membranous septum. Fairly near the midline, the lateral crura may slightly overlap the upper lateral cartilage and are attached to the lower rim of the upper lateral cartilage and the septum by connective tissue. They are also intimately attached to the overlying skin. The medial crura are loosely attached to the nasal septum and to each other by connective tissue. A few small loose cartilages (minor alar cartilages) are occasionally found laterally or just above the lateral crura.

The septum consists of both cartilage and bone. The septal cartilage is a single quadrilateral plate of cartilage that forms the anteroinferior portion of the septum. It unites with the bony portion of the septum behind with the perpendicular plate of the ethmoid, resting below on the groove of the vomer, the maxillary crest, and the maxillary spine (Figure 38–15). A tail-like posterior projection of the cartilage between the perpendicular plate and vomer is known as the sphenoid process. A small strip of cartilage, which is often absent, lies over the nasal spine and maxillary crests and is known as the vomeronasal cartilage of Jacobson.

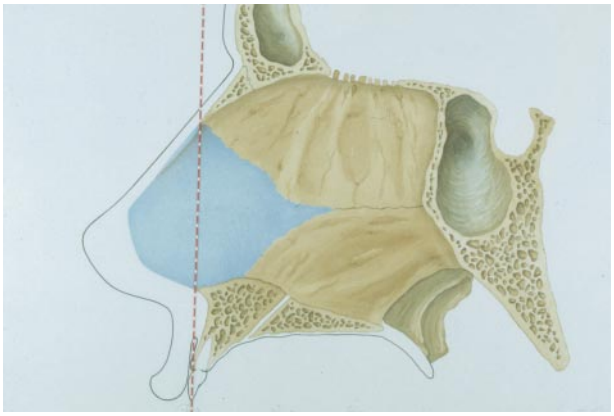


FIGURE 38–15. Normal anatomy of the nasal septum, lateral view.

The skin covering the external nose is thin and contains an areolar type of subcutaneous tissue. It is loosely attached in its upper half, but in its lower portions it is intimately bound to the lower lateral cartilages and may sometimes be thick and fatty and contain many sebaceous glands (Figure 38–16). The skin continues into the nares to supply lining to the nasal vestibule.

The muscles of the nose lie directly subjacent to the skin, and, occasionally, muscle bundles are attached to the cutis itself. The muscles comprise the procerus, nasalis, depressor septi, dilator naris posterior, dilator naris anterior, and angular head of the quadratus labii superioris (Figure 38–17).

The blood supply of the external nose is principally through the angular and lateral nasal branches of the external maxillary artery and the infraorbital branch of the internal maxillary artery. The internal nose is supplied by the sphenopalatine branches of the internal maxillary artery and the ethmoids from the ophthalmic artery. The veins terminate in the anterior facial and ophthalmic veins.

The motor nerve supply to the nose is by the facial nerve. The sensory supply includes the infratrochlear and nasal branches from the ophthalmic division of the trigeminal nerve and the infraorbital nerves from the maxillary division of the trigeminal nerve. The nasal septum is innervated by the ethmoid and nasopalatine nerves from the first and second divisions of the trigeminal nerve, respec-

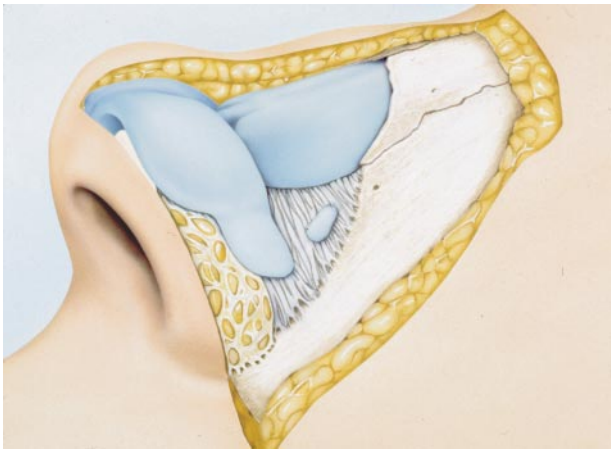


FIGURE 38–16. Rendering of a cadaver dissection displaying thick subcutaneous tissue cushioning the nasal skeleton, especially abundant at the nasal tip and supratip region.

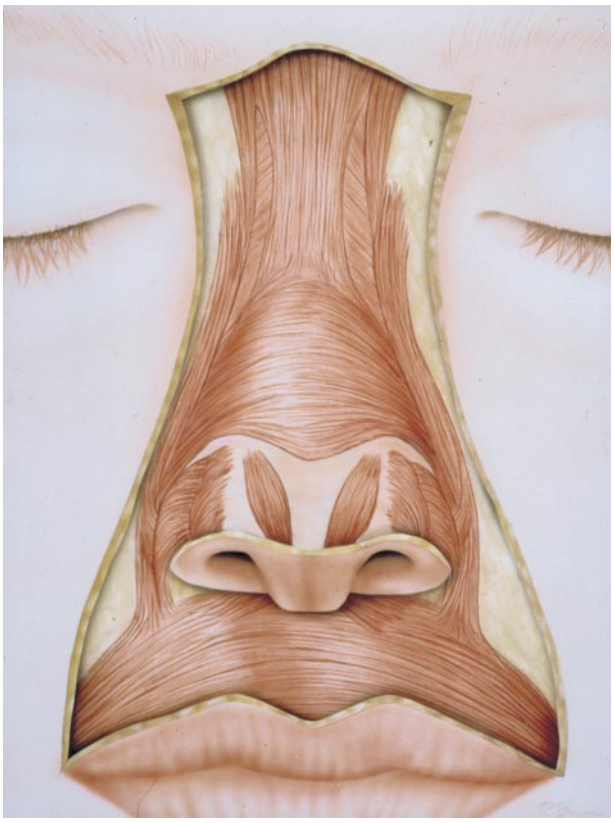


FIGURE 38–17. Mimetic musculature of the nose, interconnected in form and function in a superficial muscular aponeurotic system. Dissection in rhinoplasty should be beneath this layer, which contains the arterial and venous blood supply of the nose as well as lymphatic structures.

tively. The lateral wall of the internal nose receives fibers from the nasal branches of the palatine nerve, the ethmoid nerves, and a small branch from the anterior superior alveolar nerve.

Artistic Anatomy An artistic anatomic concept must be developed by the rhinoplastic surgeon when approaching the study of nasal anatomy.[6] The surgeon visualizes the nasal structure not only as static bone and cartilage but also as muscle tension, skin texture, relationship to surrounding structures, and the effect of related and interrelated structures on the shape of the nose. Furthermore, this concept divides the nose into an upper or immobile portion and a lower or mobile portion.

The immobile or upper portion of the nose is composed of the nasal bones and the upper lateral cartilages. The upper lateral cartilages are firmly attached to the undersurface of the nasal bones and maxillae by dense connective tissue and are fused with the septum in the midline, except for a small area at its lowermost end. The nasal bones are supported by the adjacent maxillae, the nasal process and spine of the frontal bone, and the septum below. Covering these structures are periosteum and perichondrium and the procerus and transverse nasalis muscles. The overlying superficial fascia or panniculus adiposus is thin, with scanty fat and fibrous tissue, allowing great mobility of the skin over this area.

The lower or mobile portion of the nose is composed of the two alar (lower lateral) cartilages and, occasionally, a few sesamoid cartilages superiorly and laterally. These cartilages are curved, thereby shaping the nares (Figure 38–18). The alar cartilages are more or less free-floating, being suspended and maintained in position by fibrous and muscular attachments to the two upper lateral cartilages, the septum, each other, and the overlying and surrounding skin.

The shape of the nasal tip is basically formed by the alar cartilages. Depending on the thickness of the skin, the amount of subcutaneous fat and areolar tissue, and the activity of the sebaceous glands, however, the configuration of the tip may vary considerably from the basic structure of the alar cartilages. Furthermore, the form of the nasal tip depends on its relative position to the immobile or fixed structures, which suspend and hold the tip in

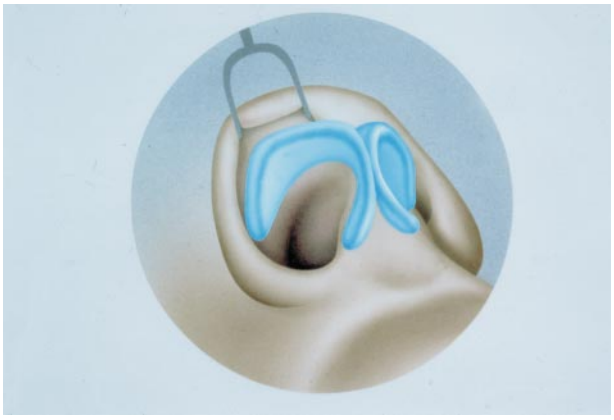


FIGURE 38–18. The paired alar (lower lateral) cartilages are essentially free-floating, mobile, and encased in soft tissues. They are loosely connected to the upper lateral cartilages, the ascending process of the maxilla, and the caudal septum.

position by firm fibrous attachments. A long, short, elevated, or depressed septum, long or short upper lateral cartilages, the shape of the nasal bones and maxillary processes, the spine of the maxilla, and the direction and tension of the surrounding muscles and fibrous connective tissue are determining factors in the position and shape of the nasal tip.

We are frequently surprised to note what remarkable changes in shape and position are accomplished by simply dissecting the tip free from the surrounding tissues. A further change is also noted when the last remaining suspensory attachment of the tip is freed by dissection of the alar cartilage from the enveloping skin. The skin at the nasal tip is immobile because of its firm adherence to the cartilage by dense fibrous connective tissue. This firm union aids in suspending and positioning the tip; thus, to mobilize it properly, the alar cartilage is detached from the overlying and surrounding skin. Occasionally, this may be all that is required to change the shape and direction of the tip. This is usually accomplished at the onset of the rhinoplastic operation. With mobilizing of the alar cartilages, the muscle fibers are also detached. The muscles involved are the nasalis (transverse and alar), dilator nares, depressor septi, and, occasionally, a few slips of the caput angularis of the quadratus labii superioris.

The muscles that surround and envelop the nose are part of the mimetic muscles of the face, which have no real fascial covering such as covers the skeletal muscles of the body. Instead, they are found below and in the panniculus adiposus or superficial fascia, and the fibers are inserted directly into the dermis and cutis of the skin. This arrangement of muscle fibers is a direct outgrowth of the highly complex development of speech, whereby small muscle fibers were needed superficially to allow multiple variations of facial movements.

Planning the Correction In planning the correction of the nasal deformity, the component features of the face and the general body build of the patient must be considered. Frequently, failure to consider the relationship of the nose with the chin, maxilla, and forehead mars an otherwise perfect rhinoplasty and results in a nose that appears unnatural for the particular individual.[7]

Procedures that have been advocated for analysis of deformity include marking angles on photographs, mechanical measuring instruments used directly on the face, and the use of computer imaging. All may be of value. It is, however, difficult to measure so subjective an abstraction as beauty in angles or degrees because of the many variables in physiognomy. Rather, one should strive to attain a harmonious relationship of all of the features. A simple but valuable method of studying the relationship of facial features is by the use of sketches made on the reverse side of the photograph. The photograph is held up to a light or placed on a roentgen film shadow box, and by pencil sketching and shading, the ultimate desired result may be depicted.

Whether or not sketches are made on them, pre- and postoperative photographs are an integral part of the surgical routine (Figure 38–19). In all plastic surgical undertakings, photographs are absolutely essential and should form a part of the patient's record. A series of standardized and uniform photographs is more informative than the most carefully detailed notes, both as a planning and self-instructional device and as a necessary medicolegal record.[8] Series of 5×7 inch black and white prints combined with color 2×2 inch color slides make ideal records in rhinoplasty photography and serve as an ideal immediate reference when projected full size in the operating room during surgery.[8] Preoperative photographs are followed by postoperative views taken at 1 week and then at 1, 3, 6, 12, 18, and 24 months. This carefully standardized photographic composite provides a clear uniform panorama of the dynamics of healing affecting the ultimate result and is a superior teaching record as well. The Frankfort horizontal line is used in proper positioning of the patient, providing uniformity and standardization from sitting to sitting. The following views are preferred (Figure 38–19): (1) frontal, (2) basal, (3) right lateral, (4) left lateral, and (5) right lateral, smiling.

Using single-lens reflex cameras, photographing with properly positioned dual electronic flash units and a pastel background complementary to skin tone (blue or green) is rapid and relatively simple. A fast, sharp lens of approximately 105 mm focal length avoids parallax distortion and provides precise uniformity. Commercial photographers are frequently used, but it is to the surgeon's advantage personally to master the details of precise photography. The advent of digital photography will provide significant recording and storage capabilities for nasal images.

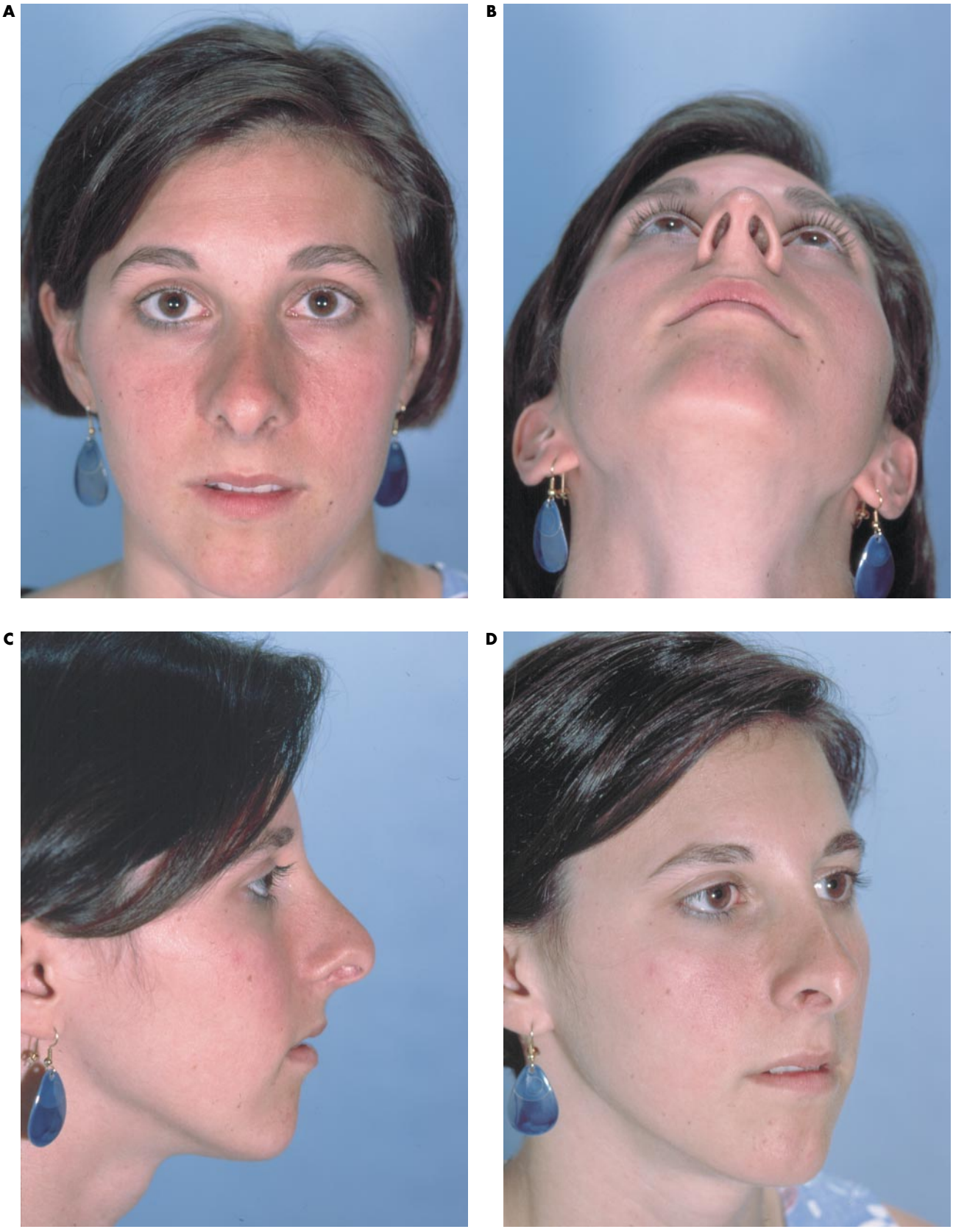


FIGURE 38–19. Typical preferred pre- and postoperative rhinoplasty photographic views (smiling view optional).

Classic Steps of Rhinoplasty Rhinoplasty remains the most challenging of all esthetic operations because no two procedures are ever identical. Each patient's nasal configuration and structure require individual and unique operative planning and surgical reconstruction. Therefore, no single technique, even though mastered, will prepare the surgeon for the varied anatomic patterns encountered. It is essential to regard rhinoplasty as one operation planned *to reconstitute and shape* the anatomic features of the nose into a new, more pleasing relationship with one another and the surrounding facial features. Rhinoplasty should be approached as an anatomic dissection of the nasal structures requiring alteration, conservatively shaping and repositioning these anatomic elements. Many more problems and complications arise from *overcorrection* of nasal abnormalities than from *conservative* correction. Inappropriate technique applied persistently without regard for existing anatomy creates frequent complications. One truism, that it is not what is removed in rhinoplasty that is important but what is left behind, remains valid. Furthermore, one must comprehend clearly the dynamic aspects of operative rhinoplasty because all surgical steps are interrelated and interdependent: *most maneuvers lead to a temporary deformity to be corrected by the steps that follow.*

Most corrections of the long nose associated with a hump or dorsal prominence follow the basic principles formulated by Jacques Joseph in the first half of the twentieth century. His monumental treatise, *Nasenplastic und Sonstige Gesichtplastik,* published in 1931, has proved its fundamental soundness for over half a century.

The procedure varies in each patient by the amount of tissue excised and/or the relative repositioning, reorientation, and/or grafting of anatomic structures. Classically, then, rhinoplasty consists of the following interrelated steps (Figure 38–20):

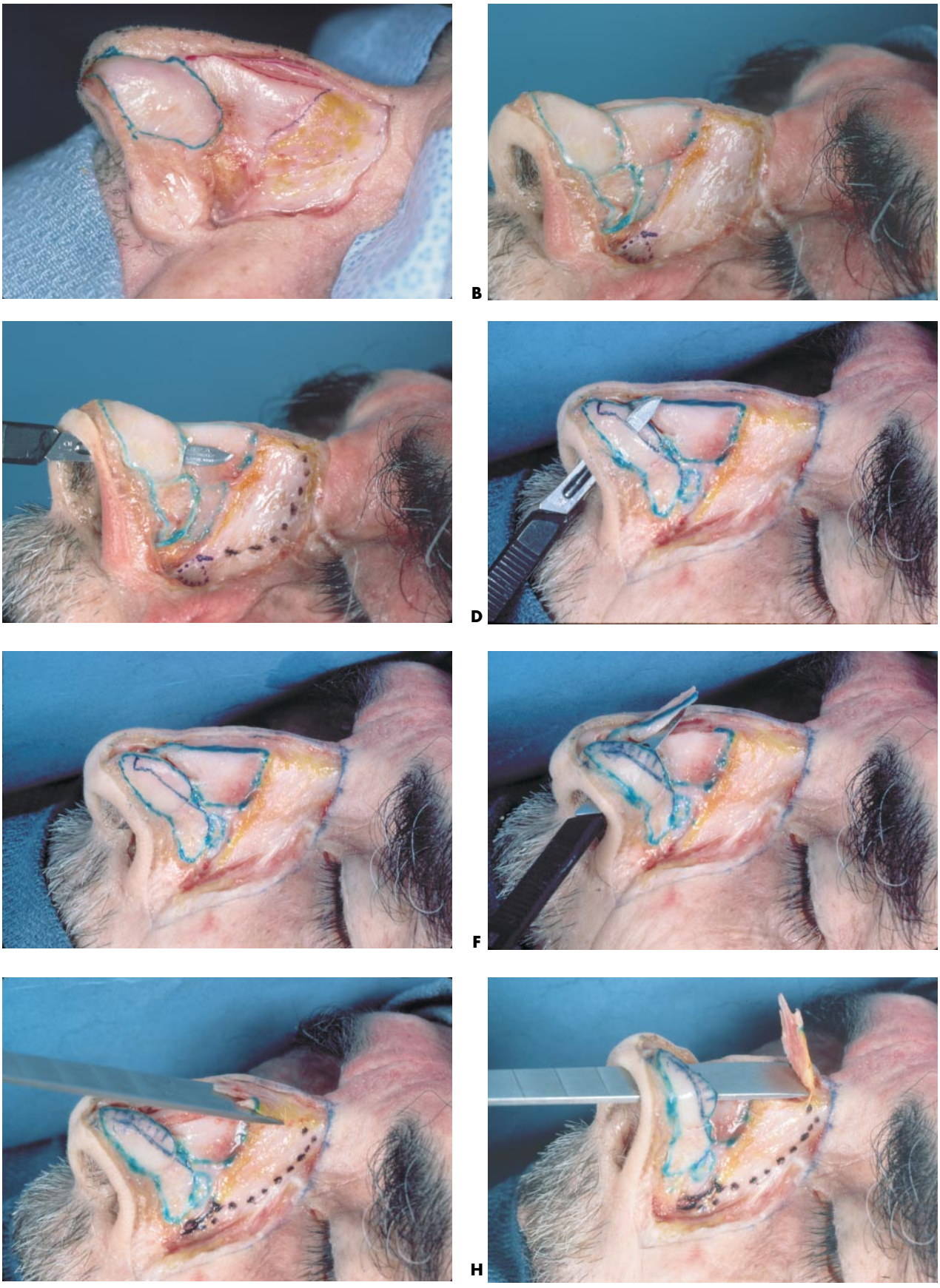
1. Septoplasty
2. Tip remodeling, projection, and cephalic rotation
3. Hump removal (establishing the profile line)
4. Narrowing of the nose with osteotomies
5. Final correction of subtle deformities

Preoperative Preparations Thorough, compulsive preparation by the surgeon before surgery ensures a streamlined operative procedure, designed to achieve esthetic satisfaction and avoid complications. The optimum time for operation in respect to the physiologic and psychological condition of the patient (and the surgeon) should be chosen. Patients are counseled that the procedure will be a meaningful and exciting event in their lives with long-range implications rather than a necessarily unpleasant experience.

A complete physical examination is indicated, including routine laboratory tests such as blood count, urinalysis, and chemical and coagulation profiles. Untoward bleeding histories should be comprehensively resolved before proceeding. Recent upper respiratory infections, skin pustules, or allergic exacerbations may be sufficient reasons to postpone surgery. Conditions such as anemia, metabolic disorders, or nutritional deficiency states contraindicate any elective procedure. A patient who is a poor anatomic surgical risk usually heals poorly and unpredictably. An exacting history is essential to determine if any medication currently taken by the patient could create bleeding or untoward drug interaction.

General planning of the operative sequence should be carried out before surgery, with a well-devised "game plan." Intraoperative improvisation may be necessary but only if complementary to a careful preoperative game plan. Photographic evaluation with the patient, clearly defining exact patient expectations and demonstrating realistic expectations as well as anatomic limitations, is mandatory before embarking on surgical repair. The patient must clearly comprehend the details, ramifications, limitations, and potential complications of the procedure. This informed consent discussion is purely a physician-patient encounter; no individual or list of written "helpful hints" supplants this critical part of the preoperative conditioning.

Preoperative Medication The cardinal rules in preoperative medication of patients are to use as few different categories of drugs as possible and preferably to employ drugs whose actions may be reversed or antagonized. Too often, patients scheduled for rhinoplasty are given a wide array of different drugs, which confuse the pharmacologic picture considerably—certainly, a therapeutic error. Reliance on predictable narcotic and phenothiazine and neuroleptic drugs only has been found safe and effective, particularly with a combination of local infiltration and intravenous analgesia and anesthesia, carefully


A
B
C
D
E
F
G
H

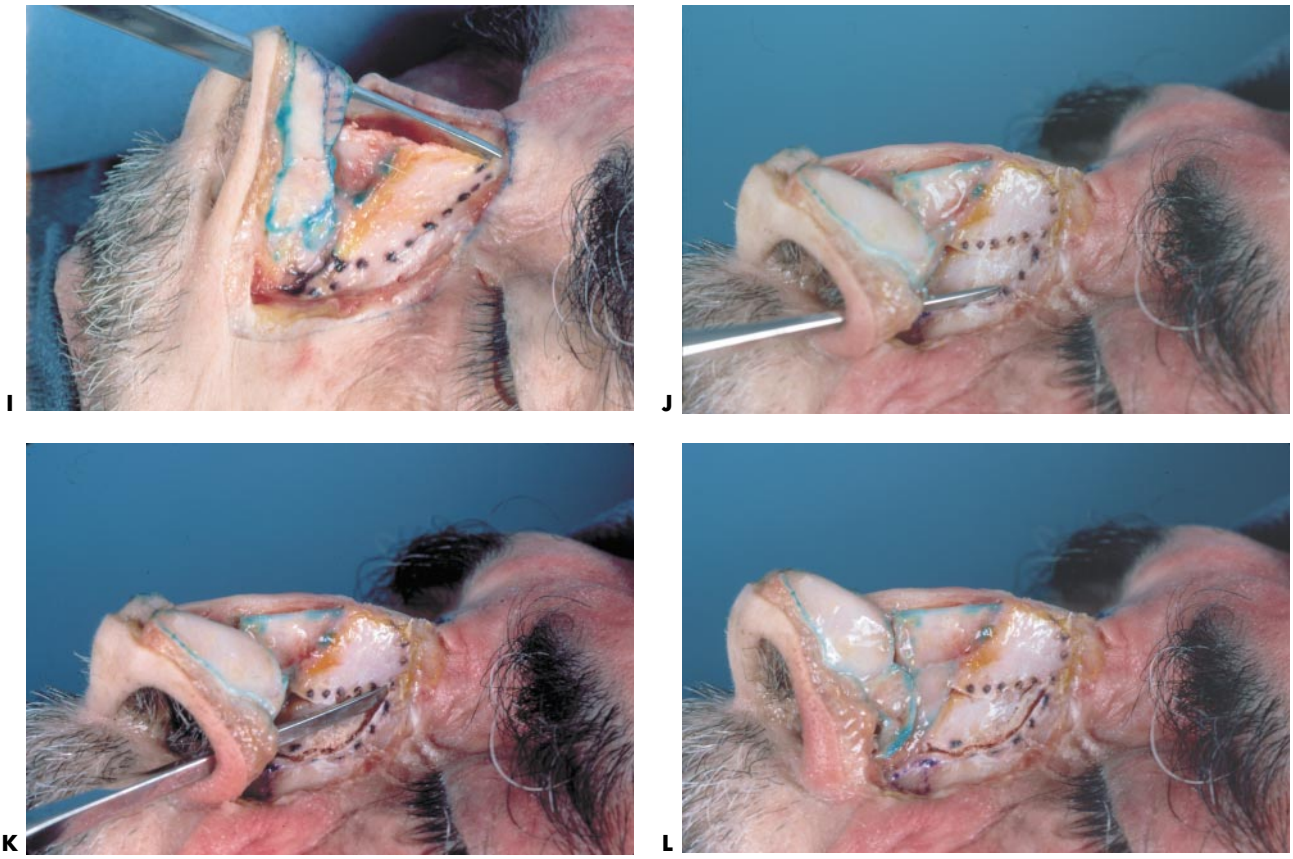


FIGURE 38–20. The classic steps of reduction rhinoplasty. *A*, Surgical anatomy of the nose. *B*, Variability of upper and lower lateral cartilages. *C*, Intercartilaginous incision. *D*, Alternative approach: transcartilaginous incision. *E*, Volume reduction of ala cartilage. *F*, Cartilaginous hump removal. *G*, Bony hump removal initiated. *H*, Bony hump removal completed. *I*, Medial oblique osteotomies. *J*, Low curved lateral osteotomy initiated. *K*, Completion of low curved lateral osteotomy. *L*, Infracture with nasal narrowing completed.

titrated, controlled, and monitored by a medical anesthesiologist. Individual drug titration in each patient, with the drug slowly infused in minimal intravenous increments, provides a safe and comfortable combination of euphoria, analgesia, and obtundency. General anesthesia may be used with equal facility.

One or two small gauze pads (neurosurgical cottonoids) moistened with 5% cocaine solution are placed in each nasal passage. Superb surface anesthesia with intense nasoconstriction and turbinate shrinkage results.

Minimal intranasal subcutaneous infiltration of standard 1% lidocaine (Xylocaine) with 1:100,000 epinephrine (freshly mixed) is conservatively accomplished. Infiltration occurs only as the No. 27 gauge needle is being *withdrawn* to avoid the possibility of direct vessel injection. The intent is to diffuse a thin plane of anesthesia and vasoconstrictive influence *in the favorable tissue plane just above the cartilage structure of the nose,* where most dissection will be performed. Small amounts of infiltration anesthesia are added at the alar margins, piriform aperture, columella, and nasal septum. No attempt is made to create a specific block anesthesia of specific sensory nerves. This method effectively anesthetizes the operative field, creates excellent vasoconstriction, and avoids distortion of the nasal structures (a common error). Seldom is more than 4 to 5 mL of anesthetic solution required. Overaggressive injection leads to tissue ballooning, feature distortion, and consequent impaired surgical judgment.

It is essential that 10 to 15 minutes be allowed to elapse between the completion of injection and the initiation of surgical steps, thereby ensuring *intense vasoconstriction,* minimal bleeding, unre-

stricted vision, and, consequently, improved surgical technique. The surgeon must use a strong head light to maximize visualization.

The patient cleanses his face and nose gently with hexachlorophene the morning of surgery, rinsing thoroughly. No disinfectant is applied to the nose or face at surgery. Vibrissae are trimmed only if excessively long.

Surgery of the Nasal Tip *Philosophy of Tip Sculpture* Sculpture of the nasal tip is regarded, and properly so, as the most exacting aspect of nasal plastic surgery. The surgeon is challenged by the essentially bilateral, animate, and mobile nasal anatomic components. Because no single surgical technique may be used successfully in correction of the endless anatomic tip variations encountered, the surgeon must analyze each anatomic situation and make a reasoned judgment about which approaches and tip modification techniques will result in a predictably natural appearance. Factored into this judgment decision must be consideration of, among other things, the strength, thickness, and attitude of the alar cartilages, the degree of tip projection, the tip skin and subcutaneous thickness, the columellar length, the length of the nose, the width of the tip, and the tip-lip angulation and relationship.

One fundamental principle of tip surgery is that normal or ideal anatomic features of the tip should be preserved and, if possible, remain undisturbed by surgical dissection, and abnormal features must be analyzed, exposed, reanalyzed, and corrected by reduction, augmentation, or shape modification.[3]

Surgeons have gradually come to understand that radical excision and extensive sacrifice of alar cartilage and other tip support mechanisms all too frequently result in eventual unnatural or "surgical" tips. What appears pleasant and natural in the early postoperative period may heal poorly because of overaggressive attempts to modify the anatomy more extensively than the tissues allow. Cross-cutting or morselization of the lateral crura may provide an excellent early appearance but commonly results in distortion or loss of tip support as the soft tissues "shrink-wrap" the weakened cartilages over time. Rhinoplasty is, after all, a compromise operation in which tissue sacrifices are made to achieve a more favorable appearance. It therefore becomes judicious to develop a reasoned, planned approach to the nasal tip based entirely on the anatomy encountered coupled with the final result intended. A philosophy of a *systematic incremental anatomic approach* to tip surgery is highly useful to achieve consistently natural results. Conservative reduction of the volume of the cephalic margin of the lateral crus, preserving a substantially complete, undisturbed strip of residual alar cartilage, is a preferred operation in individuals in whom nasal tip changes are intended to be modest. As the tip deformity or asymmetry encountered becomes more profound, more aggressive techniques are required, from weakened and complete strip techniques to significant interruption of the residual complete strip with profound alteration in the alar cartilage size, attitude, and anatomy. Cartilage structural grafts to influence the size, shape, projection, and support of the tip are often invaluable.

Tip sculpture cannot be successfully undertaken, let alone mastered, until the *major and minor tip support mechanisms* are appreciated, respected, and preserved or, when indicated, reconstructed (Table 38–1). Loss of tip support and projection in the postoperative healing period is one of the most common surgical errors in rhinoplasty. This tip

TABLE 38–1. Tip Support Mechanisms

Major
Size, shape, and resilience of the medial and lateral crura
Medial crural footplate attachment to the caudal border of the quadrangular cartilage
Attachment of the upper lateral cartilages (caudal border) to the alar cartilages (cephalic border)
Minor*
Ligamentous sling spanning the paired domes of the alar cartilages
Cartilaginous septal dorsum
Sesamoid complex extending the support of the lateral crura to the piriform aperture
Attachment of the alar cartilages to the overlying skin and musculature
Nasal spine
Membranous septum

*On occasion, because of extreme anatomic variability, a "minor" tip support may assume the importance of one of the more major supports.

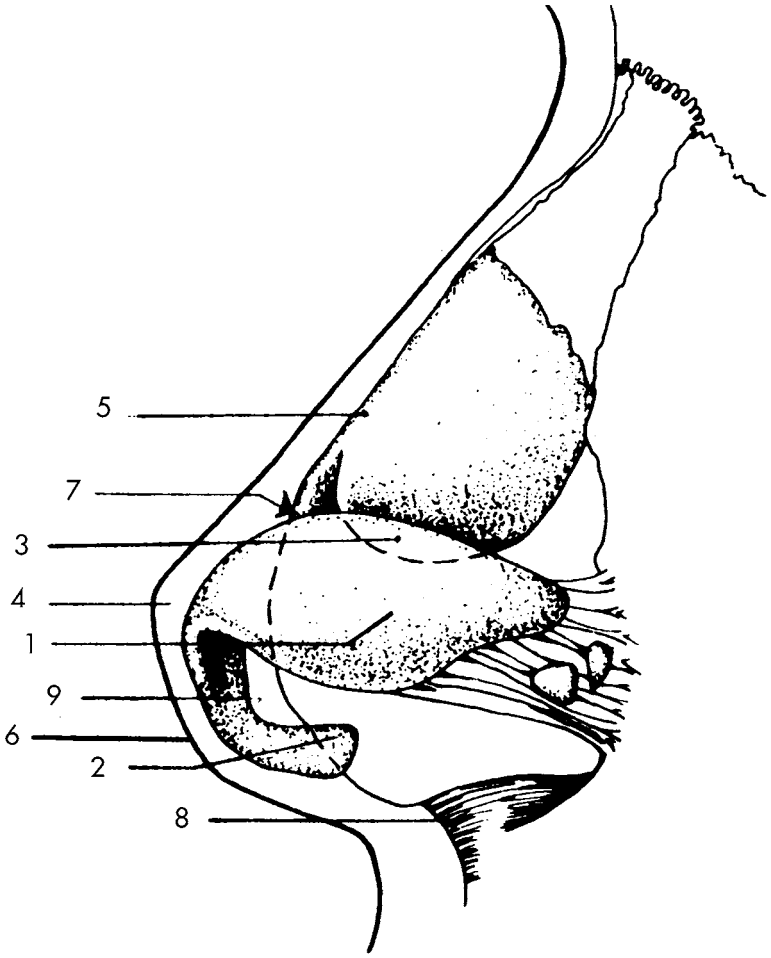


FIGURE 38–21. The *major* tip support mechanism of the nose consists of (1) the size, shape, and resiliency of the medial and lateral crura; (2) the wrap-around attachment of the medial crural footplates to the caudal end of the quadrilateral cartilage; and (3) the soft tissue attachments of the caudal margin of the upper lateral cartilage to the cephalic margin of the alar cartilage. The *minor* tip support mechanisms, which may in certain anatomic configurations assume major support importance, include (5) the dorsal cartilaginous septum, (4) the interdomal ligament, (9) the membranous septum, (8) the nasal spine, (6) the surrounding skin and soft tissues, and (7) the anterior septal angle.

"ptosis" is usually the inevitable result of the sacrifice of nasal tip support mechanisms.

In the majority of patients, the major tip support mechanisms (Figure 38–21) consist of (1) the size, shape, and resiliency of the medial and lateral crura; (2) the wrap-around attachment of the medial crural footplates to the caudal end of the quadrilateral cartilage; and (3) the soft tissue attachment of the caudal margin of the upper lateral cartilage to the cephalic margin of the alar cartilage. Compensatory re-establishment of major tip support should be considered if, during the operation, any or all of these major tip support mechanisms are compromised in any fashion.

The minor tip support mechanisms that, in certain anatomic configurations, may assume major support importance include (1) the dorsal cartilaginous septum, (2) the interdomal ligament, (3) the membranous septum, (4) the nasal spine, (5) the surrounding skin and soft tissues, and (6) the alar sidewalls.

Tip projection in every rhinoplasty operation is inevitably enhanced, reduced, or preserved in its original state. Anatomic situations in which each of these outcomes is desirable and intended are regularly encountered in a diverse rhinoplasty practice. The desirable surgical goal in every operation is preservation of the projection already existent if, as is true in the majority of rhinoplasty patients, preoperative projection of the tip is satisfactory. Other patients require an increase in the projection of the tip relative to the intended new profile line. A predictable variety of reliable operative methods exists for creating or augmenting tip projection; they are discussed later in this chapter. Finally, in a limited but clearly definable group of patients with overprojecting tips, a calculated intentional reduction of excessive tip projection is desirable to effect intentional retroprojection.

Successfully achieving these diverse surgical results requires an understanding of and a healthy respect for the major and minor tip support mechanisms, seasoned by the recognition of the intraopera-

tive surgical tip dynamic principles that interact in every tip operation. It clearly follows that the *appropriate tip incisions and approaches should be planned to preserve as many tip supports as possible.* Alar cartilage sculpturing should similarly respect this principle by conserving the volume and integrity of the lateral crus and avoiding, in all but the most extreme anatomic situation, radical excision and sacrifice of tip cartilage.

The surgeon should differentiate clearly among *incisions, approaches*, and *techniques*. Incisions are simply methods of gaining access to the underlying supportive structures of the nose and by themselves have little importance. Approaches to the nasal tip provide important exposure to the skeletal structures and consist of procedures to deliver the tip cartilages or to avoid complete delivery, or to operate on the alar cartilages without removing them from their anatomic beds. Sculpturing techniques are defined as surgical modifications: excision, reconstruction, or orientation of the alar cartilages calculated to cause significant changes in the definition, size, orientation, and projection of the nasal tip. Because of the amazing complexity of anatomic configurations encountered in nasal tip surgery, further modifications are frequently used to ensure stable refinements.

It is important to assess several factors before selecting the appropriate tip procedure. In planning tip remodeling, the surgeon must determine whether the tip requires (1) a reduction in the *volume* of the alar cartilages, (2) a change in the *attitude* and *orientation* of the alar cartilages, (3) a change in the *projection* of the tip, (4) a cephalic *rotation* with a subsequent increase in the columellar inclination (nasolabial angle), (5) a bilateral narrowing of the angle of the domes, and (6) reduction of the interdomal distance.

Ideally, conservative reduction of the volume of the cephalic margin of the lateral crus, preserving the majority of the crus while maintaining a complete (uninterrupted) strip of alar cartilage, is preferred. This procedure is satisfactory and appropriately safe when minimal conservational tip refinement and rotation are required. As the tip deformity increases in size and complexity, more aggressive techniques are required. A philosophy of a graduated incremental anatomic approach to nasal tip surgery has proved useful. This implies that no routine tip procedure is ever used; instead, the *appropriate incisions, approaches, and tip sculpturing techniques are selected based entirely on an analysis of the varying anatomy encountered.* Whenever possible, a complete strip operation is used, reserving more risky interrupted strip techniques for anatomic situations in which more profound refinement changes and significant rotation are desirable.

Surgical Approaches to the Tip. Nondelivery Approaches. In anatomic situations in which the nasal tip anatomy is favorable, only conservative refinements are necessary, and nondelivery approaches possess great value. Less dissection and less disturbance of the tip anatomy are necessary, and this reduces the chance for asymmetry, error, and unfavorable healing. Properly executed (when indicated), nondelivery approaches therefore allow the surgeon to control the healing process more accurately than when more radical approaches and techniques are chosen.

The transcartilaginous approach is preferred because of its simplicity and ease of use (Figure 38–22). The same tip refinements, however, may be accomplished through the retrograde approach. These approaches are chosen in patients whose tip anatomy is fundamentally satisfactory, requiring only volume reduction to accomplish a thinning sculpture of the cephalic or medial margin of the lateral crus. When tip projection is to be enhanced by the use of cartilage tip grafts, nondelivery approaches are preferred because precise recipient pockets may be more accurately created in the infratip lobule undisturbed by the minimal dissection inherent in nondelivery approaches. If sutured-in-place tip grafts are planned, a delivery or open approach is preferred.

Delivery Approaches. Delivering the alar cartilages as individual bipedicle chondrocutaneous flaps through intercartilaginous and marginal incisions is the preferred approach when the nasal tip anatomy is more abnormal (broad, asymmetric, etc) or when more dramatic tip refinements are necessary. Significant modifications in the alar cartilage shape, attitude, and orientation are more predictably attained when the cartilages are delivered (Figure 38–23). The basal photograph is usually helpful in determining which patients may best be approached in this manner. If the triangularity of the tip from below is satisfactory and only modest volume reduction of the lateral crus appears necessary, the nondelivery approach serves well. If, however, on basal and frontal views, the alar cartilages flare unpleasantly (Figure 38–24), tip triangularity

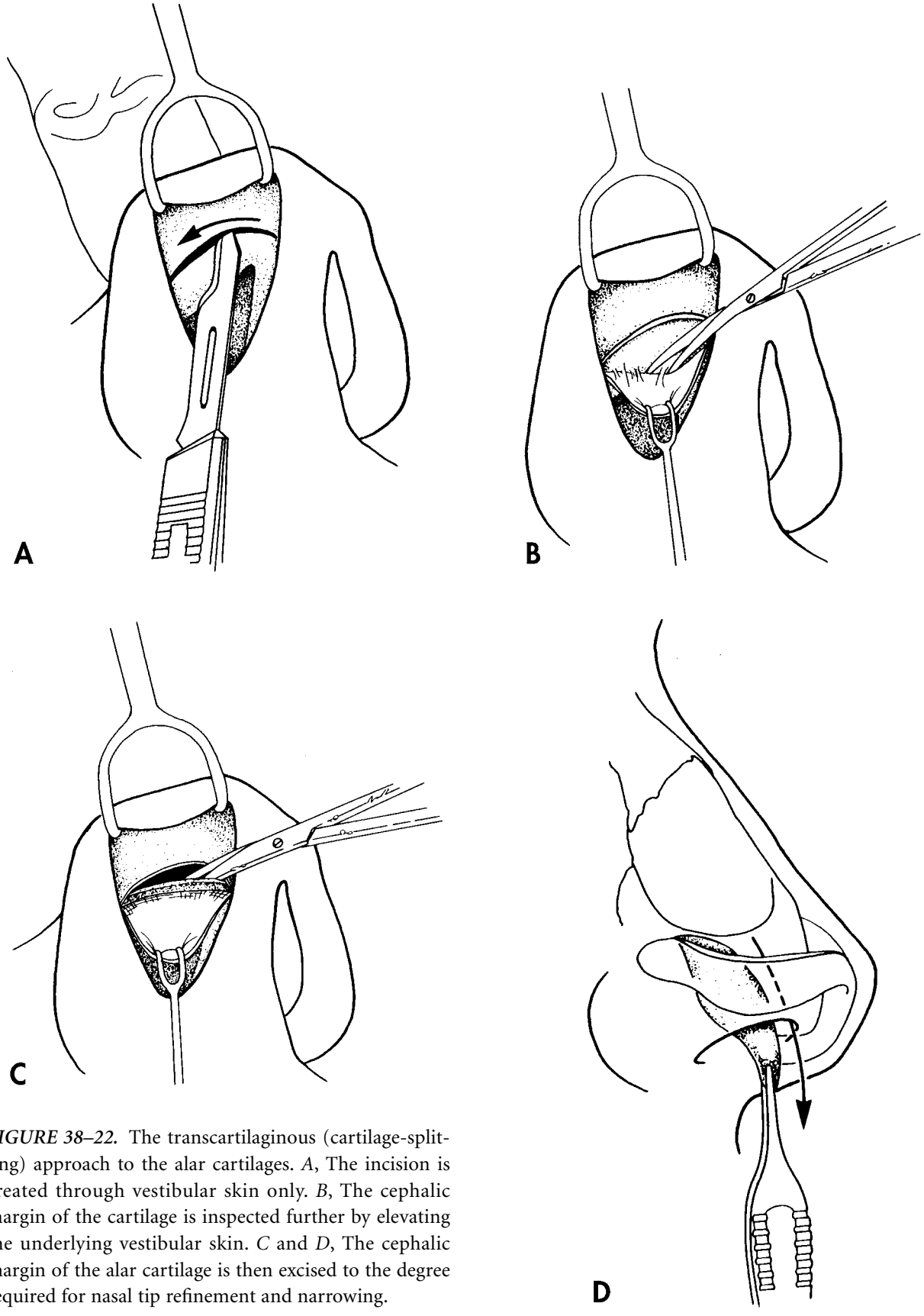


FIGURE 38–22. The transcartilaginous (cartilage-splitting) approach to the alar cartilages. *A*, The incision is created through vestibular skin only. *B*, The cephalic margin of the cartilage is inspected further by elevating the underlying vestibular skin. *C* and *D*, The cephalic margin of the alar cartilage is then excised to the degree required for nasal tip refinement and narrowing.

FIGURE 38–23. *A*, The delivery approach to the nasal tip (*B* and *C*). Intercartilaginous (*B*) and marginal (*C*) incisions are used on either side to free and deliver the alar cartilage for evaluation and refinement. *D*, Delivery of the lower lateral cartilage for analysis, inspection, and reconstruction.

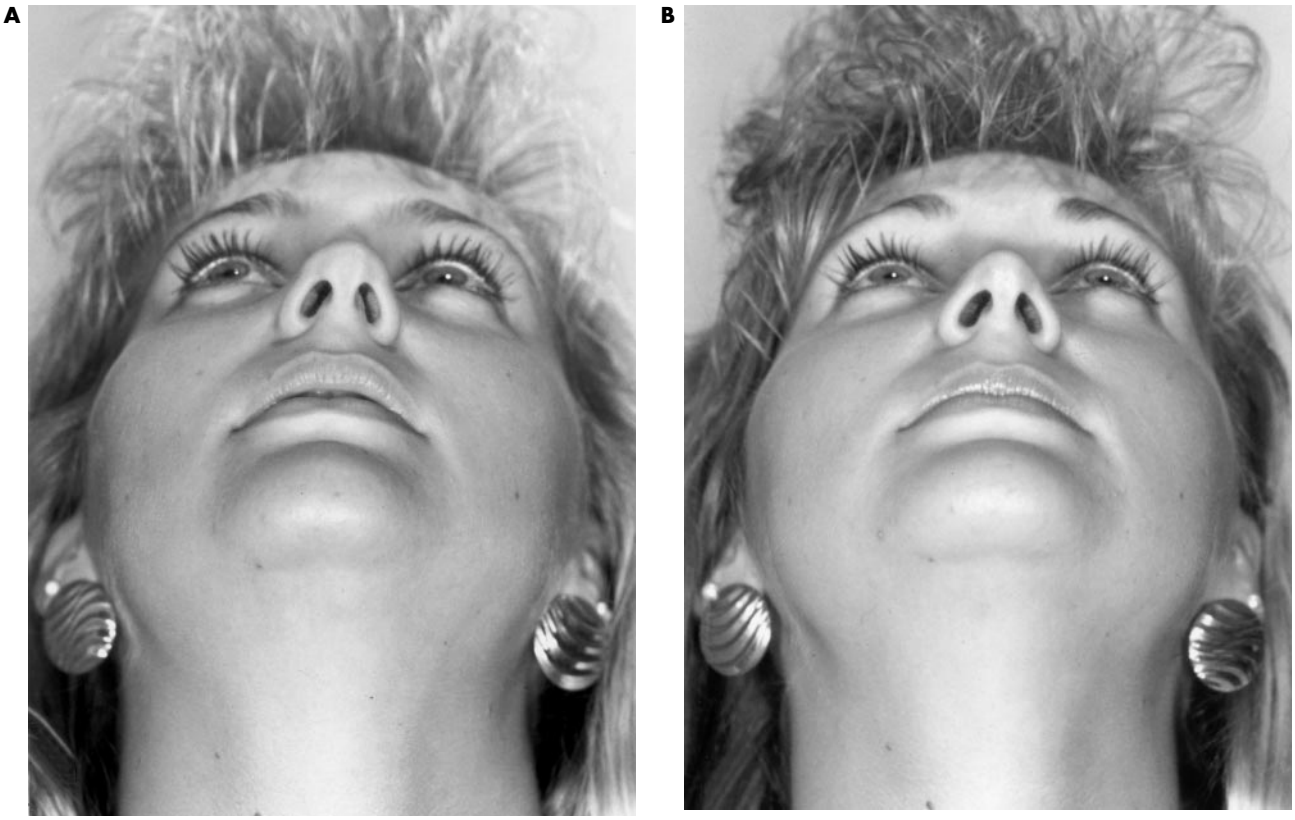


FIGURE 38–24. *A*, Basal view of a patient demonstrating a square- or trapezoid-shaped amorphous nasal tip, in which the delivery approach is preferred to refine cartilages and suture them together with a transdomal suture to achieve favorable triangularity. *B*, Improved basal view appearance 1 year after surgery.

is unsatisfactory or the tip appears too amorphous and bulbous, the domal angles are too wide, and the interdomal distance is excessive, a delivery approach is chosen to correct these esthetic deficiencies more thoroughly. Transdomal suture narrowing of broad domes (Figure 38–25), an effective and preferred technique, is effected by means of the delivery approach. In similar fashion, interrupted strip techniques (rarely employed) for more radical tip refinement and cephalic rotation are more efficiently accomplished when the cartilages are delivered. The increased surgical exposure provides the surgeon with an improved binocular view of the tip anatomy and affords the added ease of bimanual surgical modifications.

Open (External) Approach. The external or open approach to the nasal tip is in reality a more aggressive form of the delivery approach and is chosen with discretion in specific nasal tip deformities (Figure 38–26). When the nasal tip is highly asymmetric, markedly overprojected, severely underprojected, or anatomically confusing in its form (as in certain secondary revision cases), the open approach is considered. The transcolumellar scar is of negligible importance in this decision because it routinely

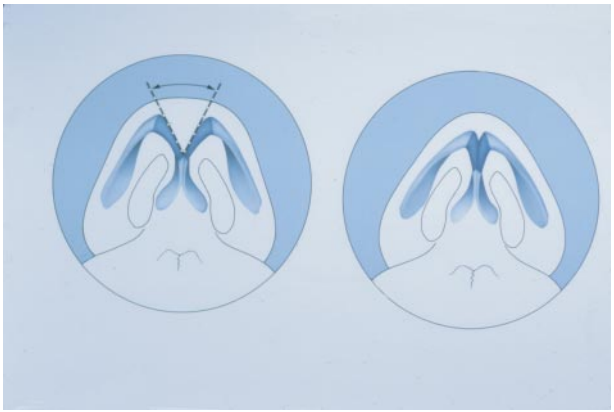


FIGURE 38–25. The triangular effect achieved by transdomal suturing of the alar cartilage domes, narrowing both the interdomal space and the bilateral domal angle.

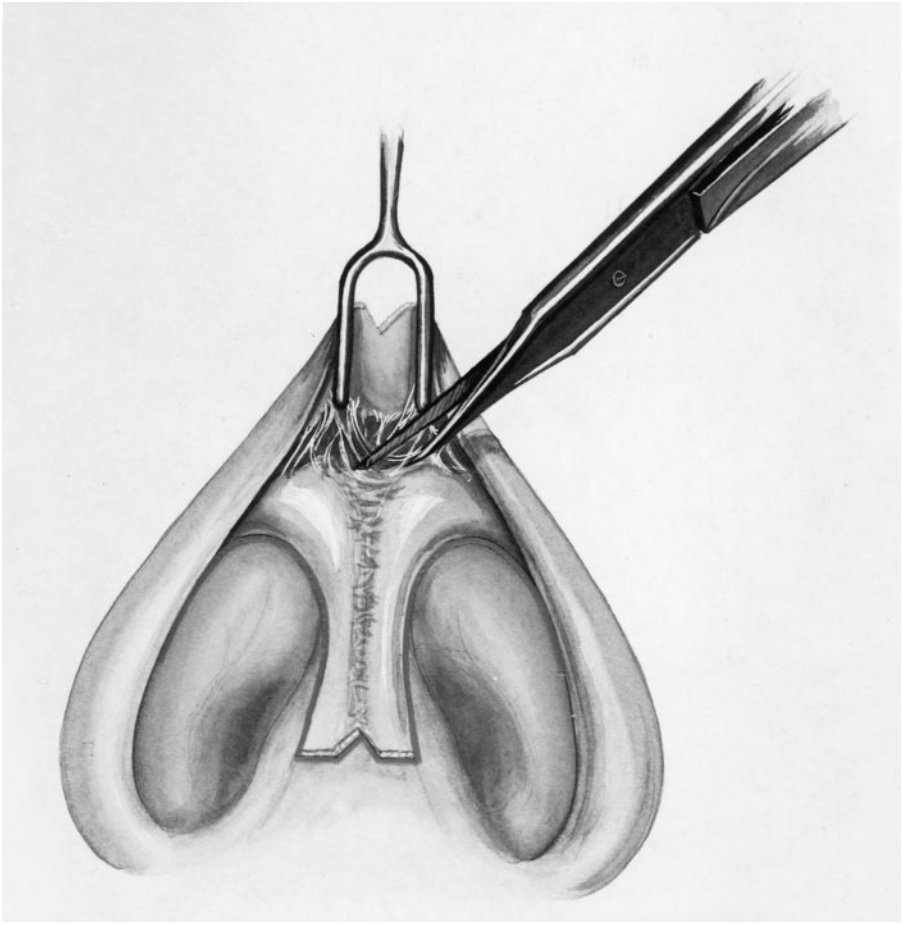


FIGURE 38–26. The open approach to the nasal tip, employing an irregularized incision across the columellar.

heals inconspicuously when meticulously repaired. The anatomic view is unparalleled through this approach, affording the surgeon diagnostic information unavailable through traditional closed approaches. These technical virtues must be balanced with the potential disadvantages of an enlarged scar bed, slightly delayed healing with some prolongation of tip edema, and an increased operating time. Clearly, when subtle and conservative tip surgery is indicated by the patient's existent anatomy, the open approach is unnecessary.

Alar Cartilage Sculpturing Techniques. The choice of the technique used to modify the alar cartilages and the relationship of the nasal tip with the remaining nasal structures should be *based entirely on the anatomy encountered* and the predicted result desired, as defined from the known dynamics of long-term healing. The astounding diversity of tip anatomic situations encountered demands a broad diversification of surgical planning and execution by the experienced surgeon.

Three broad categories of nasal tip sculpturing procedures may be identified. Although additional subtle technical variations exist, the three primary categories are (1) volume reduction with residual complete strip, (2) volume reduction with suture reorientation of the residual complete strip, and (3) volume reduction with interrupted strip.

Preserving intact the major portion of the *residual complete strip* of the alar cartilage is always preferred when the anatomy of the alar cartilages and their surrounding soft tissue investments allows. This preservative approach retains the supportive advantage of the intact cartilage strip (thus "mimicking" nature), discourages cephalic rotation when it is undesirable, eliminates many of the potential hazards of more radical techniques, and tends to produce a more natural final result (Figures 38–27 and 38–28).

Techniques involving a weakened (or suture-reoriented) residual complete strip have all the foregoing positive virtues and allow the surgeon to effect reorientation of the breadth of the domal angle, projection modification, and narrowing refinement, so

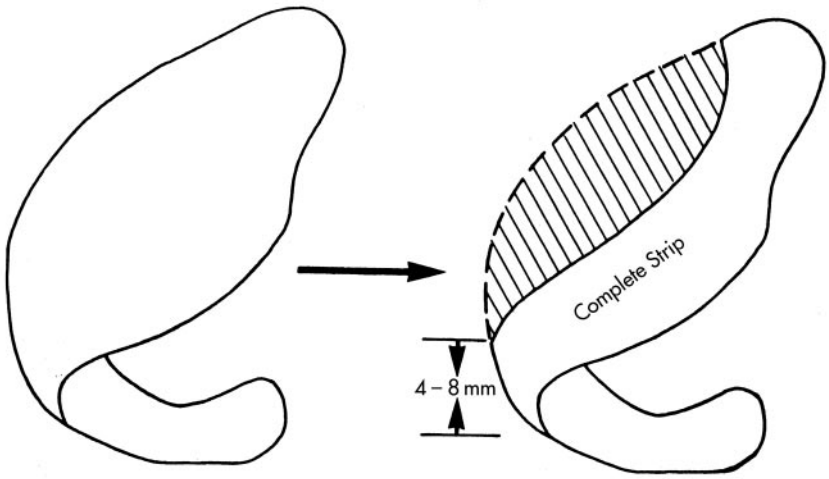


FIGURE 38–27. The minimum amount of complete strip that should remain after volume reduction of the alar cartilage.

desirable in the ideal postoperative appearance (Figures 38–29 and 38–30). The control of favorable healing is enhanced with these techniques, with the risk of complication diminished considerably.

Despite a laudable desire to preserve the integrity of the residual complete strip whenever possible, anatomic situations are occasionally encountered in which the shape, breadth, and orientation of the alar cartilages must be changed more radically by interrupting the complete strip in a vertical fashion somewhere along its extent to refine severe anatomic deficits[9] (Figures 38–31 and 38–32). When significant cephalic rotation is indicated, interrupted strip techniques are considered. The risks of asymmetric healing are higher when the alar cartilages are divided, however, and initial loss of tip support occurs immediately. The latter problem must be recognized and countermeasures taken during surgery to ensure that sufficient tip support is reconstituted. Shoring struts in the columella,

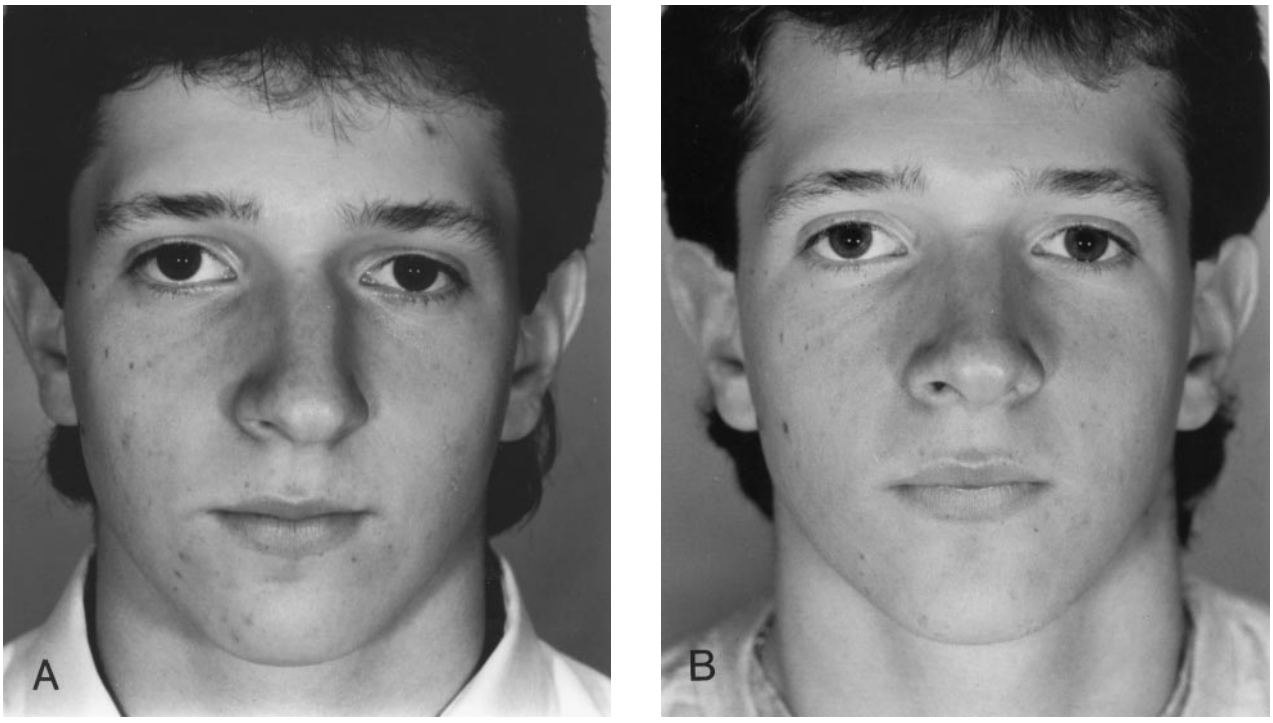


FIGURE 38–28. Patient before (*A*) and 1 year after (*B*) rhinoplasty using the transcartilaginous approach, preserving a generous complete strip of alar cartilage.

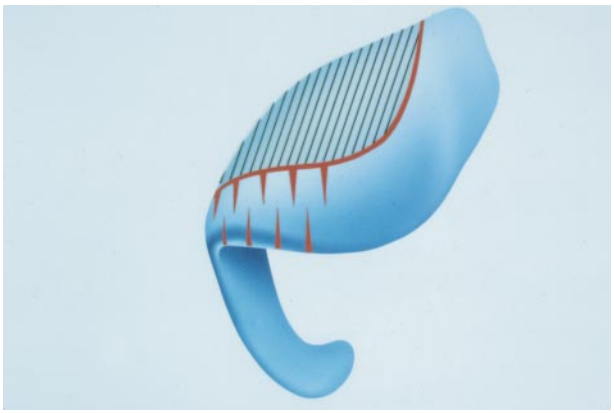


FIGURE 38–29. Weakening of residual complete strip by incomplete, noncoalescent incisions through the lateral crus. Although this technique helps to narrow bulky tips, less control over long-term favorable healing is realized.

infratip lobule cartilage grafts, and transdomal suturing are the most commonly used tip support adjuncts. Almost without exception, interrupted strips should be avoided in patients displaying thin skin with sparse subcutaneous tissue.

Tip Projection and Cartilage Tip Grafts. In addition to the creation of narrowing refinement and symmetry of the nasal tip, most evident in the frontal view, *appropriate projection* must be preserved or newly created to result in the most natural appearance possible. Clearly, the most attractive and elegant noses are those in which anterior projection is sufficient to set the tip subtly but distinctly apart from the nasal supratip areas. Ptotic or poorly projected tips produce a snubbed and indistinct appearance.[10]

Ideally on profile view, the nasal tip should be slightly elevated above the cartilaginous dorsum by 1 to 2 mm, blending gently rather than abruptly into the supratip (Figure 38–33). If the preoperative projection of the tip is normal and adequate, lowering the cartilaginous dorsum into proper alignment will achieve a satisfactory esthetic appearance provided that no loss of tip support occurs during the operative or postoperative periods. Preserving the major and minor tip support structures increases this likelihood, whereas their sacrifice without compensatory re-establishment of support inevitably leads to eventual tip ptosis. If preoperative tip projection is

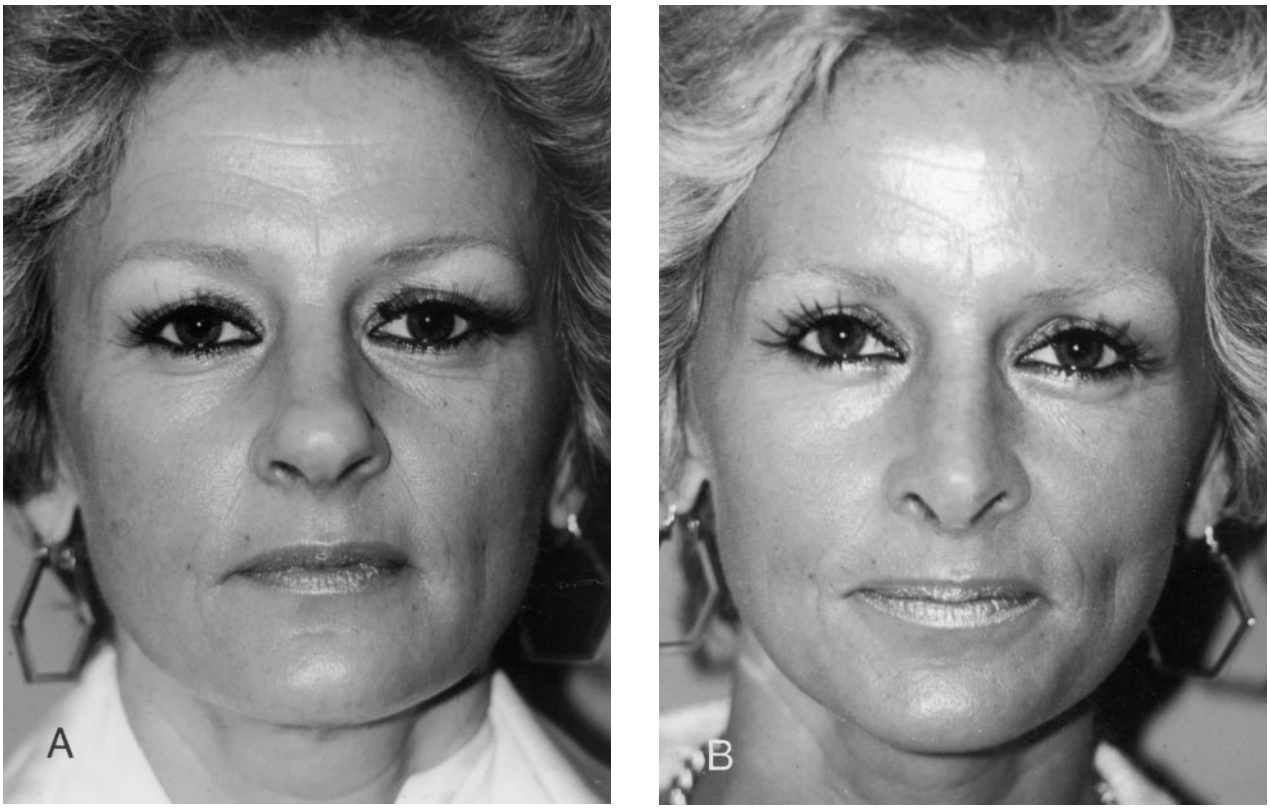


FIGURE 38–30. Patient before (*A*) and 1 year after (*B*) surgery. The nasal tip has been refined using a weakened complete strip technique with a transdomal suture repair created through a delivery approach.

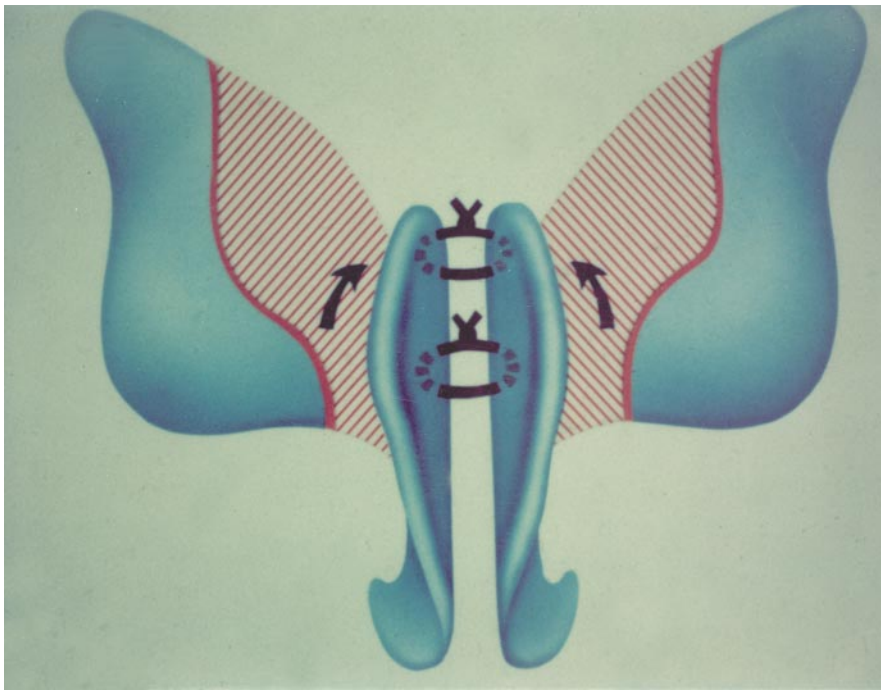


FIGURE 38–31. Interruption of the lower lateral cartilage at the domal angles, creating an interrupted complete strip.

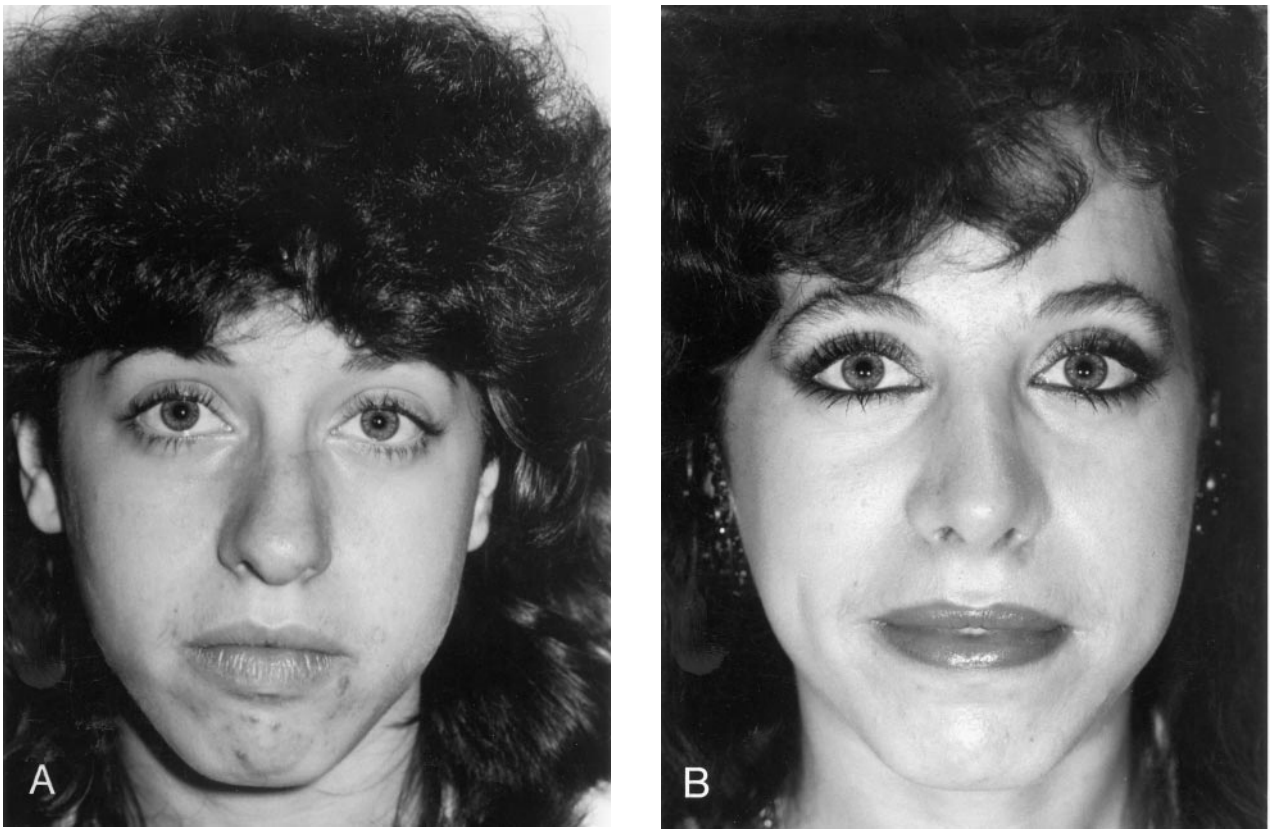


FIGURE 38–32. Patient before (*A*) and 1 year after (*B*) rhinoplasty using a delivery approach and interrupted-strip technique for refinement of the nasal tip. Heavy bulbous cartilages accompanied by thick skin and subcutaneous tissues dictated the need for this more radical technique.

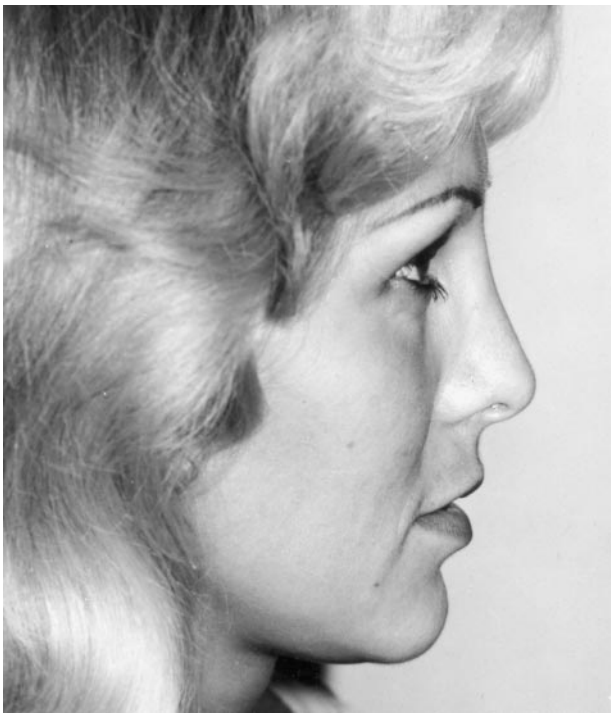


FIGURE 38–33. An ideal nasal profile in a patient in whom the nasal tip leads the supratip cartilaginous dorsum by 1 to 2 mm.

inadequate, attempts to over-reduce the supratip cartilaginous dorsum to produce pseudoprojection of the tip are inadvisable and lead to apparent flattening or widening of the middle third of the nose.

If tip projection is inadequate, several reliable methods may be used singly or in tandem to establish permanent improvement. All involve reorientation of the alar cartilages or addition of autogenous cartilage grafts to strengthen or sculpt the projection and/or attitude of the tip and infratip lobule.

Because the long-term viability and stability of sutured-in-place cartilage tip grafts are well established, they are regularly used with success if the surgical modification of existing alar cartilage configuration is not adequate to produce the desired degree of projection.[2,3] In revision rhinoplasty in particular, cartilage tip grafts are irreplaceable in skeletal reconstruction beneath scarred skin and asymmetric topography.

Two distinct varieties of tip grafts are preferred: those that directly overlie the dome profile of the alar cartilage and those that redefine and contour the skeletal anatomy of the infratip lobule (Figure 38–34). Because these grafts (single or laminated) lie in intimate subcutaneous pockets, exacting sculpture of their size and shape is mandatory. Harvested from septal or auricular cartilage, they are ideally inserted with or without suture fixation into small pockets dissected to accommodate exactly the dimensions of the graft(s). Bilateral marginal incisions beneath the anatomic dome area facilitate the careful pocket creation and render final positioning and stabilization of the graft easier than if only one unilateral incision is used. Suture fixation to the alar cartilages may be necessary if the tip region has been widely developed in a primary delivery or open approach method, disallowing the creation of a stable, limited, defined pocket. The edges of the grafts must be beveled or softened to avoid visible contour irregularities or to offset deformities. Carved in triangular, trapezoid, or shield-like shapes, tip grafts may accentuate favorable tip-defining points and highlights while imparting a more natural appearance to tips with congenital or postsurgical inadequacies (Figure 38–35).

If additional projection is required, it may be achieved in a variety of ways. Autogenous *cartilage struts* positioned below and/or between medial crura are effective in establishing permanent tip stability

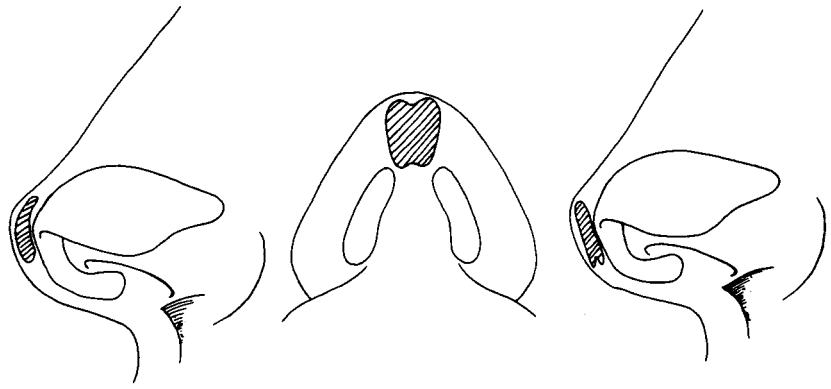


FIGURE 38–34. Cartilage grafts positioned in the nasal tip or infratip lobule serve well to contour and create ideal projection.

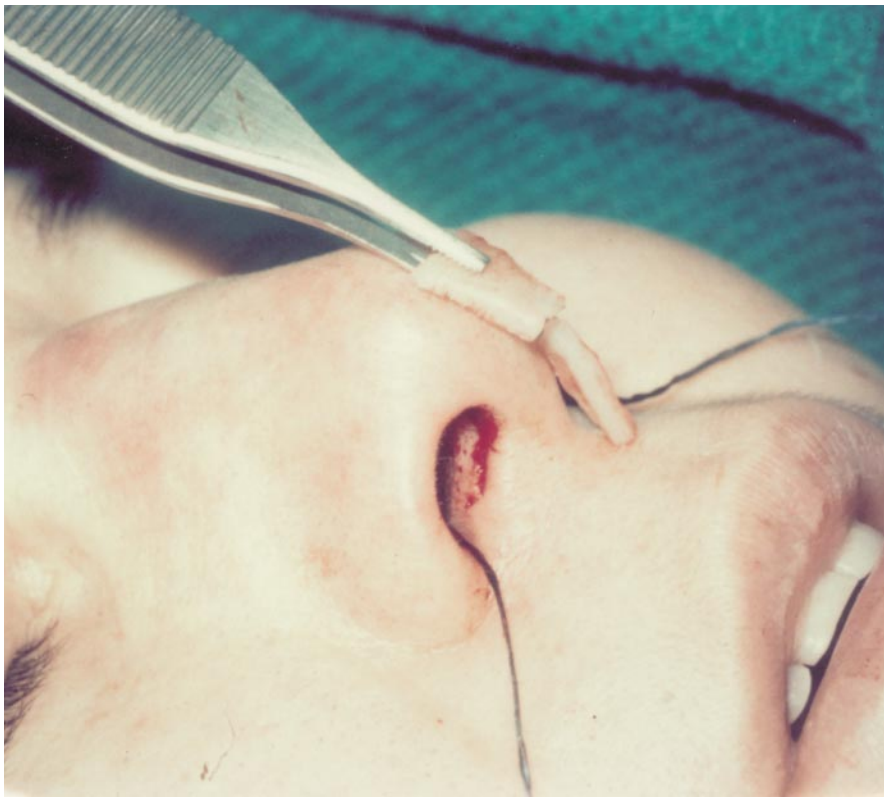
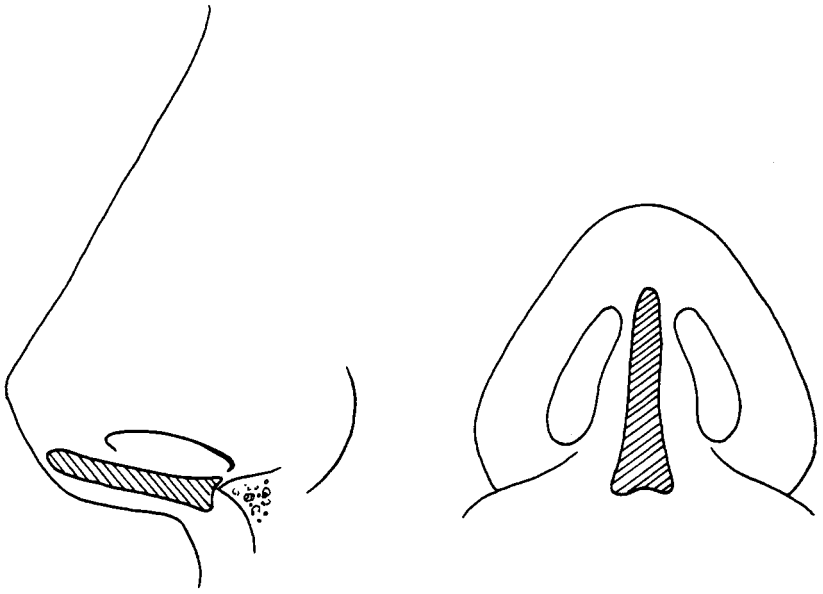


FIGURE 38–35. Sculptured tip graft and columellar strut (from nasal septal cartilage) designed to support and project the nasal tip.

and slight additional projection (Figure 38–36). *Plumping grafts* of cartilage fragments, introduced into the base of the columella through a low lateral columellar incision, provide an additional platform for the tip projection resulting from the strut. Cartilage struts should be shaped with a gentle curve to match the anatomy of the curved columella, at times aiding in the creation of a distinct "double break," but should never extend to the apex of the tip skin lest a tent-pole appearance develops. If the medial crural footplates diverge in a widely splayed fashion, further tip projection may be gained by resecting excessive intercrural soft tissue and suturing the medial crura together.

FIGURE 38–36. Autogenous cartilage struts may be positioned in precise pockets or sutured into position between the medial crura.



Tip Rotation. In many patients undergoing rhinoplasty, *cephalic rotation* of the nasal tip complex (alar cartilages, columella, and nasal base) assumes major importance in the surgical event, whereas in other individuals, the *prevention* of upward rotation is vital. Certain well-defined and reliable principles may be invoked by the nasal surgeon essentially to calibrate the degree of tip rotation (or prevention thereof). The dynamics of healing play a critical role in tip rotation principles; the control of these postoperative healing changes distinguishes rhinoplasty from less elegant procedures. In the past, over-rotation of the nasal tip created an unhealthy stigma for the rhinoplasty procedure. Most individuals recognize and prefer the esthetic advantages of the stronger nose possessed of sufficient length to impart character and suitable proportions to the face.

The planned degree of tip rotation depends on a variety of factors, which often include (1) the length of the nose, (2) the length of the face, (3) the length of the upper lip, (4) facial balance and proportions, (5) the patient's esthetic desires, and (6) the surgeon's esthetic judgment.

An important distinction must be drawn between *tip rotation* and *tip projection*. Although certain tip rotation techniques may result in desirable increases in tip projection, the converse is not true. Tip rotation and projection, in fact, complement each other, and their proper achievement in individual patients is constantly interrelated. A classic example of this interdependent relationship is illustrated by the almost inevitable loss of tip projection when interrupted strip techniques are chosen to enhance cephalic rotation; steps must be planned to restore adequate long-term tip projection by one of the several methods recommended.

Finally, a distinction must be drawn between true tip rotation and the *illusion* of tip rotation achieved by contouring cartilage grafts placed in the infratip lobule, columella, and nasolabial angle. Favorable modifications in the tip-lip complex profile areas with autogenous implants may obviate the need for any actual tip rotation, thus preserving a long, and at times more desirable, nasal appearance. Reduction of the nasal profile, particularly the supratip cartilaginous pyramid, may also impart the illusion of rotation and a shortened nose, although occasionally at the expense of a strong and narrow dorsum.

Nasal tip rotation results fundamentally from planned surgical modifications of the alar cartilages, but increments of rotation may also be realized from additional adjunctive procedures on nasal structures adjacent to the alar cartilages, which exert a favorable influence on calibrated tip rotation methods. *Shortening of the caudal edge of the septum, excision of the caudal margins of overlong upper lateral cartilages, and septal shortening with a high transfixion incision* are regularly used to enhance the effects of a planned degree of tip rotation.[11]

Because tip rotation is only one of the many objectives of rhinoplasty, decisions regarding rotation must be interrelated with planning for tip volume reduction, alar cartilage thinning reduction, and modifications in the attitude and angulation of the alar cartilages.

The techniques and healing dynamics described are not absolute but are reasonably predictable. Most tip rotation techniques may be incorporated in an organizational scheme that incorporates three procedures preserving a complete, intact strip of alar cartilage (Figure 38–37) and three additional procedures involving interrupted strip techniques (Figure 38–38). Unique anatomic situations are regularly encountered that require modifications of this scheme to achieve a more refined result, but the fundamental principles elaborated remain constant. In addition, the thickness and strength of the alar cartilages, along with the character of their enveloping soft tissue and skin, dictate, to a degree, which techniques may safely and predictably be used in each anatomic situation.

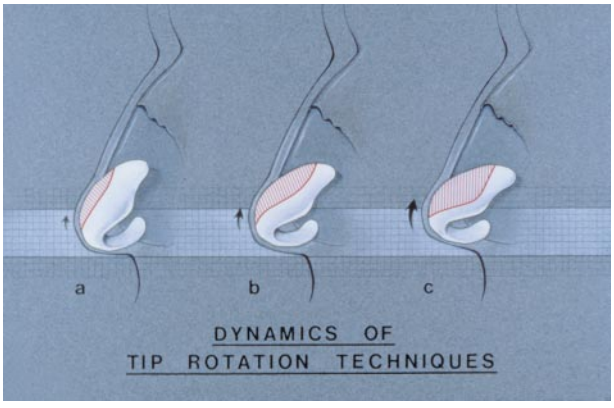


FIGURE 38–37. A to *C*, Preserving an intact strip of lower lateral cartilage tends to resist upward rotation. The greater the tissue void created, the greater is the tendency of scar tissue contracture to result in slight upward rotation as the tissue void is closed.

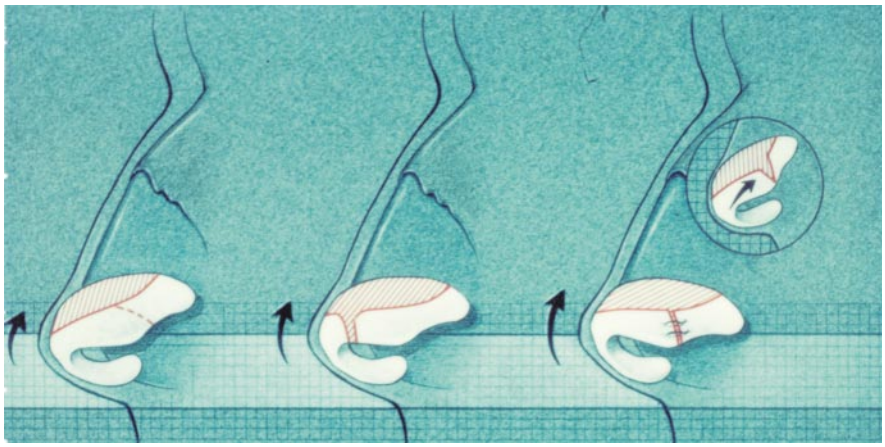


FIGURE 38–38. Interrupted strip techniques tend to foster cephalic rotation by weakening the strength of the residual cartilage strip, allowing scar contracture to pull the tip cephalically.

Complete strip techniques are always preferable tip procedures when the nasal anatomy permits, and the goals of the surgical procedure may be met without resorting to the less predictable interrupted strip procedures. Preserving a complete, uninterrupted segment of alar cartilage remnant contributes to a more stable and better supported nasal tip that tends to *resist* cephalic rotation during healing.

Interrupted strip techniques, combined with volume reduction of excessive alar cartilage, tend to result in a more substantial degree of cephalic rotation of the tip complex. Once the complete strip of residual alar cartilage is divided (interrupted), the result is a relative instability of the nasal tip, on which the forces of upward scar contraction create a variable degree of cephalic rotation, underscoring the principle that during scar contracture, tissues are generally moved from areas of instability (in this case, the unstable nasal tip cartilages) toward areas of stability (the bony-cartilaginous nasal pyramid).

Tip Rotation with Complete Strip Techniques. Volume reduction of the alar cartilage causes a tissue deficit of minimal, moderate, or maximal proportions, depending on the degree of cartilage removal indicated or desirable. Essentially, no cephalic tip rotation results from minimal volume reduction alone, whereas the greater tissue void resulting from moderate to maximal volume reduction tends to create progressively greater degrees of minimal tip rotation. Indeed, preservation of the complete strip is regularly indicated and preferred to *resist* the forces of upward rotation when the preoperative nasal length is to be maintained.

Substantial planned tip rotation when complete strip techniques are used depends therefore on the addition of adjunctive procedures to achieve cephalic elevation of the tip such as septal shortening with a high septal transfixion or on designing illusions of tip rotation with a high septal transfixion or the use of cartilage battens, struts, or plumping implants.

Tip Rotation with Interrupted Strip Techniques. When the integrity and spring of the alar cartilage are broken, cephalic rotation is fostered by virtue of upward scar contracture forces acting inexorably on alar cartilage segments, which are now more flail and less supported. These techniques are particularly useful when the attitude of the alar cartilages is one of a profound downward inclination, imparting a depressed or snarl-like appearance to the nose. Caution must be exercised constantly in the use of interrupted strips in patients with thin skin and/or more delicate cartilages because the absence of good tip-supporting structures sets the stage for loss of projection, alar collapse, notching, pinching, and asymmetry.

Lateral Interruption Techniques. In anatomic situations in which cephalic rotation is desirable and the anatomy of the bridge between the medial and lateral crus (the "dome") is esthetically pleasing, lateral interruption of the residual complete strip has merit[9] (see Figure 38–38). Avoiding interruption of the strip medially fosters symmetric healing and reduces the likelihood that uneven tip-defining points will become evident months after the operation. The lateral interruption allows the reduced alar cartilage to be pulled moderately upward by scar tissue during healing, but because the dividing cut is sited more laterally and therefore more deeply in the soft tissues of the tip, notching, pinching, and other asymmetries

are essentially prevented. If modification of the dome is necessary, transdomal suture techniques to narrow, refine, and even slightly project the tip may be favorably combined with lateral interruption.

MEDIAL INTERRUPTION TECHNIQUES. Many different methods of interruption of the residual complete strip at or near the dome have been described; each predictably leads to some degree of cephalic rotation, and the complete strip is converted to two or more segments of flail cartilage (see Figure 38–38). Planned rotation with this approach is reserved for patients with thicker skin and supporting structures to minimize undesirable consequences of asymmetric healing and even over-rotation. Elevation of the medial aspect of the nostril margin, an onerous stigma of nasal surgery, is more common with medial strip interruption.

Medial interruption techniques almost always result in a moderate to major loss of tip projection, and adjunctive procedures are required to restore or augment tip projection to avoid tip ptosis.

LATERAL INTERRUPTION TECHNIQUES WITH SUTURE ROTATION. An ideal method for significant tip rotation would combine lateral strip interruption to preserve the integrity of the strip medially with a calibrated triangular excision of cartilage laterally and stabilization of the cut cartilage edges with suture fixation[3] (Figure 38–39). The degree of rotation realized here is controllable by the surgeon, essentially eliminates most of the undesirable sequelae of interrupted strip techniques, and changes in a predictable and permanent way the attitude of the alar cartilages. The suture fixation helps to diminish loss of tip support inherent in most interrupted strip techniques. In individuals with thin or moderately thin skin with more delicate cartilages, this method is highly predictable and desirable.

As with the rotation concepts discussed in connection with complete strip techniques, the same adjunctive techniques for enhancing tip rotation may be useful to combine with interrupted strip techniques. Included among these adjunctive techniques to cephalic rotation are (1) shortening of the caudal edge of the septum, (2) shortening of the caudal margins of the upper lateral cartilages, (3) high septal incision for wedge excision of septum, and (4) reduction of overly convex medial crura (caudal margin) (Figure 38–40, A to D).

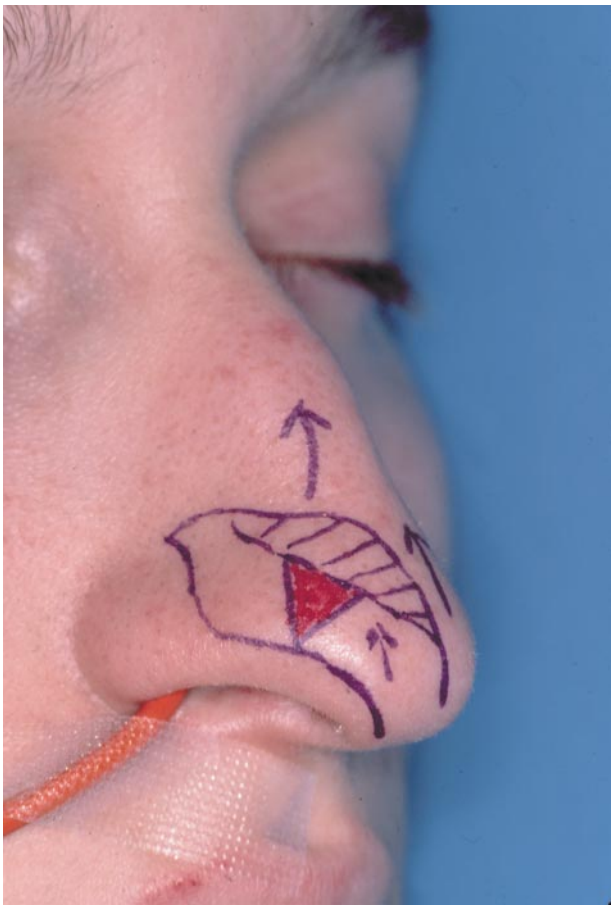


FIGURE 38–39. Lateral interruption of the residual lateral crus, with calibrated excision of a base-up triangle, allows a *controlled* cephalic rotation dependent on the geometry of the excised triangle.

Over-rotation, however, must be assiduously avoided because the correction of this undesirable postoperative situation is often difficult or impossible.

CORRECTION OF THE OVERPROJECTING TIP. Profound facial and nasal disharmony may result from the anatomic facial feature variant termed "the overprojecting nose." Because the entire nose, and especially the normal nasal tip, is composed of distinct, interrelated anatomic components, any one or a combination of several of these components may be responsible for the tip that projects too far forward of the anterior plane of the face. The guidelines for determining appropriate and inappropriate tip projection are now well accepted (Figure 38–41). When numerous patients with overprojecting tips are analyzed, it becomes apparent that no single anatomic component of the nose is constantly responsible for

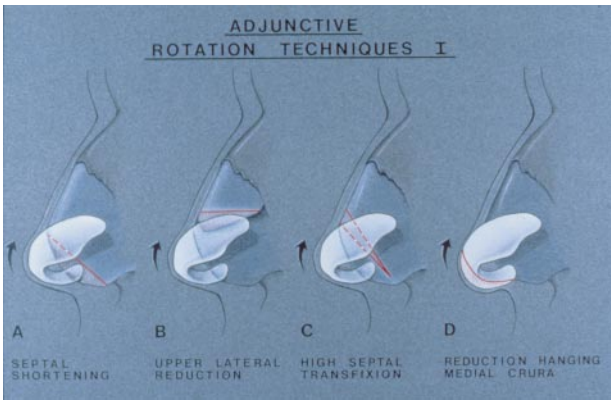


FIGURE 38–40. *A* to *D*, Adjunctive techniques tend to foster cephalic rotation of the nasal tip.

overprojection; therefore, no single surgical technique is uniformly useful in correcting all of the problems responsible for the various overprojection deformities.[3,4]

Accurate anatomic diagnosis allows preoperative development of a logical individualized strategy for correction and tip retropositioning. In almost every instance, *weakening or reducing of normal tip support mechanisms is required* to achieve normality, supplemented by reduction of the overdeveloped components. The following anatomic variants are commonly responsible individually or collectively for overprojection of the nasal tip.

Overdevelopment of the alar cartilages, commonly associated with thin skin and large nostrils, is frequently encountered in the overprojecting nose. The junction between the medial and lateral crura may form an overlarge dome of significant convexity, or the anatomic dome area may be sharply angulated, twisted, or even buckled, frequently demonstrating significant asymmetry of the entire tip and its tip-defining points. The hypertrophied cartilages must be delivered, their abnormalities visually diagnosed, and overall volume reduction of both the lateral and medial crura accomplished. Portions of the medial crus may require resection to retroposition the nasal tip satisfactorily.

Overprojection and obliteration of a definitive nasolabial angle may be the result of *overdevelopment of the caudal part of the quadrilateral cartilage* (Figure 38–42). The nasal spine may, in fact, be of normal size, but if it is even slightly overlarge, it compounds the problem of overprojection. In effecting repair, the caudal septal margin abutting the nasal spine should always be inspected and shave-reduced to normal proportion before sacrificing any of the nasal spine.

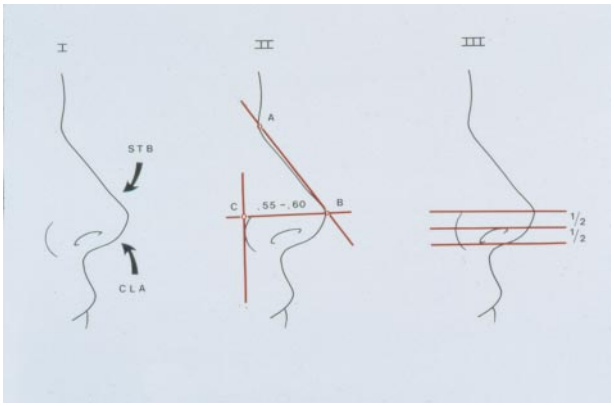


FIGURE 38–41. Goode has characterized the guidelines for nasal tip projection. If the tip projects from the alar-facial junction to the nasal tip–defining point more than 0.55 to 0.6 of a line connecting the tip-defining point and nasal frontal angle, the nose may be considered overprojected.

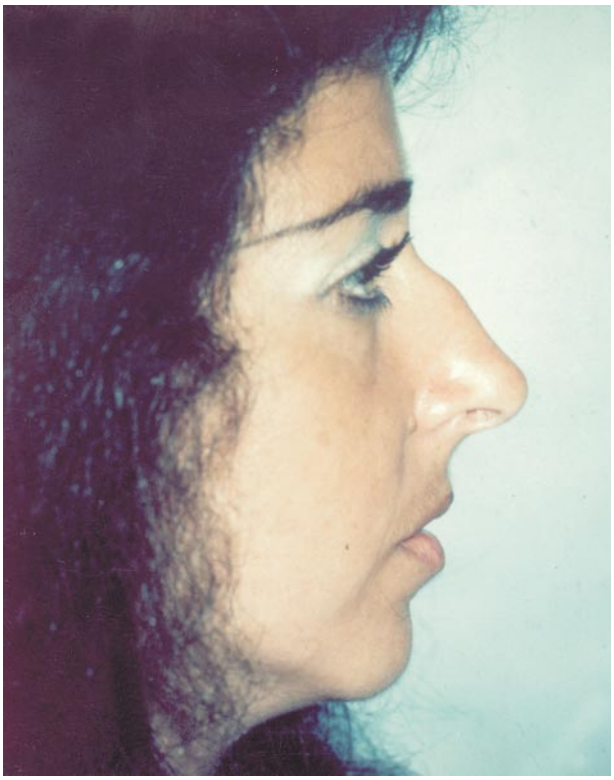


FIGURE 38–42. Overprojection of the nose as a result of overdevelopment of the caudal part of the quadrilateral cartilage.

A *high anterior septal angle* caused by an overdeveloped quadrilateral cartilage component may spuriously elevate the tip to an abnormally forward projecting position (Figure 38–43), even when associated with otherwise perfectly normal tip anatomy.[12] This condition tends to "tent" the tip away from the face and "tether" the upper lip, producing an indefinite nasolabial angle and, on occasion, creating abnormal exposure of the maxillary gingiva, particularly on smiling. Correction demands a departure from the normal operative sequence of correcting the tip first. The initial surgical steps are planned to lower the cartilaginous profile first, releasing the tip from an abnormal overprojected influence. Further tip refinement measures can then be carried out as desired and indicated by the alar cartilage anatomy.

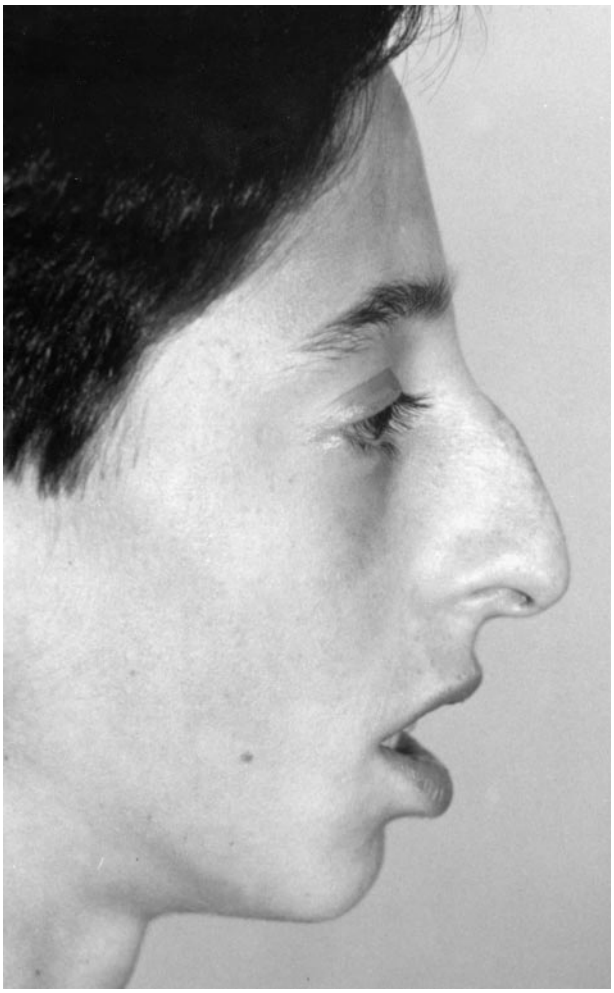


FIGURE 38–43. Overprojection of the nose secondary to an overdeveloped dorsal part of the quadrilateral cartilage component.

A less common cause of excess nasal tip projection is an *overlarge nasal bony spine*, which seemingly imparts an upward thrust of the tip components (which may otherwise be of normal dimensions). Compounding this abnormal appearance is often a coexistent blunting of the nasolabial angle, which may appear full, webbed, and excessively obtuse, with no obvious demarcation between the tip and columella. The upper lip may appear short, tethered, and tense, often exposing excessive gingiva in facial repose as well as in animation. Rongeur or osteotome reduction of the overlarge spine and associated caudal quadrilateral cartilage and soft tissue is a surgical prerequisite to tip retrodisplacement.

Tip overprojection may occur as a result of an *overly long columella* associated with excessively long medial crura. In this deformity, the infratip lobule is commonly insufficient, creating the effect of extremely large and disproportionate nostrils.[13] This deformity suggests the use of an external approach to the nasal tip to shorten the columellar length as well as that of the medial crura.

Various *combinations* of the foregoing hypertrophic anatomic problems may contribute to the overprojecting tip problem. In preoperative analysis, each nasal component must be compulsively identified and analyzed; only then can a definitive plan for natural correction be conceived. Generally, a combination of weakening of the major tip support mechanisms associated with reduction of the components responsible for the tip overprojection is carried out incrementally and as conservatively as possible to achieve the desired normal final result in a progressive fashion. The various components capable of creating or contributing to overprojection of the nose are shown in Table 38–2.

Iatrogenic overprojection may occur when surgeons intent on profoundly increasing tip projection produce an unnaturally sharp and projected tip con-

TABLE 38–2. Causes of Overprojecting Nasal Tip

Alar cartilage overdevelopment
Nasal spine overdevelopment
Overdevelopment of the caudal aspect of the septum
Overdevelopment of the dorsal aspect of the septum
Elongated columella and medial crura
Combined anatomic abnormalities
Iatrogenic overprojection

figuration (often associated with over-rotation of the tip).[4] These misadventures commonly result from overaggressive tip surgery in which portions of the lateral crus are borrowed and rotated medially to increase medial crus projection.

ALAR BASE REDUCTION. Appropriate retroprojection of the projecting nose typically requires diminishing the various major and minor tip support mechanisms to retroposition the tip closer to the face. A concomitant reduction of the alar component length and lateral flare (occasioned by tip retropositioning) is usually required to improve nasal balance and harmony. Alar wedge excisions of various geometric designs and dimensions are necessary to balance alar length and position (Figure 38–44). The exact geometry of these excisions is determined by the present and intended shape of the nostril aperture, the degree and attitude of the lateral alar flare, the width and shape of the nostril sill, and the thickness of the alae. It is axiomatic that the surgeon creating alar reduction by excision of alar or nostril floor tissue should always err on the side of conservatism and strive for symmetric repair because overaggressive and asymmetric tissue resection leads to an almost irreversible situation of disharmony and even nostril stenosis.

Profile Alignment Three anatomic nasal components are responsible for the preoperative profile appearance: the *nasal bones, the cartilaginous septum*, and *the alar cartilages*. Generally, all three must undergo modification to create a pleasing and natural profile alignment. If the nose is overlarge with a convex profile, reduction of the three segments is required. Less commonly (except in revision rhinoplasty), profile *augmentation* with autograft materials must be accomplished.

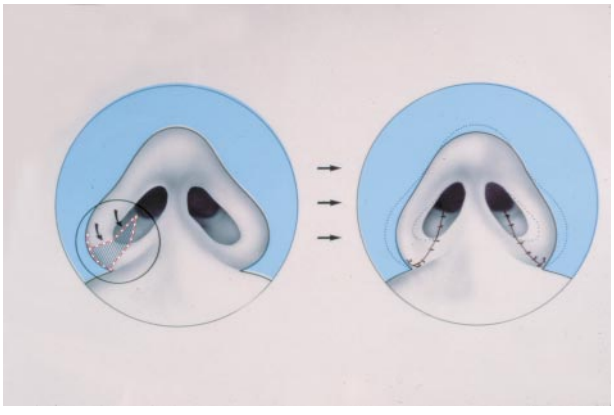


FIGURE 38–44. Alar wedge excisions of various geometric designs and dimensions are necessary to improve the alar flaring found in certain patients. The exact design of each tissue removal depends on the dimensions of the alar sidewalls and the nostril itself.

The surgeon visualizes the ultimate intended profile, extending from the nasofrontal angle to the tip-defining point, and then on around the infratip lobule and columella to the nasolabial angle. The extent of reduction of bone, cartilage, and soft tissue always depends on and should be guided by stable tip projection; *therefore, positioning the projection of the tip at the outset of the operative procedure is beneficial.*[13] Because the thickness of the investing soft tissues and skin varies at different areas of the profile and from patient to patient, dissimilar portions of cartilage and bone must be removed to create a straight or slightly concave profile ultimately.[14] Strong, high profiles generally suit the patient best in the long term, contributing to a more elegant nose on profile and oblique views and also a more narrow nasal appearance on frontal view (Figure 38–45). Over-reduced profiles result in a washed-out, indefinite, and widened appearance from the front, separating the eyes inadequately and reflecting light poorly.

In planning profile alignment, the two stable reference points are the existing (or planned) *nasofrontal angle* and the *tip-defining point*. Esthetics are generally best served when profile reduction results in a high, straight-line profile in men and the leading edge of the tip just slightly higher in women. A gentle slope of no more than 2 to 3 mm should exist between the caudal part of the cartilaginous dorsum and the most anteriorly projecting aspect of the nasal tip. Reversal of the usual preoperative tip-supratip relationship is required to achieve this esthetic ideal.

The degree and angulation of the "hump removal" depend on various factors, the most important of which are the size of the various involved anatomic components and the surgeon's confidence in the stability of postoperative tip projection.[14] These must be balanced with the personal preference for profile appearance combined with the surgeon's value judgment of facial esthetics.

Surgical access to the nasal dorsum is gained through the transcartilaginous, intercartilaginous, or transcolumellar incision, depending on which approach was used during tip refinement. In endonasal approaches, a complete transfixion inci-

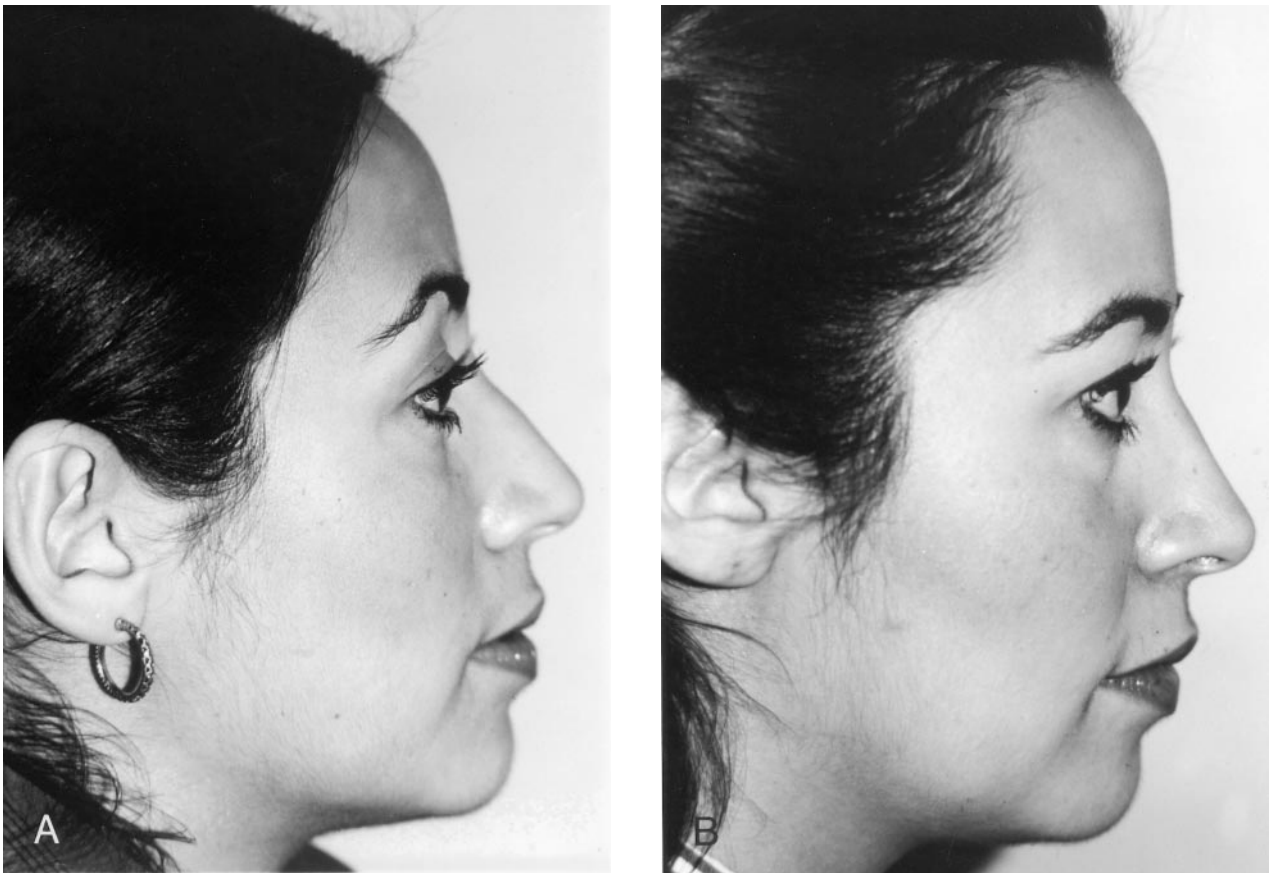


FIGURE 38–45. Patient before (*A*) and 1 year after (*B*) rhinoplasty. A strong, high profile has been maintained.

sion for exposure is unnecessary and may compromise tip support by sacrificing the attachment of the medial crural footplates to the caudal septum.

The plane of tissue elevation over the nasal dorsum is important for several reasons. A relatively avascular potential plane exists intimate (superficial) to the perichondrium of the cartilaginous vault and just below the periosteum of the bony vault.[3,6] Elevating the soft tissue flap in this important plane (Figure 38–46) preserves the thickest possible ultimate epithelial soft tissue covering to cushion the newly formed bony and cartilaginous profile. Generally, only sufficient skin is elevated to gain access to the bony and cartilaginous profile; therefore, wide undermining is unnecessary in the typical rhinoplasty. In older patients with redundant and less elastic skin, or when access is needed for major autograft augmentation, wider undermining is carried out.[15] Even in the latter instance, the periosteal soft tissue layer over the intended site of the low lateral osteotomies is preserved intact to help stabilize the mobile bony pyramid after infracture osteotomy maneuvers.

The soft tissues over the cartilaginous dorsum are elevated by means of scalpel dissection with a No. 15c blade (Figure 38–47), and the periosteum over the bony pyramid is lifted from its stable bony attach-

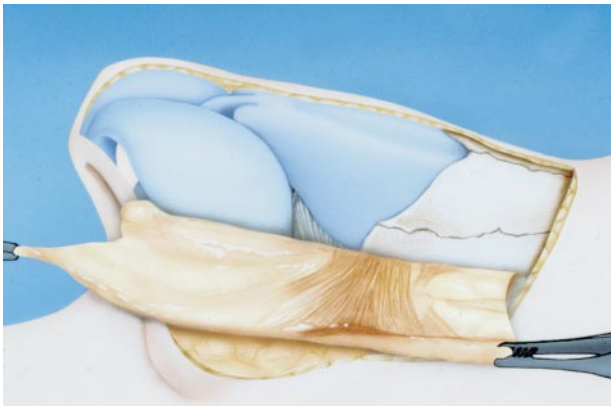


FIGURE 38–46. Favorable tissue dissection plane of the nose, located intimate to the cartilaginous structure of the nose, beneath the overlying superficial musculoaponeurotic system fascial network.

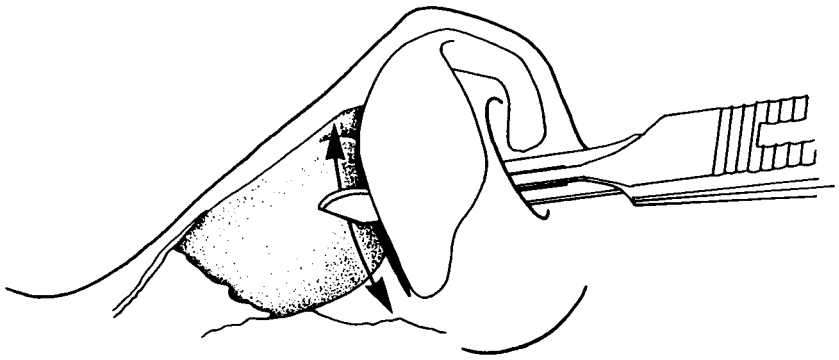


FIGURE 38–47. Elevation of the nasal skin and subcutaneous tissue layers by knife dissection, staying intimate to the perichondrium overlying the cartilaginous pyramid.

ment with the knife and sharp Joseph elevator (Figure 38–48). Because the periosteum inserts into the internasal suture line in the midline, the periosteum is lifted on either side of this suture and the space is brought into continuity with the sharp scissors. Little or no bleeding should ensue during uncovering of the nasal skeleton in these important planes, allowing direct visual assessment of the anatomy encountered.

Either of two methods of profile alignment is preferred: *incremental* or *en bloc.*[14] In the first method, the cartilaginous dorsum is reduced by incrementally shaving away the cartilaginous dorsum until an ideal tip-supratip relationship is established (Figure 38–49), followed by sharp osteotome removal of the residual bony hump.

If only minimal hump removal is contemplated, the knife is positioned at the osseocartilaginous junction, plunged through this area, and then advanced caudally to and around the anterior septal angle of the caudal portion of the septum. In

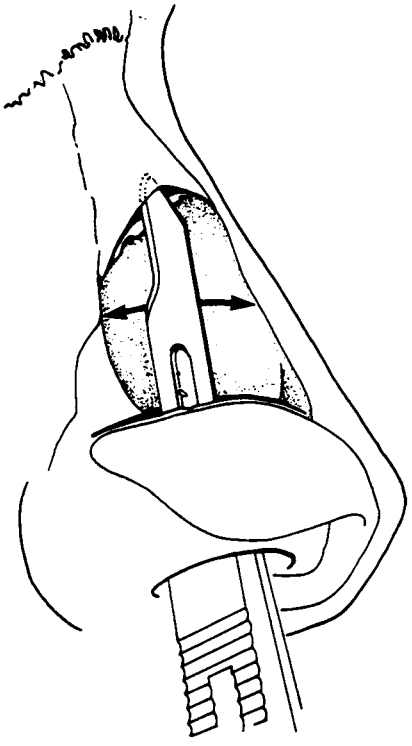


FIGURE 38–48. Elevation of the periosteum overlying the bony pyramid.

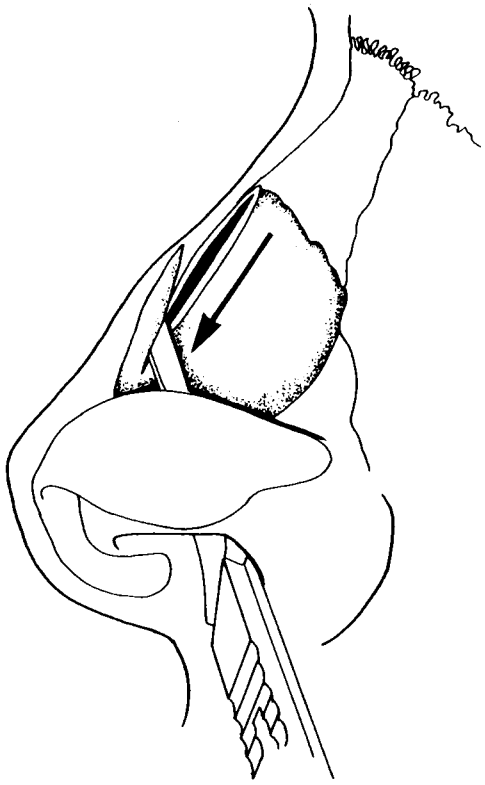


FIGURE 38–49. Incremental reduction of the cartilaginous pyramid.

large cartilaginous reductions, a portion of the upper lateral cartilage attachment to the quadrilateral cartilage must be removed with the dorsal septum, but leaving these two cartilaginous components attached by the intact underlying mucoperichondrial bridge.

A Rubin osteotome, honed to razor sharpness for each procedure and seated in the opening made by the knife at the osseocartilaginous junction, is advanced cephalically to remove the desired degree of bony hump in continuity with the cartilaginous hump (Figure 38–50).

Any remaining irregularities are corrected under direct vision with a knife and sharp tungsten carbide rasp. Palpating the skin of the dorsum with the examining finger moistened with hydrogen peroxide often provides clues to unseen irregularities, as does intranasal palpation of the profile with the noncutting edge of the No. 15 blade.

Except in large or severely twisted noses, it is unnecessary and potentially harmful to separate the upper lateral cartilages from the septum by cutting through the mucoperichondrial bridge of tissue connecting them at the nasal valve.[16]

Redundant soft tissue around the anterior septal angle may be trimmed away to improve tip-supratip definition. The caudal edge of the septum, assessed by stretching the partial transfixion incision posteriorly, lies exposed for geometric shortening or repositioning.

In patients in whom the nasofrontal angle is poorly defined or in need of retropositioning, weakening of the bone in the desired area is accomplished before bony hump removal. At the exact site in the midline where the nasofrontal angle is desired, a 2 mm osteotome is plunged transcutaneously into the midline of the nasal bones[3] (Figure 38–51). By angulating this small osteotome laterally on either side, the exact cephalic extent of bony hump removal may be controlled by *scoring the bone* in a horizontal line at the nasofrontal angle. During the bony hump removal phase of profile alignment, the nasal bones fracture cephalically where this weakening maneuver has established a bony dehiscence, allowing the surgeon some additional control over the ultimate site of the nasofrontal angle. Creating a more caudally placed angle provides the illusion of a shorter nose without actually shortening it, whereas establishing a more cephalically placed angle creates the appearance of a longer nose.

In patients in whom the nasofrontal angle is overly deep, augmentation with residual septal cartilage or remnants of the excised alar cartilages provides a beneficial esthetic refinement (Figure 38–52).

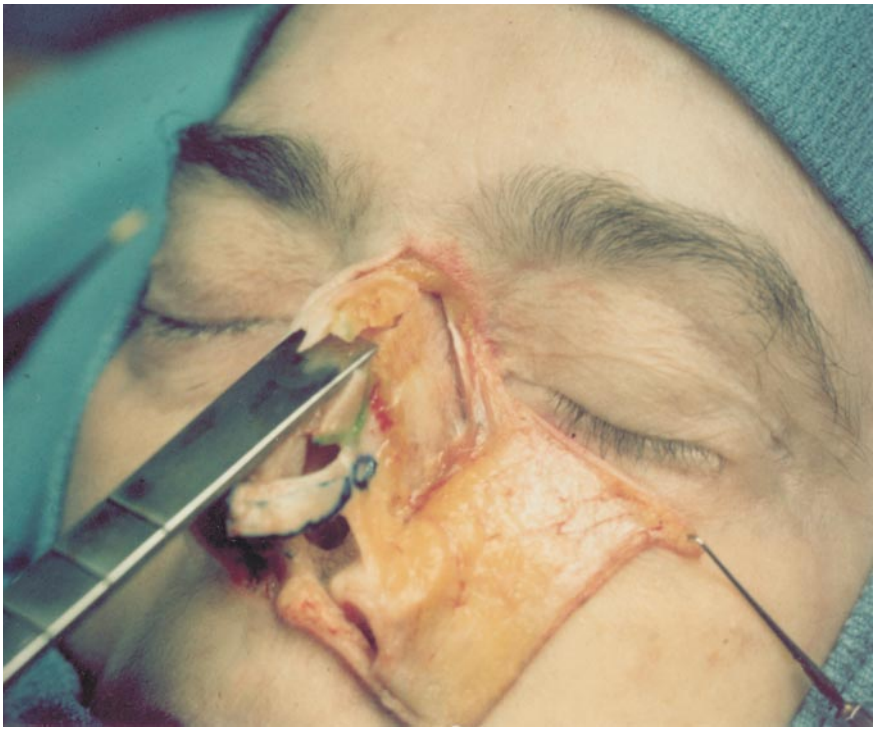


FIGURE 38–50. Sharp osteotome excision of the nasal bony hump.

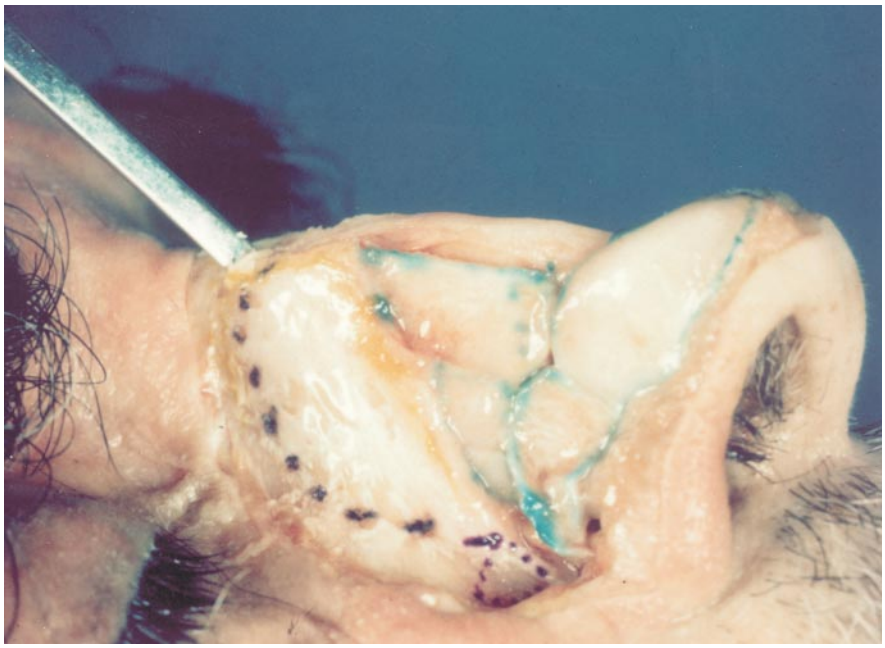


FIGURE 38–51. Weakening of the bone at the proposed nasofrontal angle by means of a percutaneous osteotomy using a 2 mm osteotome. The bone is scored and weakened to facilitate ultimate fracture at the intended site of the nasofrontal angle.

Further profile enhancements may be favorably developed with contouring cartilage grafts positioned along the dorsum, supratip area, infratip lobule, columella, and nasolabial angle. In the last site, so-called "plumping" grafts are commonly used to open an otherwise acute or unsatisfactory nasolabial angle and thereby contribute to improved profile appearance.[17] The illusion of tip rotation and nasal shortening results from this maneuver (Figure 38–53), reducing the degree of actual shortening required and preserving a longer and often more elegant nose.

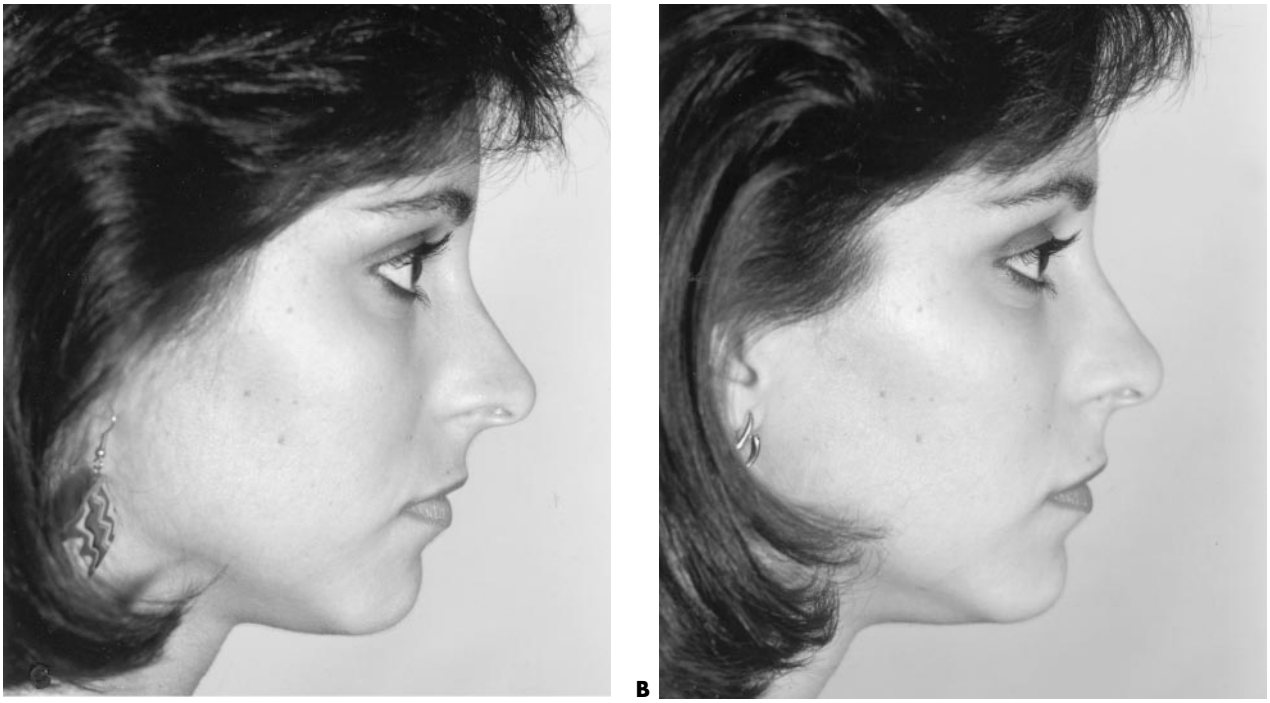


FIGURE 38–52. Patient before (*A*) and 1 year after (*B*) rhinoplasty whose profile has been improved by cartilaginous grafts positioned at the nasolabial angle and supratip dorsum.

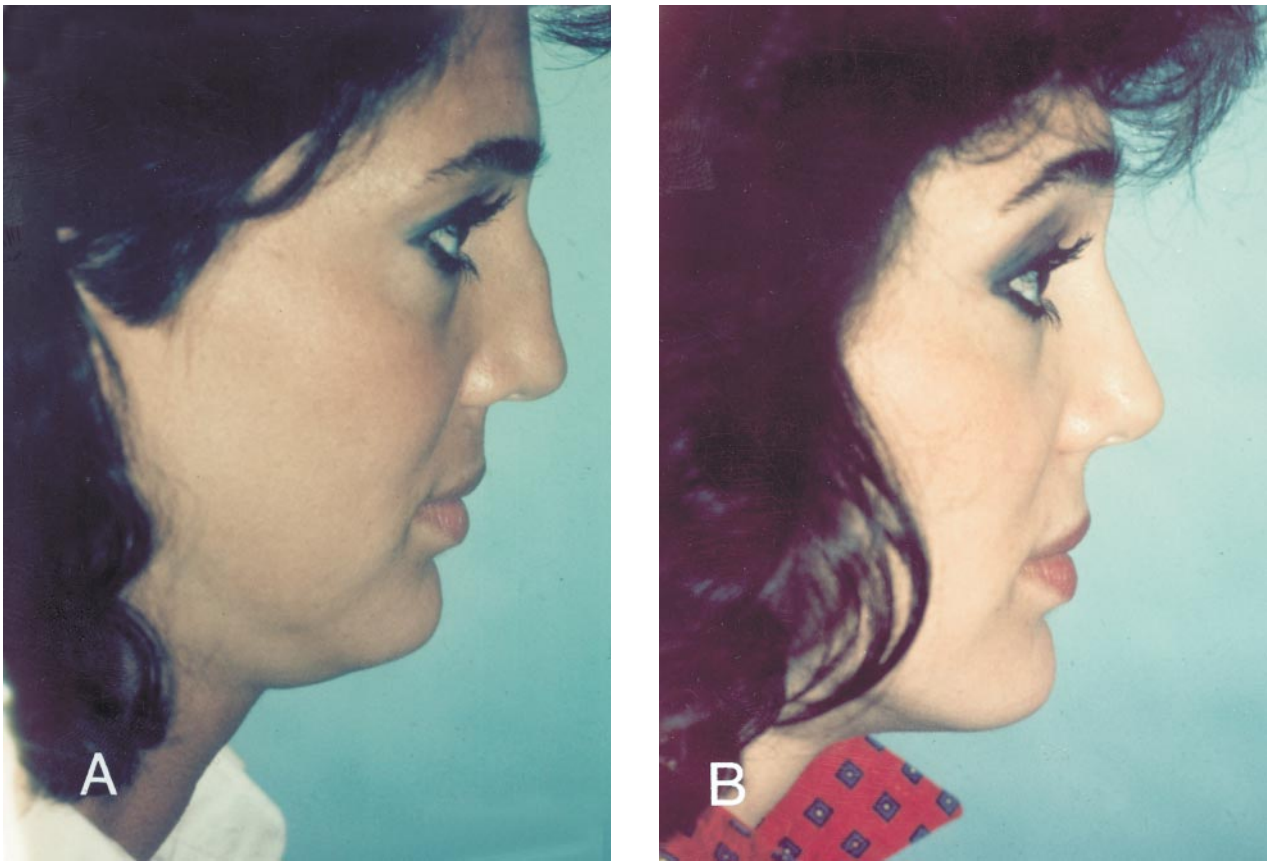


FIGURE 38–53. Patient before (*A*) and 1 year after (*B*) rhinoplasty, illustrating the benefits of augmentation of the retracted nasolabial angle and supratip dorsum.

Bony Pyramid Narrowing and Alignment Significant advances have been made over the past two decades in the reduction of osteotomy trauma in rhinoplasty surgery. Osteotomies, the most traumatic of all nasal surgical maneuvers, are best delayed until the final step in the planned surgical sequence, when vasoconstriction exerts its maximal influence and the nasal splint may be promptly positioned.[3,18,19]

Profile alignment in the typical reduction rhinoplasty inevitably results in an excessive plateau-like width of the nasal dorsum, requiring narrowing of the bony and cartilaginous pyramid to restore a natural and more narrow frontal appearance to the nose. The lateral bony sidewalls (consisting of the nasal bones and maxillary ascending processes) must be completely mobilized by nongreenstick fractures and moved medially (exceptions may exist in older patients with more fragile bones in whom greenstick fractures may be acceptable or even preferable). The upper lateral cartilages are also moved medially because of their stable attachment cephalically to the undersurface of the nasal bones.

To facilitate atraumatic low lateral osteotomy execution, medial oblique osteotomies angled laterally 15 to 20 degrees from the vertical midline are preferred (Figure 38–54). By creating an osteotomy dehiscence at the intended cephalic apex of the lateral osteotomy, the surgeon exerts added control of the exact site of backfracture in the lateral bony sidewall. A 2 or 3 mm sharp micro-osteotome is positioned intranasally at the cephalic extent of the removal of the bony hump (if no hump removal has been necessary, the site of positioning is at the caudal extent of the nasal bones in the midline). The osteotome is advanced cephalo-obliquely to its intended apex at an angle of 10 to 15 degrees, depending on the shape of the nasal bony sidewall. Little trauma results from medial oblique osteotomies, which prevent the ever-present possi-

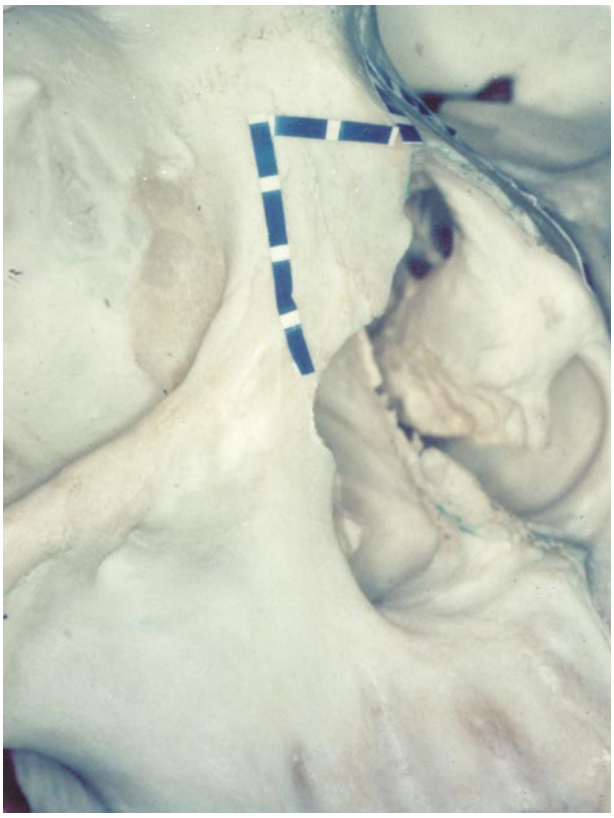


FIGURE 38–54. Medial oblique osteotomies course laterally, approximately 15 to 20 degrees from the vertical midline.

bility of eccentric or asymmetric surgical fractures from developing when lateral osteotomies alone are performed. In addition, bony narrowing to accomplish desired infracture as a consequence of lateral osteotomies combined with medial oblique osteotomies occurs without strong manual pressure exerted on the nasal bones, a traditional but unnecessary traumatic maneuver.

Trauma may be significantly reduced in lateral osteotomies if 2 or 3 mm micro-osteotomies are used to accomplish a controlled fracture of the bony sidewalls.[3] No need exists for elevation of the periosteum along the pathway of the lateral fractures because the small osteotomies require little space for their cephalic progression. Appropriately, the intact periosteum stabilizes and internally splints the complete fractures, facilitating stable and precise healing. The low curved lateral osteotomy is initiated by pressing the sharp osteotome through the vestibular skin to encounter the margin of the piriform aperture at or just above the inferior turbinate. This preserves the bony sidewall along the floor of the nose, where narrowing would achieve no favorable esthetic improvement but might compromise the inferior nasal airway without purpose. The pathway of the osteotome then progresses toward the base of the maxilla, curving next up along the nasal maxillary junction to encounter the previously created small medial oblique osteotomy (Figure 38–55). A complete, controlled, and atraumatic fracture of the bony sidewall is thus created, allowing infracture without excessive traumatic pressure. Immediate finger pressure is applied bilaterally over the lateral osteotome sites to forestall further extravasation of blood into the soft tissues. *In reality, little or no bleeding occurs during micro-osteotomies because the soft tissues embracing the bony sidewalls remain essentially undamaged.*

In most rhinoplasty procedures, controlled nasal fractures as a result of osteotomy should cause slight but definite mobility of the bony sidewalls stabilized by the internal and external periosteum,[19] which bridges the nasal fragments on either side of

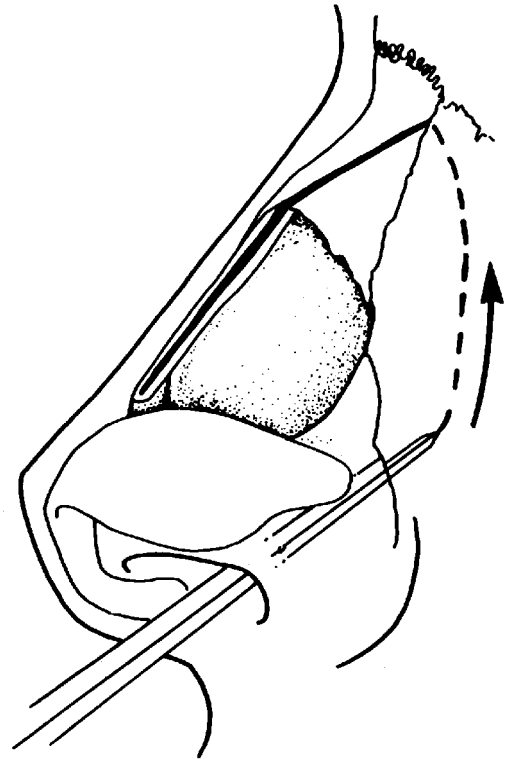


FIGURE 38–55. Curved low lateral osteotomies, created with a 2 or 3 mm osteotome, cause minimal trauma to the surrounding tissues.

the osteotome pathway. Large guarded osteotomies destroy this vital periosteal sling, potentially rendering the bony fragments unstable and susceptible to eccentric or asymmetric healing. In addition, trauma from large osteotomes may produce increased bleeding, edema, and unnecessary ecchymosis.

In deviated noses characterized by essentially convex or concave bony asymmetries, excessively wide or extremely thick bones (including revision rhinoplasty), *double* lateral osteotomies or *percutaneous* osteotomies may be considered for improved mobilization and regularization.[18] This decision is best determined preoperatively to allow the higher osteotomy to be accomplished before the low lateral osteotomy. Reversing this order necessitates attempting the higher osteotomy on an already mobilized lateral bony sidewall, a more difficult task.

On completion of satisfactory osteotomies and suitable esthetic nasal narrowing, the profile line is finally inspected and palpated for irregularities or inadequate alignment. Because the upper lateral cartilages move medially with the bony sidewalls following osteotomies, their dorsal margins should be trimmed to sit flush with or slightly lower than the cartilaginous profile line. Excess soft tissue, if present, is excised and the new nasal dorsum is inspected for and cleaned of any debris. Any profile grafts to be placed to improve the ultimate intended profile line are now scalpel-sculptured to size and placed accordingly. If limited undermining of the overlying dorsal epithelium has occurred, grafts may be placed with no requirement for suture fixation. If a large subcutaneous pocket exists, however, transcutaneous pull-out 5-0 mild chromic catgut sutures are used to fix the grafts into the intended site. The sutures, cut flush with the skin at 5 to 7 days, retract into the subcutaneous space and are absorbed.

Final subtle nasal refinements are now completed and may include caudal septal reduction, resection of excessive vestibular skin and mucous membrane, trimming of the caudal margins of the upper lateral cartilages (only if overlong or projecting into the vestibule), columellar narrowing, and bilateral alar base reduction. These final maneuvers are carried out with the assistant maintaining constant finger pressure over the lateral osteotomy sites to prevent even minimal oozing and intraoperative swelling. All incisions are closed completely with 5-0 chromic catgut suture. No permanent sutures are used.

Nasal dressings are now applied. No intranasal dressing or packing is necessary in routine rhinoplasty. If septoplasty has been an integral part of the operation, a folded strip of Telfa is placed into each nostril along the floor of the nose to absorb drainage. The previously placed transseptal quilting mattress suture (Figure 38–56) acts as a sole internal nasal splint for the septum,[3] completely obliterating the submucoperichondrial dead spaces and fixing the septal elements in place during healing. The external splint consists of a layer of compressed Gelfoam placed along the dorsum and stabilized in place with flesh-colored Micropore tape, extending over and laterally beyond the lateral osteotomy sites. A small aluminum and Velcro splint applied firmly over the nasal dorsum completes the operation (Figure 38–57).

Postsurgical Considerations The care of the postrhinoplasty patient is directed toward patient comfort, reduction of swelling and edema, patency of the nasal airway, and compression/stabilization of the nose.

Whether the patient is discharged on the afternoon of or the morning after surgery, all intranasal dressings are removed from the nose before the patient leaves. A detailed list of instructions is supplied for the patient or accompanying family member (Table 38–3); the important aspects of these "do's" and "don'ts" are emphasized. Prevention of trauma to the nose is clearly the most important consideration. Oral decongestant therapy is helpful,

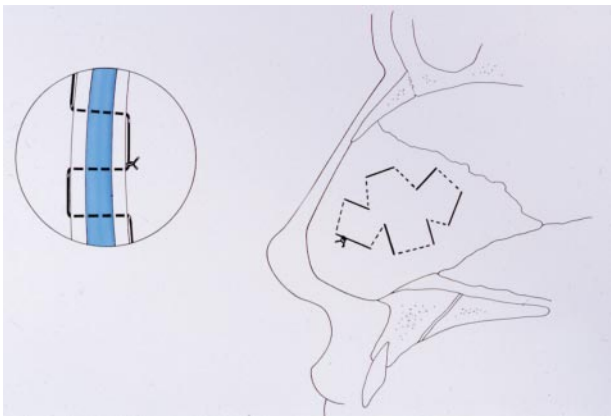


FIGURE 38–56. Quilting mattress transseptal suture woven back and forth through the dissected septum to close dead space, coapt the mucoperichondrial flap(s), and internally support the reconstructed septal components. *Inset:* Sagittal view.

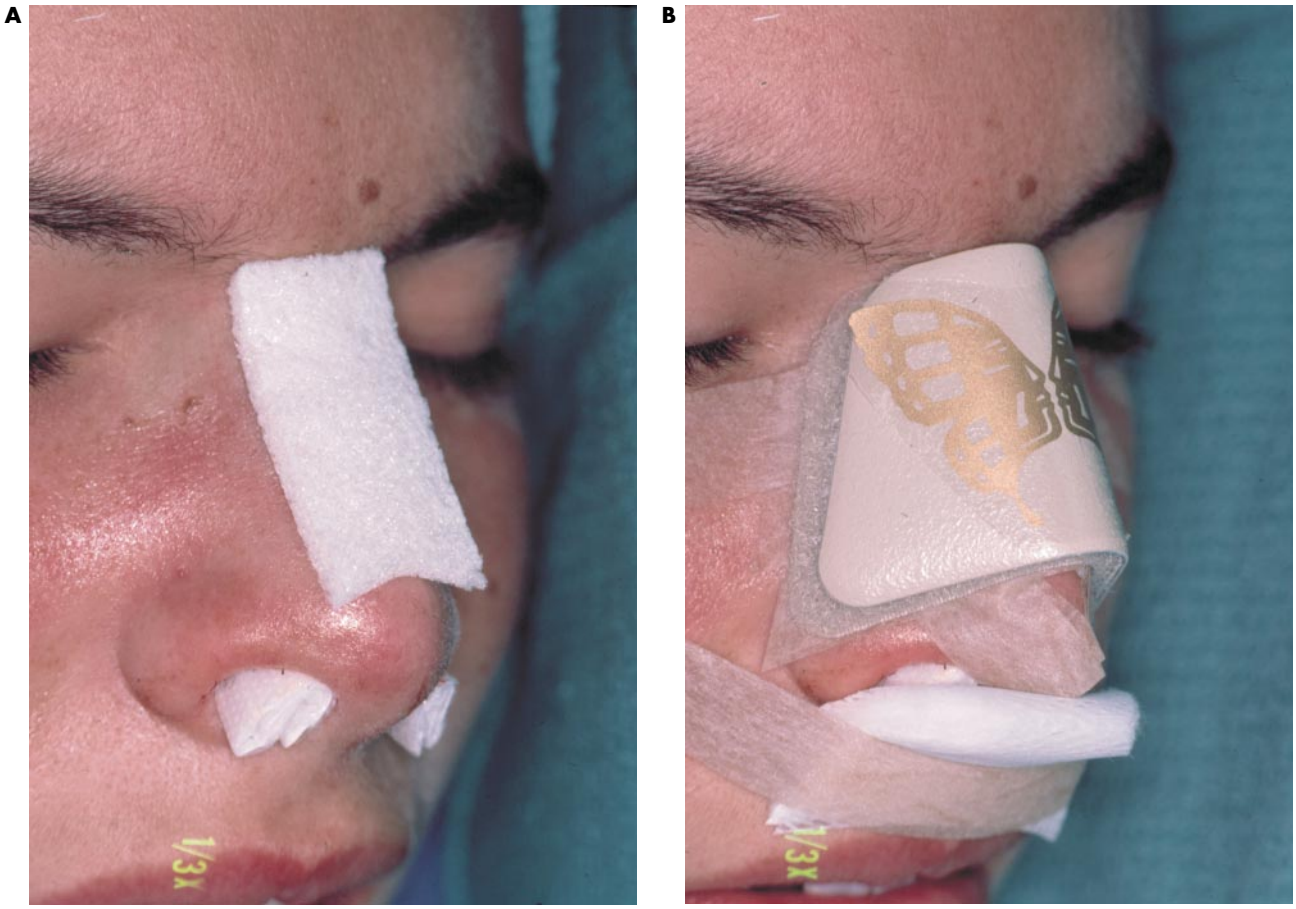


FIGURE 38–57. Nasal splint applied to a reconstructed nose and maintained for 5 to 7 days to support, splint, and protect the nose. A strip of Gelfoam is initially laid over the dorsum to protect and compress the skin.

but the value of corticosteroids and antibiotics in routine rhinoplasty is conjectural.

The external splint is removed 5 to 7 days after surgery. An important consideration should be gentle removal of the tape and splint by bluntly dissecting the nasal skin from the overlying splint with a dull instrument without disturbing or tenting up the healing skin. Failure to follow this policy may lead to disturbance of the newly forming subcutaneous fibroblastic layer over the nasal dorsum, with additional unwanted scarring and even abrupt hematoma.

Septum in Rhinoplasty

The role of the septum in rhinoplasty has always been subject to debate. The difficulty arises because of a lack of definition, namely, between the septal operation performed as an entity and its being performed as a combined operation with rhinoplasty. Septal resection without rhinoplasty found many advocates for radical removal of the septum at the turn of the century. In 1937, Peer expressed the opinion that radical removal might be necessary for certain types of deflection but emphasized the need for a section of cartilage in the anterior position to replace the septum for nasal support.[20] Fomon and associates in 1946, 1948, and 1951 described a similar method but incorporated the rhinoplastic approach of Joseph.[21] The adherents to this method did not clearly define the limitations of this procedure with regard to rhinoplasty, although this was foreshadowed. These techniques apply primarily to the correction of the deviated/dislocated septum alone and are rarely indicated when combined with rhinoplasty, except in the badly traumatized nose.

A more complex problem arises with combined correction of both an external and a septal deformity. Radical removal of the septum performed

TABLE 38–3. Patient Instructions Following Nasal Plastic Surgery

A. INTRODUCTION

Please read and familiarize yourself with these instructions both BEFORE and AFTER surgery. By following them carefully, you will assist in obtaining the best possible result from your surgery. If questions arise, do not hesitate to communicate with me and discuss your questions at any time. Take this list to the hospital with you and begin observing these directions on the day of surgery.

B. INSTRUCTIONS

1. Do not blow nose until instructed. Wipe or dab nose gently with Kleenex, if necessary.
2. Change dressing under nose (if present) as needed.
3. The nasal cast will remain in place for approximately 1 week and will be removed in the office. Do NOT disturb it; keep it dry.
4. Avoid foods that require prolonged chewing. Otherwise, your diet has no restrictions.
5. Avoid extreme physical activity. Obtain more rest than you usually get and avoid exertion, including athletic activities and intercourse.
6. Brush teeth gently with a soft toothbrush only. Avoid manipulation of upper lip to keep nose at rest.
7. Avoid excess or prolonged telephone conversations and social activities for at least 10–14 days.
8. You may wash your face—carefully avoid the dressing. Take tub baths until the dressings are removed.
9. Avoid smiling, grinning, and excess facial movement for 1 week.
10. Do not wash hair for 1 week unless you have someone do it for you. DO NOT GET NASAL DRESSINGS WET.
11. Wear clothing that fastens in front or back for 1 week. Avoid slipover sweaters, T-shirts, and turtlenecks.
12. Absolutely avoid sun or sun lamps for 6 weeks after surgery. Heat may cause the nose to swell.
13. Don't swim for 1 month since injuries are common during swimming.
14. Don't be concerned if, following removal of dressing, the nose, eyes, and upper lip show some swelling and discoloration; this usually clears in 2–3 weeks. In certain patients, it may require 6 months for all swelling to completely subside.
15. Take only medications prescribed by your doctor(s).
16. Do not wear regular glasses or sunglasses that rest on the bridge of the nose for at least 4 weeks. We will instruct you in the method of taping the glasses to your forehead to avoid pressure on the nose.
17. Contact lenses may be worn within 2–3 days after surgery.
18. After the doctor removes your nasal cast, the skin of the nose may be cleansed gently with a mild soap or Vaseline Intensive Care Lotion. BE GENTLE. Makeup may be used as soon as bandages are removed. To cover discoloration, you may use "ERACE" by Max Factor, "COVER AWAY" by Adrien Arpel, or "ON YOUR MARK" by Kenneth.
19. DON'T TAKE CHANCES! If you are concerned about anything you consider significant, call me immediately.

with extensive rhinoplasty can result in a depressed nose because of a lack of septal support. The exception is the severely traumatized nose. In these cases, the dorsum is depressed, and it is impossible to shift the septum without sacrifice of large sections of cartilage. Radical removal of the septum may then be required, but a dorsal implant of a bone or cartilage graft is necessary to correct the depression.

To correct a deviated-dislocated septum associated with a rhinoplastic correction in the developmental type of deformity safely, techniques advocating the shifting and reconstruction of the septal cartilage with minimal removal of septum are currently considered preferable to radical septal resection.[22,23]

Deviated or Scoliotic Nose

Deviation of the entire nose involves both the bony and cartilaginous vaults; deviations of the lower half of the nose are caused mainly by septal derangements accompanied by secondary effects of the cartilaginous vault (Figure 38–58).

Associated asymmetries and deformities of the upper lateral cartilage commonly accompany severe septal scoliosis. The septal and external deformities are corrected together whenever feasible. Because the procedures are interdependent, a better evaluation of the problem is possible when they are carried out simultaneously, and the patient is saved considerable time and discomfort (Figure 38–59).

Correction of the cartilaginous component precedes the management of the bony deformity to allow adequate space for readjustment and infracturing of the bony vault. Furthermore, seldom can an externally twisted nose be adequately straightened without concomitant septal realignment. The admonition "as goes the septum, so goes the nose" holds much truth.

For a better understanding of the pathologic changes and mechanics involved in correcting this type of nasal deformity, a review of etiologic and anatomic factors is presented.

Etiologic and Anatomic Factors

At birth, the septum is almost completely cartilaginous, with the exception of the vomer and the two premaxillae and their processes.[24] The vomer develops bilaterally on each side of the cartilaginous nasal septum from a pair of ossification centers, which are present at the beginning of the third month of fetal life. The bilateral plates of the vomer unite from behind and grow forward at the expense of the

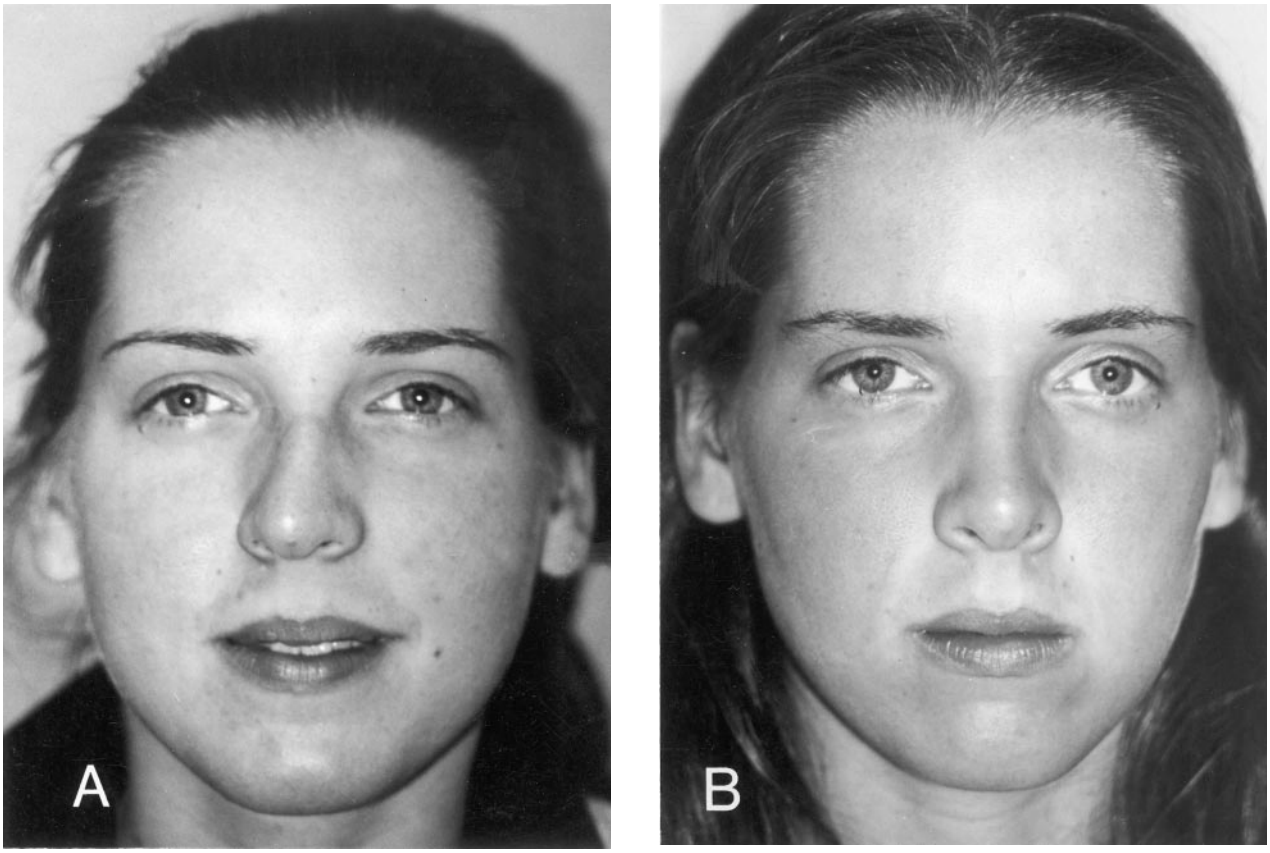


FIGURE 38–58. Deviation of the external nose caused by nasal septal deflection, before (*A*) and after (*B*) correction.

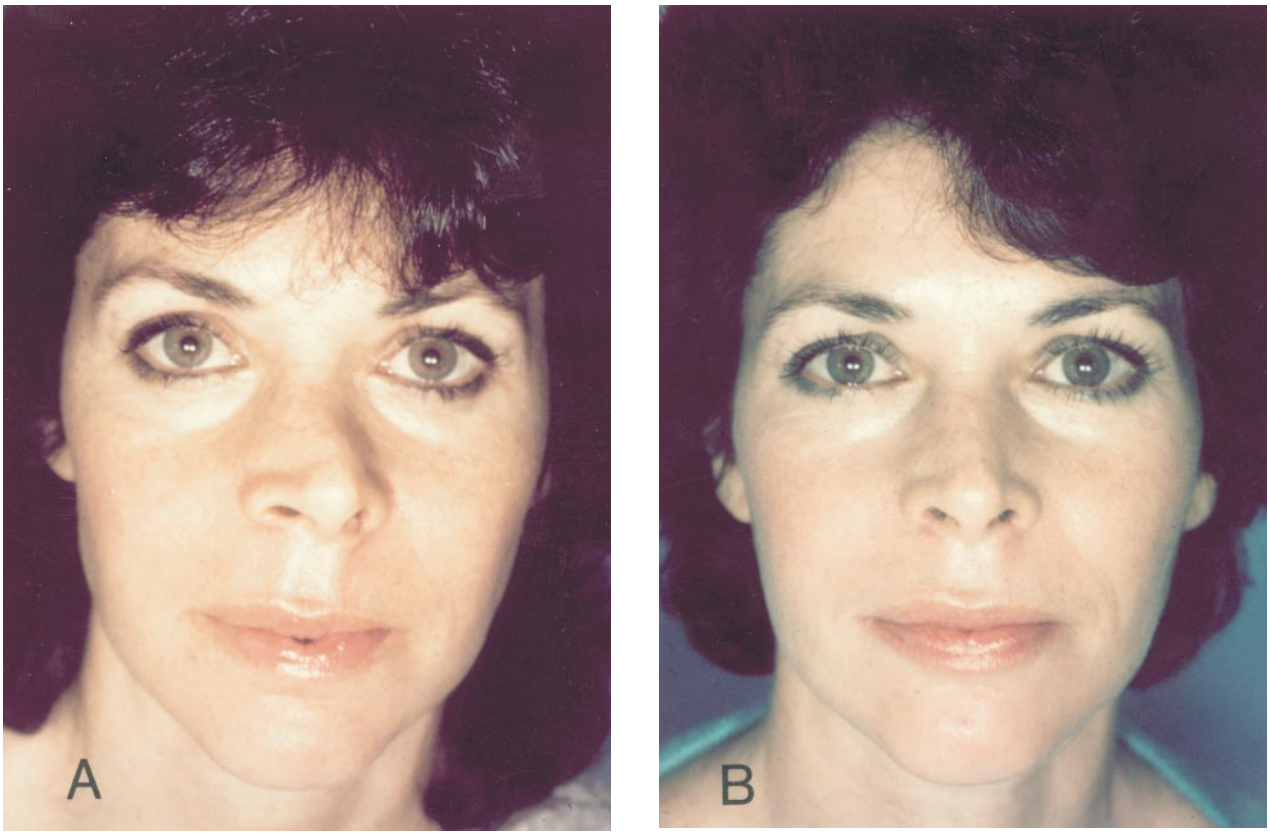


FIGURE 38–59. Patient before (*A*) and 1 year after (*B*) septorhinoplasty using preservative reconstruction of the nasal septum.

imprisoned cartilage. Their development is completed at age 15 years. Indication of the bilateral origin of the vomer is evidenced in the infant by a deep groove between the two plates. The groove is somewhat flattened in adults.

The premaxilla parallels the development of the vomer, but after the child is 6 years old, it develops rapidly. The ethmoid bone begins to ossify during the first year of postfetal life, and ossification is not completed until the end of the seventeenth year.

At birth, neither the palate bones nor the superior maxillae rise into a crest for the support of the lower edge of the septum, but in the adult, these bones have marked crests. The upward growth of the crests, along with the development of the premaxilla and vomer, combined with the downward growth of the septum from the ethmoid ossification centers and the downward expansion of the cranial cavity, may account for many deviations and dislocations of the septum. This disproportion of growth with pressure on the cartilaginous septum has been emphasized as a causative factor in septal deformities.

The vomer and premaxilla are enveloped by a periosteal covering that separates the bony portion of the septum from the cartilaginous septum, which, in turn, is enveloped in perichondrium. Microscopic studies of sections removed from the junction of bone and cartilage in the vomeroseptal maxillary crest region revealed the perichondrial envelopment of the septal cartilage and its fusion with the periosteum of the bone below. The two opposing membranes form a smooth surface between bone and cartilage, especially at the junction of the vomer, maxillary crest, and maxillary spine, where the groove of the vomer is usually shallow or flat. This region is therefore a weak point and a frequent site of traumatic dislocations. Furthermore, the bones themselves form a smooth surface. The lower end of the vomer is smooth and rests on the smooth concave surface of the maxillary crest, which, in turn, rests on the smooth concave surface of the nasal spine.

From observations made at the operating table, in the majority of patients, a dislocation is

usually accompanied by a deviation of the vomer and maxillary crest and spine along with the septal cartilage. These findings are either developmental in origin or traumatic at age 6 or 7. Trauma in early childhood is an important factor because the maxillary crests and vomer are not completely ossified, and a slight shifting of these tissues may cause the crest and vomer, as they develop, to grow to the side. This situation results in a flattening of the vomerian groove and the loss of its lip on the side of deviation. Some of these deviations may have their origin in birth trauma. Cases in which the septal cartilage is displaced from the vomer and maxillary crest are usually traumatic in origin, the injury occurring, in most instances, some time after the child reaches the age of 6 or 7. Complicated septal dislocation, accompanied by buckling, twisting, and reduplication of cartilage and marked internal deformity of depression and deviation, invariably results from trauma.[24]

Some developmental factors in the etiology of the dislocated septum may be gathered from the embryonal description of the nasopalatal relationship and the influence of palatal development on the floor of the septum. Formation of the premaxilla, eruption of the permanent incisor teeth, asymmetric development of the maxillary sinuses, thumb sucking and tongue pressure habits with resultant shifting of the alveolar ridge, mouth breathing, and congenital deformities such as cleft lip and palate are some causative factors that may account for developmental disturbances. The eruption of the permanent incisor teeth and its effect on the septum were well described by Mosher, who believed that deformities of the septum are infrequent and rarely marked in children before the second dentition. The disproportion in growth among the premaxilla, vomer, and ethmoid, with downward encroachment of the cranial cavity, is probably the dominant factor, however.

Nasal Septal Reconstruction *Introduction and Philosophy* The functions of the nasal septum have been described as follows: (1) support of the external nose, (2) regulation of air flow, and (3) support of nasal mucosa.

Since the turn of the last century, procedures designed to improve the nasal airway while preserving these functions have been developed (and occasionally discarded). The classic submucous resection with sacrifice of considerable portions of bone and cartilage has fortunately given way to more conservative procedures involving structural preservation and reconstruction. Aside from historical significance, there remains little purpose in retaining the term submucous resection, with its more radical implications. The term reconstruction of the nasal septum or septoplasty conveys the meaning of contemporary septal operations more clearly.

Preservative septal surgery is further justified by *normal* anatomy. Few "normal" septa are perfectly straight, existing without imperfection. Minor septal irregularities after appropriate reconstruction are inconsequential provided that they create no obstruction and contribute to no external nasal deformity. Radical septal resections in pursuit of a "straight" septum are therefore generally without virtue.

Septal reconstruction is usually carried out with (and usually as an integral part of) rhinoplasty. Reduction rhinoplasties invariably diminish breathing space. Reconstructive septoplasty can rescue an airway otherwise potentially compromised by a purely esthetic procedure. Invariably, the deformed septum contributes to the anatomic deficit inherent in the twisted nose and is best corrected at the outset of septorhinoplasty. lt is frequently remarkable at the operating table how initial septoplasty transforms the crooked nose into near-perfect cartilaginous alignment. As in all surgery for which perfection is the goal, a secondary procedure of lesser magnitude may be occasionally required (less than 5% of all procedures), and all patients deserve to know that fact before undergoing a reconstructive procedure ("the crooked nose has a memory").

Finally, preservative conservation septal surgery should totally negate the most severe sequelae of more radical septal resections: columellar retraction, saddle nose, airway collapse, loss of tip support, and septal perforation. The latter conditions present complex difficult reconstructive exercises and are better avoided than risked.

Principles of Septal Reconstruction. If the preceding philosophy is accepted, certain precepts emerge as cardinal to all septal surgery. These precepts have as their basis an increasingly detailed, atraumatic dissection and mobilization of the septal components, an assessment of the obstructive problem, and, finally, a reconstruction and realignment after

minimal tissue sacrifice. These major steps involve the following surgical phases, not always carried out in this sequence:

1. Whenever possible, elevate the mucoperichondrial flap only on one side.
2. Atraumatically disarticulate the attachment of the quadrilateral cartilage to the perpendicular plate of the ethmoid and to the vomer. If a vertical angulation exists just caudal to this articulation, a common site of deviation, the disarticulation can be positioned at this angulation.
3. Mobilize the quadrilateral cartilage along the floor of the nose at the maxillary crest. A narrow horizontal strip of cartilage may be removed to facilitate this mobilization without compromising septal support.
4. Isolate the *bony* septum between its bilaterally elevated mucoperiosteal flaps and medially reposition or resect the portion creating obstruction (bone grafts are commonly taken at this juncture).
5. Realign the cartilaginous septum with the conservative manipulations to be described subsequently.
6. Stabilize all realigned septal segments with quilting mattress transseptal absorbable sutures during the healing phase.

Certain vital fundamental technical concepts are required to accomplish these goals of septal reconstruction:

1. Perform all septal surgery under direct vision. Intense fiber-optic head lighting, long nasal specula, complete septal mobilization, and effective vasoconstriction make this an easily realized surgical prerequisite.
2. Preserve the contralateral mucoperichondrial flap for support, septal cartilage nutrition, and stability. Exceptions to this principle exist but are infrequent.
3. Preserve the caudal septal relationship with the membranous columella and feet of the medial crura. Severe caudal subluxation may negate this principle.
4. Dissect and mobilize the septal components fully before final deformity diagnosis. Only now should the extent of a conservative resection be planned.
5. Resist the temptation to resect more radically in pursuit of a "perfectly straight" septum.
6. In septorhinoplasty of the twisted nose, generally, the septal realignment is best created before tip and profile reconstruction. Septoplasty in combined procedures frequently requires more technical ingenuity than rhinoplasty.
7. Assiduously avoid removal of septal cartilage contributing important strength to the structure of the external nose.
8. Dissect from the "known to the unknown" to best avoid development of tears and flap perforation. If perforations are created, repair with fine suture technique.
9. Unless it contributes to deformity, preserve the upper lateral cartilage attachment to the septum.
10. Control and stabilize final septal alignment with judiciously placed transseptal sutures, thereby preventing cartilage segment overrides, hematoma, and unfavorable fibrosis. The need for long-term septal splints is thereby lessened or negated.
11. Avoid long-term intranasal packing and tamponade, a more traditional than useful exercise.
12. Constantly reassess and diagnose the obstructive problem during the course of septal surgery with inspection *combined with intranasal palpation*. As in rhinoplasty, the anatomic metamorphosis in septoplasty is dynamic and interdependent, each surgical step often dramatically influencing the next. Remain flexible in ideas and approach, ready to incorporate surgical options as required.
13. Understand that airway improvement, particularly in the twisted nose (or as the sequela of old comminuted fractures), may require nasal osteotomies to achieve optimum breathing space.

Progressive Integrated Technique of Septal Reconstruction. The surgical steps described subsequently have been evaluated over a period of 30 years and found to be effective in 90% of all septal deformities.[3,25] They are based on the principles of structural preservation and septal realignment rather than resection, with final suture stabilization of the realigned septal components. Deviation from these principles is not usually required, except in the rarer anatomic circumstances to be described.

Anesthesia and analgesia. With few exceptions, septoplasty and septorhinoplasty are performed under a combination of monitored intravenous analgesia, local infiltration, and topical anesthesia. The

patient's comfort is ensured, little bleeding develops, and the patient's condition is constantly monitored by the anesthesiologist. Intravenous sedation is administered in titrated fashion until patient comfort and tranquilization are achieved.

Neurosurgical cottonoids containing 5% cocaine solution (color coded for easy identification) are positioned under direct vision in each nasal cavity. One or two on either side ensure adequate topical anesthesia, vasoconstriction, and nasal tamponade. The mucoperichondrial flap on the intended side of dissection (ordinarily the concave side) is generously infiltrated with a standard solution of 1% lidocaine with epinephrine. Accurately piercing the submucoperichondrial plane with a No. 27 gauge needle provides a "hydraulic dissection," facilitating a rapid and bloodless flap elevation. Additional infiltration anesthesia is used according to the planned extent of septal dissection.

Twelve to 15 minutes of delay should now ensue while vasoconstriction becomes intense. During this period, final inspection and palpation of the cocainized internal nose often reveal irregularities high in the nasal vaults not previously seen. Radiograph evaluation at this juncture may aid in surgical planning. The value of *palpation* of the deformed septum and of the structural strength of the nasal tip supportive elements cannot be overemphasized.

MOBILIZATION OF QUADRILATERAL CARTILAGE. Elevation of the mucoperichondrial septal flap is usually initiated through a vertical incision placed 2 to 3 mm cephalic to the caudal end of the septal cartilage on the concave side of the septum (Figure 38–60). Actually, it makes little difference on which side this incision is placed. Preserving the normal attachment of the mucoperichondrial flap on the convex side of the septal deformity, however, seems technically easier. If severe caudal septal dislocation exists, a hemitransfixion or complete transfixion incision is generally used.

The submucoperichondrial plane is located and developed with a semisharp Dunning elevator under direct vision (Figure 38–61). If fracture adhesions, cartilage overlaps, or severe scarring interferes, the dissection is carried out by bypassing these vexing areas and dissecting above or below as required initially to circumvent areas of difficult dissection. Proper hydraulic lidocaine infiltration, as described, facilitates this elevation technique.

This dissection plane is continued onto the ethmoid perpendicular plate and vomer, elevating the continuous ipsilateral mucoperichondrial mucoperiosteal space. Moderate pressure with the tip of the elevator at the junction (articulation) of the cartilaginous and bony septal components disarticulates these elements and initiates the dissection

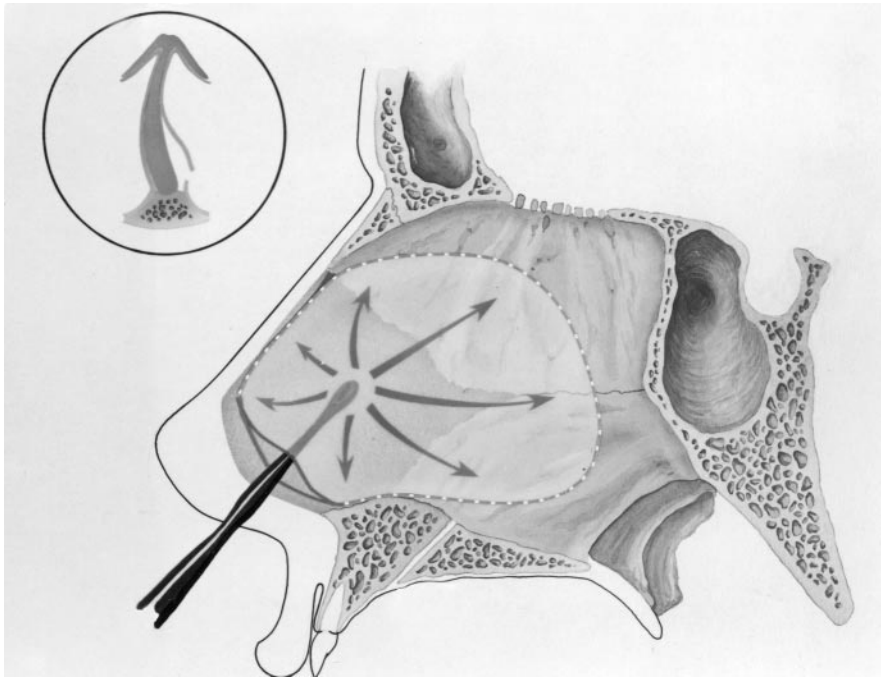


FIGURE 38–60. Septal incision is created just cephalic to the mucocutaneous junction. *Inset:* sagittal view.

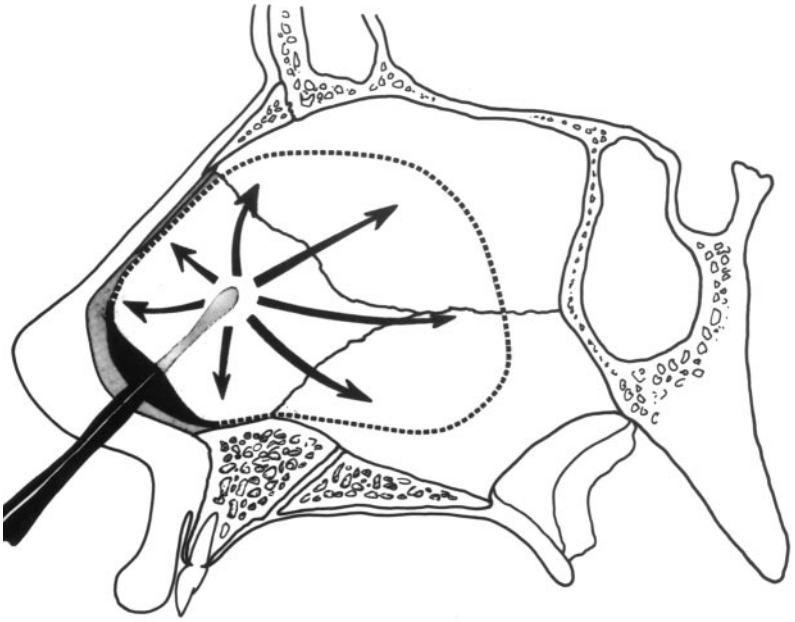


FIGURE 38–61. The submucoperichondrial plane up to and in continuity with the subperiosteal plane overlying the perpendicular plate of the ethmoid and vomer is created with a semisharp elevator.

of the contralateral mucoperiosteal flap (Figure 38–62). Because this area is a common site of septal fractures and angulation, the disarticulation can also be accomplished in a previous fracture line.

A reassessment of the septal anatomy (and deformity) is now in order. This initial dissection frequently has allowed the quadrilateral cartilage to return to the midline without further surgery, being freed from a deviated ethmoid or vomer. The cartilage may be held by a subluxation along the maxillary crest, which is overcome by resecting a horizontal strip of cartilage (and frequently bone) from along the floor of the nose. Any fibrous tissue bands further contributing to a displaced position of the cartilage are divided with sharp knife dissection. These steps will now have completely freed and mobilized the quadrilateral cartilage and usually allow its return to the midline without further surgical manipulation.

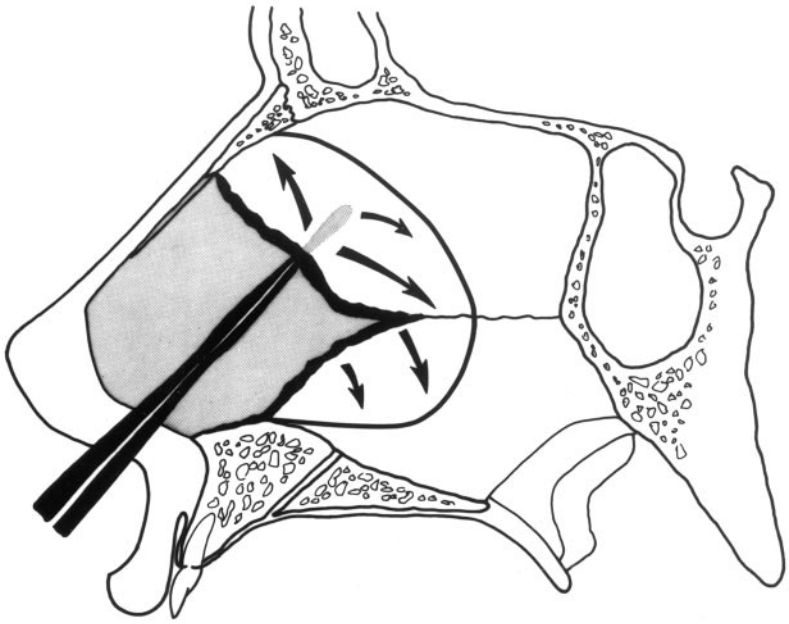


FIGURE 38–62. Disarticulation of the bony and cartilaginous septum.

CORRECTION OF BONY OBSTRUCTION. Deformities of the bony septum, particularly those high in the nasal vault, are now exposed between the blades of a long nasal speculum; selective resection of obstructing bone is gently carried out with small biting (Takahashi) forceps until the airway is clear (Figure 38–63). Larger pieces of bone required for graft or implants may be resected with fine osteotomes, heavy scissors, or bone-biting forceps. Great care should be taken in older patients (particularly osteoporotic women) to remove only small portions of the perpendicular plate of the ethmoid. The vomer provides an excellent source of bone-grafting material; flaps should be completely and bilaterally elevated to avoid creating flap tears or avulsions during vomerine resection.

These steps in surgical dissection of the septum are relatively routinely carried out in most patients and frequently are all that is required to establish a greatly improved airway. At this point in the septal reconstruction, a *reassessment* is in order before further surgery. Visually and palpably, the septal alignment is evaluated. If alignment is satisfactory and the airway adequate, flap suture closure is carried out, as described later. If the airway remains unsatisfactory, what creates the problem and how is it most easily and conservatively corrected? Once these questions are answered, a variety of the following corrections are variously carried out, depending on the existent deformity, its location, and its extent.

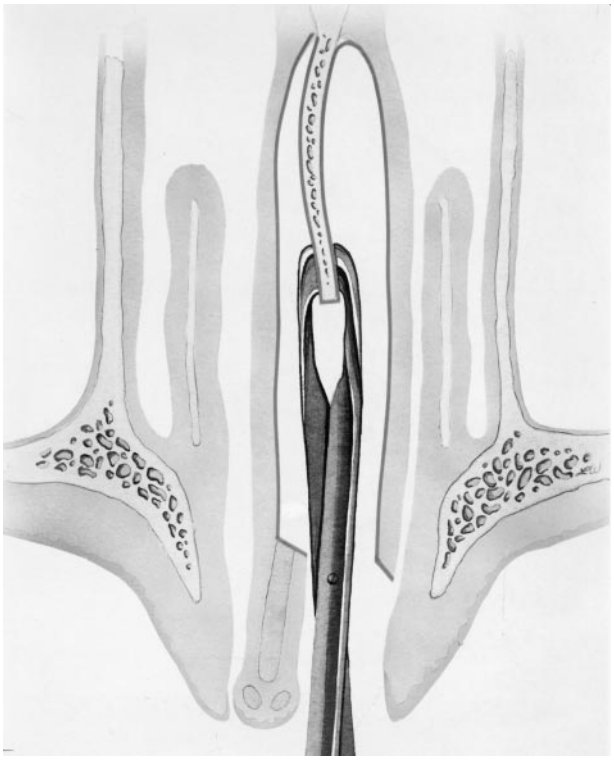


FIGURE 38–63. Removal of obstructing portions of the bony septum with a small biting forceps.

ANCILLARY CORRECTIVE PROCEDURES. Significant obstructing deformities along the floor of the nose (cartilaginous or bony spurs and ridges, which often contribute to drying, ulceration, and bleeding) may require an approach created by elevating a second compartment of periosteum along the floor of the nose (Figure 38–64). The periosteum overlying the premaxilla and maxillary crest is incised, and a submucoperiosteal pocket is created with the elevator. The two compartments (subperiosteal and subperichondrial) are then connected, exposing the maxillary crest deformity for osteotome and/or rongeur removal.

A vertical angulation of the septum, if contributing to deformity and obstruction (often the site of an old fracture), may be removed with a conservative vertical cartilage strip resection or wedge resections, with realignment carried out by absorbable suture technique (Figure 38–65).

Similarly, horizontal septal angulations or spurs are resected by means of conservative horizontal strips or wedges, again leaving all remaining cartilage segments securely attached to the contralateral undissected mucoperichondrium. Thinning the quadrilateral cartilage with *shave excision*

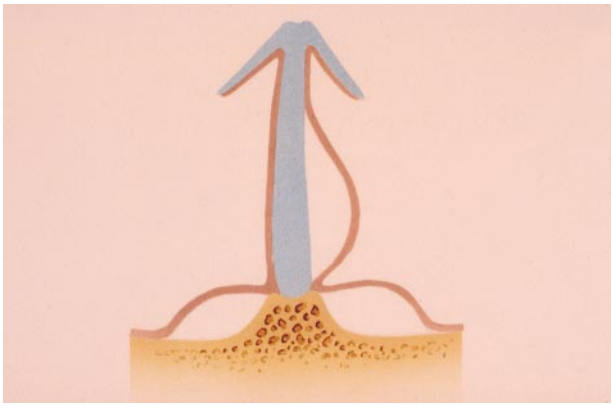


FIGURE 38–64. Elevation of an inferior compartment of periosteum along the floor of the nose to gain better access to deviated bony septal components.

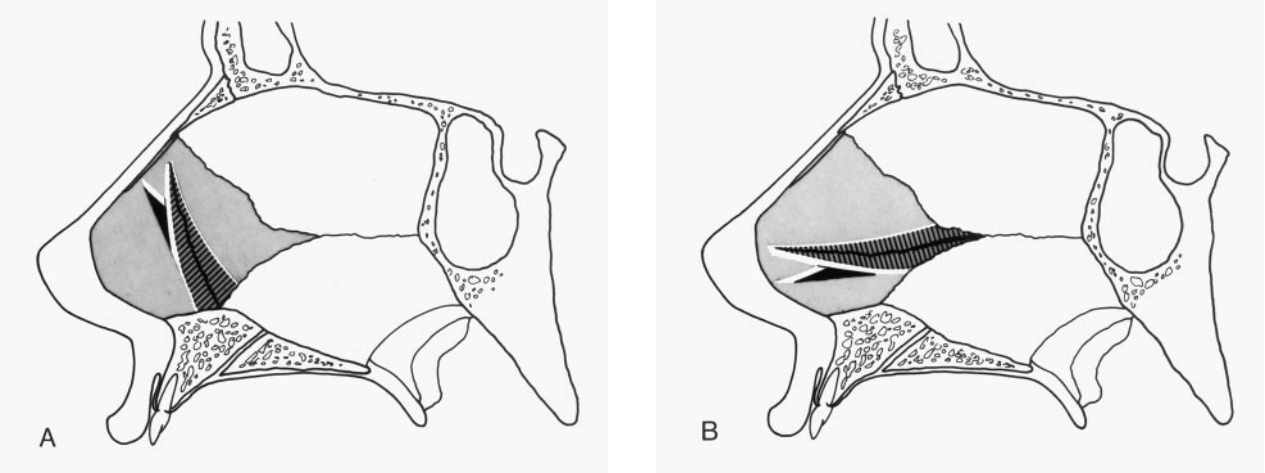


FIGURE 38–65. Vertical (*A*) or horizontal (*B*) angulations of the nasal septum may be removed with wedge resections.

techniques in a horizontal or vertical plane may contribute to airway improvement without significantly altering the supportive strength of the cartilage (Figure 38–66).

Significant dislocation of the caudal septum from the midline, whether developmental or traumatic, requires ingenuity in reconstructive techniques to resist the temptation to resect significant caudal segments, thereby risking columellar retraction and loss of tip support.

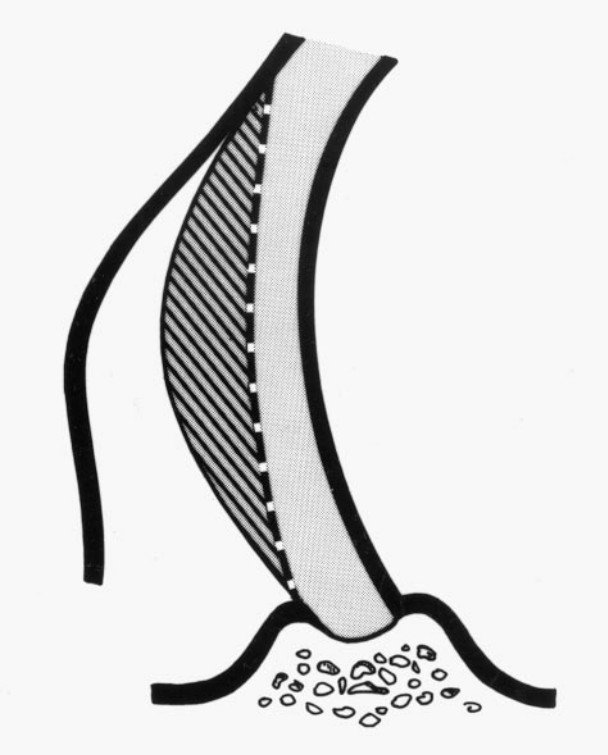


FIGURE 38–66. Shave excision techniques in overly thick quadrilateral cartilages may substantially improve the nasal airway.

A vertical incision (or narrow cartilage excision) cephalic to the point of angulation, combined with freeing the caudal segment from the nasal floor, generally creates a caudal "swinging door" segment that is sutured to the midline columella and occasionally to the midline mucoperiosteum. Further weakening of a deformed caudal segment without total structural sacrifice may be created by generous "cross-hatching," shave excision and even conservative morselization. Limited contralateral mucoperichondrial flap elevation at the caudal segment facilitates its realignment.

An overlong caudal septum, malaligned or not, may require thinning or shortening to facilitate repositioning, particularly during septorhinoplasty procedures.

Occasionally, a previously fractured nasal spine, positioned off the midline, will require refracturing and fixation to the midline to reposition the caudal septum adequately. Unless it deforms the nasolabial angle, the spine is seldom sacrificed, to avoid potential loss of tip support and columellar retraction.

Deformed upper lateral cartilages, the product of maldevelopment or trauma, are occasionally responsible for misalignment of the cartilaginous septum. Inspection and palpation confirm this suspicion. Submucosal dissection and freeing of the upper lateral cartilages from the septum allow further septal mobilization and realignment. This situation, however, is rarely encountered. Preserving the upper lateral septal attachment is preferred.

Final correction of some varieties of badly twisted noses may require realignment of the bony nasal pyramid to mobilize and straighten the septum (and therefore the nose) completely. Such instances are ordinarily the consequence of old trauma.

Persistent airway blockade after careful septal reconstruction is occasionally caused by enlargement of the inferior turbinate to fill partially the concavity created by septal deviation. Return of the septum to midline further diminishes the airway. Partial sacrifice of the inferior turbinate (resection, cautery, freezing, crushing, etc) has never seemed appropriate or physiologic. The author prefers a submucosal elevation of the turbinate tissue with resection of the bulky bone of the inferior concha (Figure 38–67). The turbinate is reattached with absorbable sutures and now occupies considerably less space than originally.

Infrequently, in the most severe quadrilateral cartilage deformities, extensive cross-hatching, crushing, or morselization of the cartilage is used to "break the spring" totally.[26] These procedures are generally reserved for the more caudal portions of the quadrilateral cartilage and are used only in the severest deformities. Bilateral mucoperichondrial flap elevation is usually required, and precise healing is often less predictable than when more conservative measures are employed.

Implantation of autogenous cartilaginous supports in the columella is frequently used, lending support to the tip, correcting retracted columellar deformities, and enhancing the nasolabial angle anatomy.[27] Thin cartilage wafers positioned along the cartilaginous profile line are useful in effacing a slight depression and can be particularly helpful in camouflaging a slightly deviated dorsum without risking breaking the dorsal support with surgical manipulations.

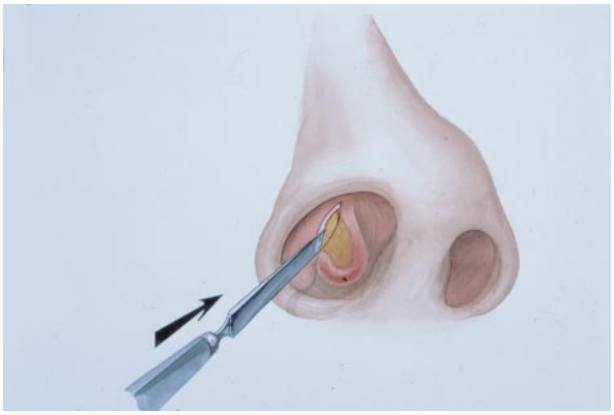


FIGURE 38–67. Reduction of an enlarged turbinate mass is effected by submucosal resection of the enlarged bony concha, preserving the majority of turbinate tissue.

Finally, the complete nasal surgeon should be aware of the septal reconstructive advantages offered by the technique of *external rhinoplasty.*[2] Advantages include a superb direct vision approach, direct anterior appraisal of the deformed elements, and superior suture realignment of reconstructed septal elements.

PRESERVATION OF REALIGNMENT. With progressive use of the foregoing integrated techniques, most deviated septa may be appropriately reconstructed rather than resected, and septal functions may be preserved without embarrassing septal support. Controlling the healing process totally is not possible, but it can be favorably influenced by the suture techniques to be described.

All incisions, including any small flap perforations that occasionally occur, are closed with 5-0 absorbable suture.

A series of similar transseptal transperichondrial through-and-through sutures are now positioned to coapt the septal flap(s), thereby closing all dead space (see Figure 38–56). Hemostasis is promoted, and hematoma is avoided. Further fixation sutures to preserve the realignment of cartilage segments are positioned after the manner of Wright to prevent slippage and overriding. The presence of one intact undissected mucoperichondrial flap provides a great advantage in stabilizing the septal reconstruction.

Septal splints are required only rarely when these methods are used. Telfa gauze rolls are placed in each nasal vault to absorb drainage and splint lightly the dissected flaps and are removed in 12 to 24 hours.

If osteotomies have been performed, or after the occasional external septorhinoplasty, an external cast is applied for 7 days.

REFERENCES

1. Burget GC, Menich FJ. Aesthetic reconstruction of the nose. St. Louis: CV Mosby; 1994.
2. Johnson CM, Toriumi DM. Open structure rhinoplasty. Philadelphia: WB Saunders; 1990.
3. Tardy ME. Rhinoplasty: the art and the science. Philadelphia: WB Saunders; 1997.

4. Sheen JH, Sheen A. Aesthetic rhinoplasty. 2nd ed. St. Louis: CV Mosby; 1987.
5. Goin JM, Goin MK. Changing the body. Baltimore: Williams and Wilkins; 1981.
6. Tardy ME. Surgical anatomy of the nose. New York: Raven Press; 1991.
7. Tardy ME, Thomas JR, Brown RJ. Facial aesthetic surgery. St. Louis: CV Mosby; 1995.
8. Tardy ME, Brown RJ, Childs C. Principles of photography in facial plastic surgery. New York: Thieme; 1992.
9. Goldman IB. Nasal tip correction with special reference to the medial crura. Trans Am Acad Ophthalmol Otolaryngol 1964;68:854–66.
10. Tardy ME. Rhinoplasty tip ptosis: etiology and prevention. Laryngoscope 1973;83:923–9.
11. Parkes ML, Brennan HG. High septal transfixion to shorten the nose. Plast Reconstr Surg 1970;45:487–92.
12. Aiach G, Levignac J. Aesthetic rhinoplasty. London: Churchill Livingstone; 1991.
13. Nolst-Trenite G. Rhinoplasty. The Hague: 1993.
14. Wright WK. Study on hump removal in rhinoplasty. Laryngoscope 1967;77:508–17.
15. Smith TW. The selection of patients for rhinoplasty. Arch Otolaryngol 1971;94:56–8.
16. Frye H. Interlocked stresses of cartilage. Br J Plast Surg 1966;19:276–81.
17. Walter C. Composite grafts in nasal surgery. Arch Otolaryngol 1969;90:622–36.
18. Wright WK. Lateral osteotomy in rhinoplasty. Arch Otolaryngol 1963;78:680–9.
19. Farrior RT. Corrective surgery of the nasal framework. J Fla Med Assoc 1968;45:276.
20. Peer LA. An operation to repair lateral displacement of the lower border of the septal cartilage. Arch Otolaryngol 1937;25:475–90.
21. Fomon S, Gilbert JG, Silver AG, Syracuse VR. Plastic repair of the obstructing nasal septum. Arch Otolaryngol 1948;47:7–23.
22. Frye H. Nasal skeletal trauma and the interlocked stress of the septal cartilage. Br J Plast Surg 1967;20:146.
23. Becker OJ. Problems of the septum in rhinoplastic surgery. Arch Otolaryngol 1951;53:622–41.
24. Farrior RT. Septorhinoplasty in children. Otolaryngol Clin North Am 1970;3:345–64.
25. Tardy ME. Reconstructive surgery of the deviated septum and nose. Memphis (TN): Richards Manufacturing Company; 1982.
26. Frye H. Cartilage and cartilage grafts: basic properties of tissue and its components responsible for them. Plast Reconstr Surg 1967;40:426–37.
27. Denecke HJ, Meyer R. Corrective and reconstructive rhinoplasty. New York: Springer; 1967.

CHAPTER 39

Facial Fractures

Paul J. Donald, MD, Jonathan Sykes, MD

With high-speed auto travel, the increasing participation in sports by people of all ages and both genders, and especially the high incidence of violent crime, facial fractures continue to be important injuries in our society. Management of facial fractures, contrary to the pattern of care in other countries of the world, is, in the United States, spread across the disciplines of oral surgery, plastic surgery, and otolaryngology. Because of the comprehensive training in head and neck anatomy and physiology, the otolaryngologist is uniquely prepared to best deal with these injuries. However, it is vitally important for the specialist to have a working knowledge of dental occlusion and facial esthetics and an understanding of the healing of membrane bone.

The first encounter with the patient with a facial fracture is usually in the emergency department. The patient is often the victim of an accident that involves many body systems, and, almost always, attention to these injuries takes precedence. The exception to this is the initial attention to the airway. Extensive soft tissue contusion, bilateral mandibular body fractures, and Le Fort fractures of the maxilla can all result in airway obstruction. In mandibular fractures, a nasotracheal intubation is appropriate; however, in maxillary fractures, there is always a risk of fracture to the cribriform plate or the fovea ethmoidalis. Intubation by the nasal route presents the danger of intracranial passage of the tube so that oral intubation, cricothyroidotomy, or tracheostomy should be used to secure the airway. Rarely will fractures of the facial skeleton present with life-threatening hemorrhage. The only time in the author's experience in which a problem of serious hemorrhage occurred was in an instance of a Le Fort III fracture in which both of the internal maxillary arteries were severed. Rarely, associated fractures of the temporal bone may rupture the petrous portion of an internal carotid artery, or a fracture through the basisphenoid bone may tear the cavernous portion of the internal carotid artery.

After the patient's condition has been stabilized, careful physical and radiographic examination of the cervical spine must be performed to rule out cervical spinal injury. At this point, a careful evaluation of the other important systems is performed.

Treatment of the patient with facial trauma should include a thorough history and physical examination to determine the location and extent of all injuries.[1] The goal of treatment of patients with craniomaxillofacial injuries should be reconstitution of all injured regions.[2] Both soft tissue and bony injuries should be assessed, and a treatment plan should be established. The goals of treatment should be the restoration of function and appearance. The premorbid form and function of dental, skeletal, and soft tissues should be re-established as much as possible. Recent photographs and dental records, if available, are most helpful to establish the pretraumatic appearance.

If the initial injury involves skin or mucosal lacerations, attempts should be made to use these lacerations when possible for the approach to fracture repair. If no epithelial injury is present, approaches to fractures should attempt to maximize exposure while minimizing scarring and risk to adjacent neurovascular structures.

As thorough a history as possible is taken. Often, however, these patients have suffered multiple injuries and may be either unconscious or intubated. Information may be gleaned from any friends or relatives at the scene or the police or ambulance attendants. Physical examination includes a careful inspection of dental occlusion, examination of the facial contour, and palpation of the facial bones. A computed tomographic (CT)

scan of the facial skeleton is crucial in securing the correct diagnosis.

After fracture diagnosis is completed, a systematic treatment plan is established. Surgical treatment of facial fractures involves adequate exposure, meticulous reduction, and stable fracture fixation. Surgical approaches should use transmucosal incisions or incisions that are camouflaged in relaxed skin tension lines or at the junctions of facial esthetic units.[3] Of course, the incision and approach chosen must not compromise the basic principle of providing adequate exposure for diagnosis and treatment of any fracture. After exposing the facial fracture, meticulous reduction must be performed and maintained until adequate fixation of the fracture can be done. Precise reduction is imperative when rigid fixation is used. Fixation techniques should allow for complete bone healing with reconstitution in three dimensions: height (superior-inferior), width (lateral-medial), and depth or projection (anterior-posterior).

Repair of soft tissue injuries is often as important as fracture treatment in the complete restoration of the patient with craniomaxillofacial injuries. This is especially true in the periorbital region, where injuries such as telecanthus, enophthalmos, and dysopia often accompany the skeletal injury. Complete treatment of facial injuries requires attention to these soft tissue injuries and accurate reattachment of soft tissue fascial layers after fracture repair.

The basic principles of diagnosis and initial treatment of patients with facial fractures are listed in Table 39–1. The principles of fracture repair are listed in Table 39–2. Adherence to these principles will maximize function and appearance.

TABLE 39–1. Principles of Diagnosis and Initial Treatment of Facial Fractures

Establishment and securing airway
Control of bleeding
Detailed history
Careful physical examination

TABLE 39–2. Principles of Fracture Repair

Repair of both skeletal and soft tissue injuries
Use of lacerations when possible
Use of mucosal incisions when possible
Exposure of all fractures adequately for fracture reduction
Reduction of all fractures
Stabilization of fractures
Fixation of all fractures adequately to allow bone healing

DENTAL OCCLUSION

The key to proper fracture reduction is the restoration of the patient's premorbid occlusion. The understanding of dental occlusion is an essential element in the management of facial fractures. Dental occlusion, simply stated, is the relationship of the maxillary to the mandibular teeth. The most important functional relationship is that of their cutting and grinding surfaces. This relationship largely depends on the relative position of the teeth and their angulation to one another. In 1899, Angle described three basic types of occlusion.[4] Certain subtypes of this system, as well as other classifications of occlusion, have been devised, sometimes adding to the confusion rather than clarifying the problem. In the Angle system, the reference point is the relationship of the first maxillary molar tooth with the mandibular molar tooth below it. Each molar commonly has four grinding surfaces called cusps. The cusps adjacent to the tongue are called lingual and those adjacent to the cheek buccal. Those cusps situated toward the oropharynx in the posterior aspect of the oral cavity are described as distal, and those located more anteriorly are mesial (Figure 39–1).

Class I occlusion is the ideal form (Figure 39–2). In this type, the mesial buccal cusp of the first maxillary molar fits in the groove on the lateral or buccal surface of the first mandibular molar tooth. The buccal cusps of the maxillary teeth overlap the buccal surfaces of the mandibular molars.

In Class II occlusion, the mesial buccal cusp of the first maxillary molar is mesial, or in front of the buccal groove of the first mandibular molar (Figure 39–3). It may be over the first mandibular molar's mesial cusp or even between it and the second premolar (bicuspid). In Class III occlusion, the reverse of class II is seen. The mesial buccal cusp of the first maxillary molar now sits over the corresponding mandibular molar's distal cusp or between it and the second molar (Figure 39–4).

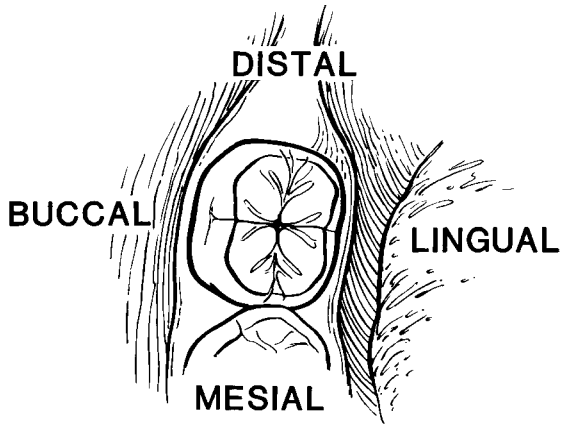


FIGURE 39–1. Orientation of molar cusps. Reproduced with permission from Donald PJ. The surgical management of structural facial dysharmony: a self-instructional package. No. 85100. Washington (DC): American Academy of Otolaryngology-Head and Neck Surgery; 1985.

Usually, these relationships are maintained in the remaining teeth as progression toward the anterior aspect of the dental arches occurs. In Class II and III occlusions, this results in aberrations in the relationship of the incisor teeth that are reflected in characteristic facial deformities.

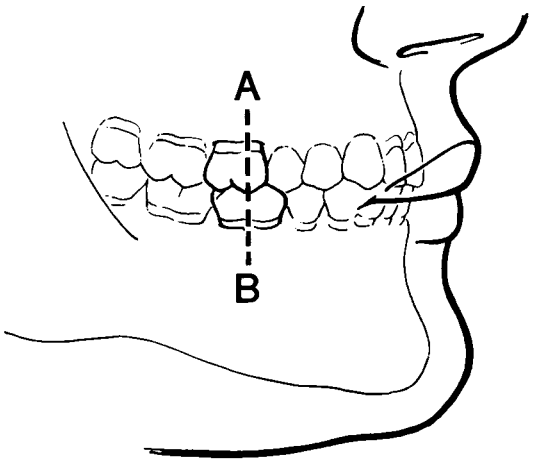


FIGURE 39–2. Class I occlusion. *A*, Mesial buccal cusp of a first molar. *B*, Buccal intercuspal groove of a mandibular first molar. Reproduced with permission from Donald PJ. The surgical management of structural facial dysharmony: a self-instructional package. No. 85100. Washington (DC): American Academy of Otolaryngology-Head and Neck Surgery; 1985.

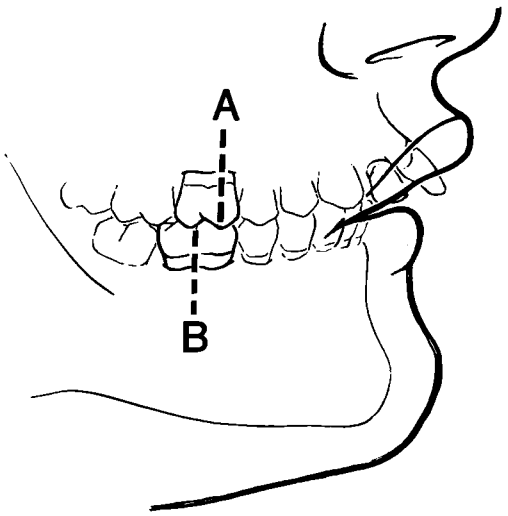


FIGURE 39–3. Class II occlusion. *A*, Mesial buccal cusp of a first molar. B, Buccal intercuspal groove of a mandibular first molar. Reproduced with permission from Donald PJ. The surgical management of structural facial dysharmony: a self-instructional package. No. 85100. Washington (DC): American Academy of Otolaryngology-Head and Neck Surgery; 1985.

In Class II, the mesial or anterior relationship of the maxillary teeth may result in a protrusion of the upper incisors beyond the lower. Not only does one see a jutting forward of these upper central teeth, a condition called overjet, but often the lower incisors bite more deeply toward the palate, a condition called overbite (Figure 39–5). The profile this condition produces is described as "bucktooth" or "weak-chin" look. This look is caricatured in cartoons by such characters as "Andy Gump" and "Sad Sack." If this condition comes about through the lack of mandibular development with a backward positioning of the lower jaw, it is called retrognathism. The setback chin often prevents the upper lip from completely covering the upper incisor teeth. Lack of lip protection of the upper incisors, caused by lip procumbency, renders them more vulnerable to injury; fractured incisors are not an uncommon finding in this group.

In Class III occlusion, the abnormal distal relationship of the maxillary molars, most commonly brought about by an abnormally protrusive mandible, translates into a position of the maxillary

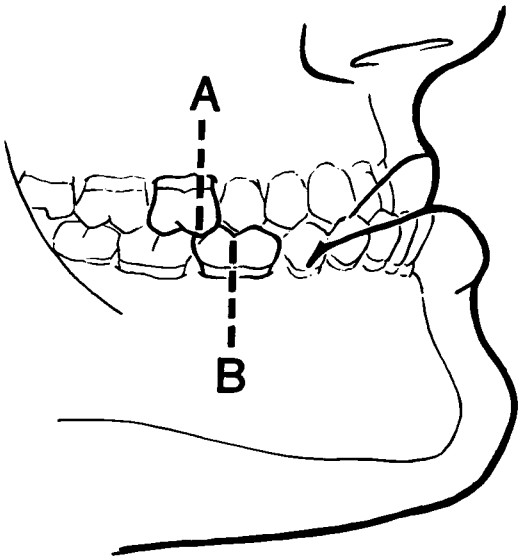


FIGURE 39–4. Class III occlusion. *A*, Mesial buccal cusp of a first molar. *B*, Buccal intercuspal groove of a mandibular first molar. Reproduced with permission from Donald PJ. The surgical management of structural facial dysharmony: a self-instructional package. No. 85100. Washington (DC): American Academy of Otolaryngology-Head and Neck Surgery; 1985.

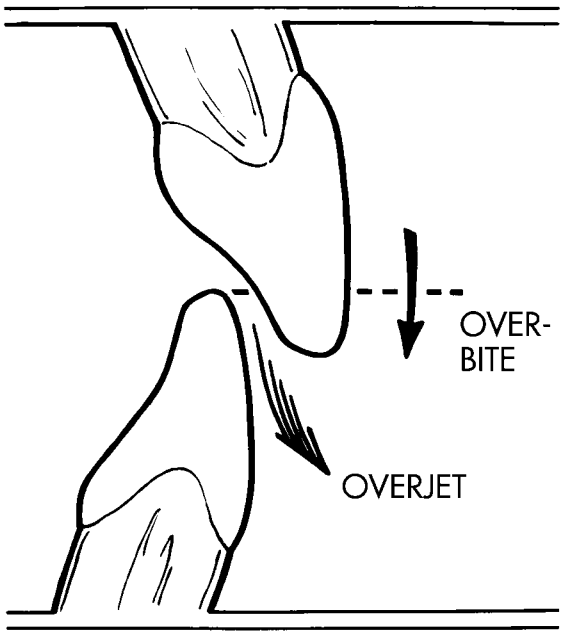


FIGURE 39–5. Cuspal relationships of central maxillary incisor with central mandibular incisor. Directions of overbite and overjet are indicated (*arrows*). Reproduced with permission from Donald PJ. The surgical management of structural facial dysharmony: a self-instructional package. No. 85100. Washington (DC): American Academy of Otolaryngology-Head and Neck Surgery; 1985.

incisors behind or distal to the mandibular incisors. This produces a so-called "bulldog" or "Dick Tracy" profile. The overbite results in the biting of the upper incisors into the gingival lingual sulcus below. The malocclusion is often so severe that the patient has an extremely difficult time chewing solid food. Often the occlusion of two or three cuspal pairs is all that is possible. In edentulous patients, it is often exceedingly difficult to fit them with a denture.

The opposite of the closed bite of overbite deformities may be seen in both class II and III types of malocclusion. This so-called open-bite deformity produces a marked functional disturbance and is extremely unsightly. Even though the molar teeth are in some form of apposition, the anterior teeth never meet.

Both Class II and III malocclusions may come about from discrepancies in development of either the mandible or the maxilla. Class II may arise from an underdeveloped mandible or from an abnormally protuberant maxilla. Similarly, a class III bite may come about as the result of a large, overdeveloped mandible or a retruded or retroplaced maxilla. Retrognathism is classically used to describe a class II type of occlusion and prognathism the class III type.

In addition to aberrations in the anteroposterior relationship of the dentition are deformities involving malposition in the medial and lateral direction, as well as abnormalities involving a lateral tilt of the occlusal plane. The maxillary molar teeth, instead of having their buccal cusps positioned lateral to the buccal face of the opposing mandibular molars, may be biting end to end or even over the lingual cusps of these latter teeth (Figure 39–6). This is called a lingual crossbite deformity. If both arches are symmetric, the maxillary molars of the opposing side will be put into a more buccal relationship with the opposing mandibular teeth, thereby creating a buccal crossbite on this particular side. Because an abnormality in the arch configuration of either or both of the dental arches is often present, it is not uncommon to have a crossbite on one side with a fairly normal occlusal relationship on the other. For instance, a

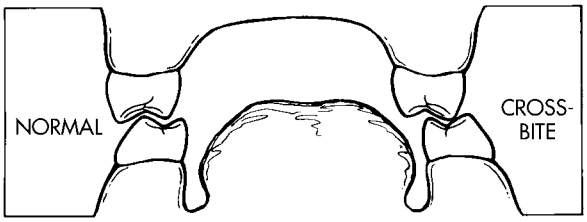


FIGURE 39–6. Coronal view of oral cavity showing normal relationship (*left*) with a maxillary buccal cusp related to the buccal surface of an underlying molar. An example of a lingual crossbite is shown (*right*). Reproduced with permission from Donald PJ. The surgical management of structural facial dysharmony: a self-instructional package. No. 85100. Washington (DC): American Academy of Otolaryngology-Head and Neck Surgery; 1985.

narrow maxillary arch coupled with a fairly normal mandibular arch, such as that seen in hemimandibular hypoplasia, often produces a buccal crossbite on one side and normal occlusion on the other. In patients with hypoplasia, the plane of the bite is commonly abnormal as well. These patients commonly have a diagonal bite, so when they bite down on a tongue blade placed laterally across the teeth, the blade tilts at an angle that is oblique to the horizontal plane (Figure 39–7). Of course, it is not uncommon to see a combination of many of these occlusal abnormalities in any given patient.

FRACTURES OF THE MANDIBLE

Anatomy

The mandible, a horseshoe-shaped bone, is the largest bone of the face. It is divided into regions well described by Dingman and Natvig.[5] The condyle is the portion extending from the mandibular notch to the condylar head, which articulates in the glenoid fossa. The coronoid process is the anterior superior extension of the mandibular ramus projecting above the mandibular notch into the infratemporal fossa. Below the mandibular notch is the ramus, which is composed of thin cortical bone. The angle is a non–tooth-bearing portion of the mandible between the ramus and the body. The parasymphyseal region is composed of the anterior arch of the mandible and is bounded by the two mental foramina. The bodies and the parasymphyseal areas are where the teeth are found. The alveolar ridge or process is composed of thin cortical bone (lamina dura) that encompasses the teeth and atrophies when the teeth are gone.[5] In the edentulous mandible, the tooth-bearing bone resorbs, decreasing mandibular height.

The inferior alveolar nerve enters on the medial (lingual) aspect of the mandibular ramus and passes through its own canal to the mental foramen. While traversing this canal, it gives off sensory innervation to the dentition and gingiva, as well as provides sensation to the ipsilateral lower lip and chin. After a fracture of the angle and body, this canal is often affected. Unless there is a complete

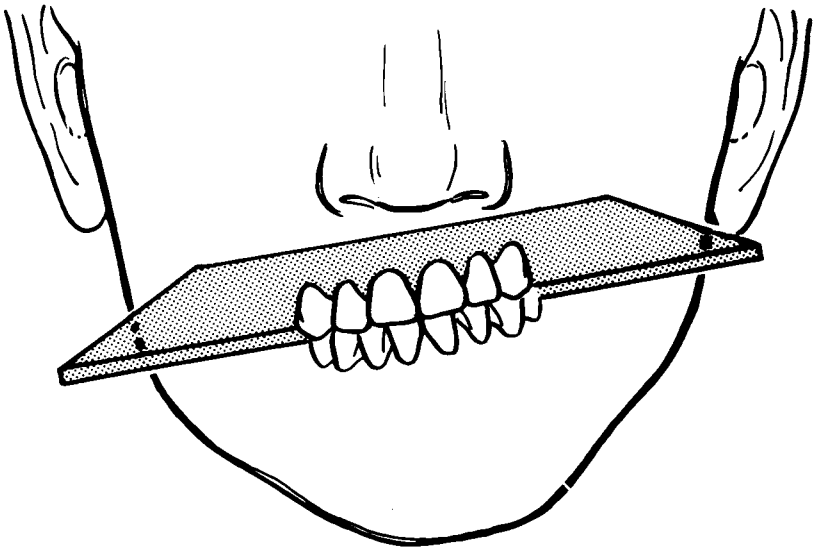


FIGURE 39–7. Obliquely oriented bite. Reproduced with permission from Donald PJ. The surgical management of structural facial dysharmony: a self-instructional package. No. 85100. Washington (DC): American Academy of Otolaryngology-Head and Neck Surgery; 1985.

transection or avulsion of the nerve, full sensation often returns within 9 months to a year.

The muscles of mastication inserting on the mandible include the temporalis, internal pterygoid, external pterygoid, and masseter. All of these muscles contribute to the movement of the temporomandibular joint, a synovial joint with both hinge and gliding action. This joint contains a capsule with a fibrocartilaginous disk. The capsule is densely innervated with proprioceptive and sensory fibers, which are extremely sensitive to subtle changes in movement of one or both joints. A slight alteration in occlusion from muscle spasm or a displaced fracture may alter the central perception of joint position. Feedback loops in the central nervous system force the contralateral muscles of mastication to compensate. This series of events may lead to chronic temporomandibular joint syndrome in an unrepaired or poorly repaired mandibular fracture.

Understanding the various attachments of the aforementioned muscles of mastication, as well as the mylohyoid, geniohyoid, genioglossus, and digastric muscles, is important in understanding the forces of displacement in a mandibular fracture. The floor of the mouth and extrinsic tongue muscles tend to displace fractures posteriorly and inferiorly. The medial pterygoid and masseter muscles act as a sling in the posterior body and angle area and tend to elevate a displaced fragment of the angle or posterior body. Condylar fractures are displaced by the pull of the lateral pterygoid muscle, which rotates and dislocates the fracture medially.[6]

Classification

Fractures of the mandible may be classified according to specific characteristics and anatomic location. A fracture is considered simple when both the external skin and oral mucosa are intact or compound (open) when a laceration in the skin or intraoral mucosa is present. If the patient is dentulous and the fracture line passes into the tooth root, the fracture is theoretically compounded because the periodontal pocket of that tooth often extends to the fracture site. If the fracture is incomplete and involves only one cortex, it is termed "greenstick." The comminuted mandible fracture is one with several fragments of bone. Mandibular fractures are most commonly characterized by anatomic location (Figure 39–8). The most frequent location of fractures of the mandible is the condylar-subcondylar region. Other common sites include the body and mandibular angle. The coronoid process is rarely fractured. An alveolar ridge fracture involves the occlusal surface of a part of the mandible but does not extend into the inferior cortical surface (Figure 39–9).

Mandibular fractures can also be classified as dentulous, edentulous, or pediatric. The latter category is important because of the vulnerability of unerupted dentition and the conical shape of the unerupted teeth, which do not hold a wire ligature well. The accuracy in approximation of the edentulous mandible is not as critical as is the exacting task of establishing the preinjury occlusion of the dentulous mandible. A denture can compensate for minor irregularities in the edentulous jaw.

The final classification may be made according to the stability of the fracture. Vertical instability results from the pull of the temporalis, masseter, and pterygoid muscles. The angle of pull of these muscles will tend to impact a jaw fracture that is obliquely inclined from distal to mesial (posterior to anterior), rendering it stable. On the other hand, if the inclination is the opposite direction, then the forces of these

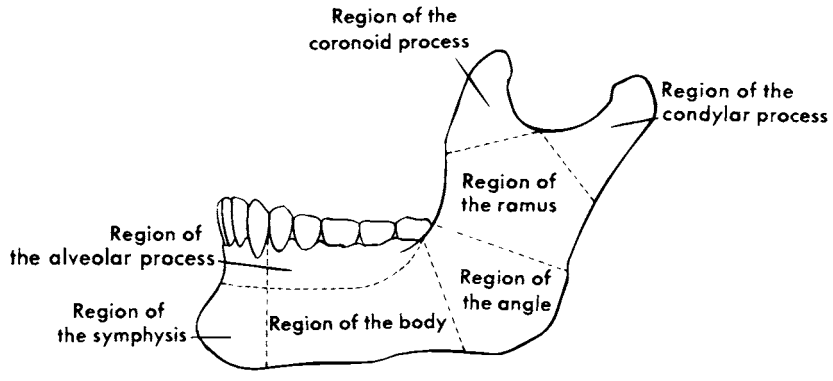


FIGURE 39–8. Schematic diagram of the regions of the mandible. The canine line can be seen to distinguish the parasymphyseal region from the body of the mandible. Reproduced with permission from Dingman RO and Natvig P.[5]

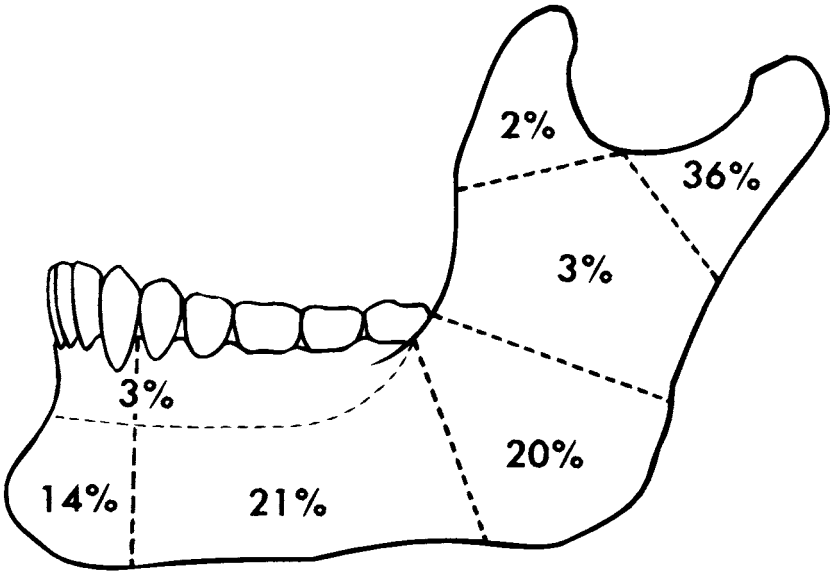


FIGURE 39–9. Schematic diagram of fractures of the mandible according to the incidence of fracture. Reproduced with permission from Dingman RO and Natvig P.[5]

muscles will distract the distal segment in a superior and medial direction. An interesting scenario presents itself regarding an unfavorably inclined fracture in the distal part of the body, near the angle, just anterior to the third molar tooth. While the tooth is present, the fracture will not dislocate, but if the tooth is extracted, the fracture becomes unstable (Figure 39–10). A fracture becomes horizontally unstable by virtue of its obliquity in the occlusal plane. A fracture with an angulation running from the buccal to the lingual surface in a posterior to anterior direction is favorably aligned, whereas that in the opposing obliquity is unstable by virtue of the pull of the mylohoid muscle (Figure 39–11).

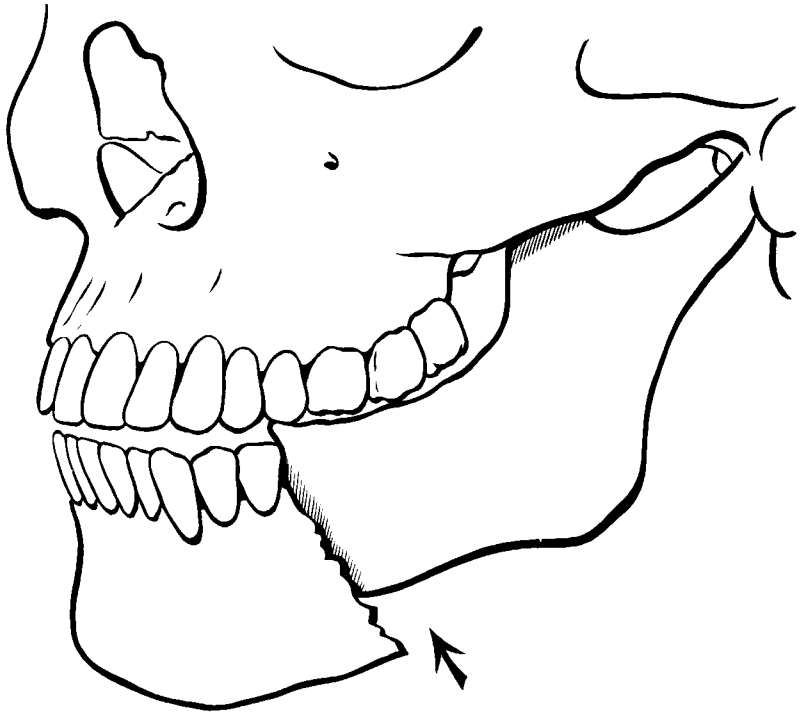


FIGURE 39–10. Diagram illustrating how absence of teeth in the distal segment of an unfavorably aligned mandibular fracture will result in displacement. *Arrow* indicates direction of dislocation. Reproduced with permission from Dingman RO and Natvig P.[5]

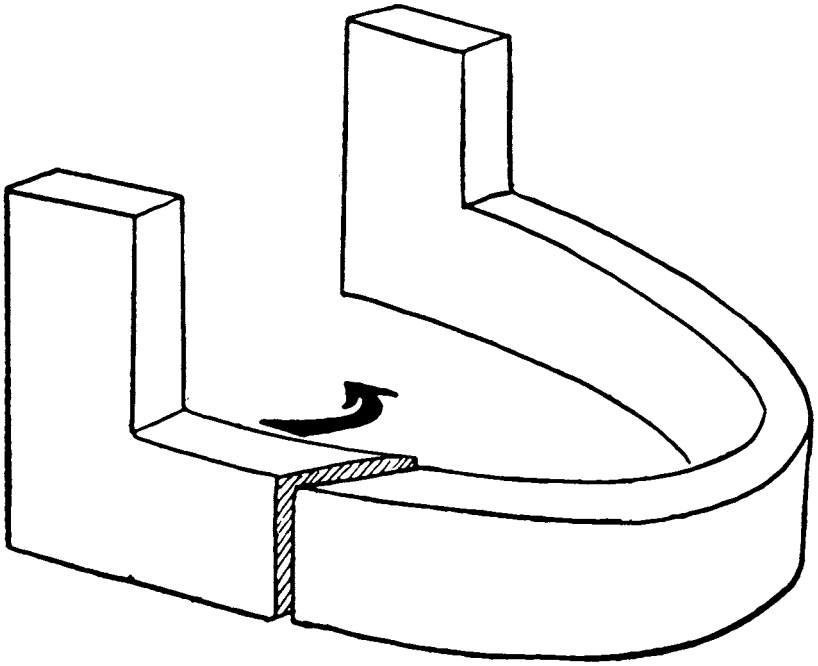


FIGURE 39–11. Horizontally unfavorable fracture. Reproduced with permission from Converse JM.[35]

Diagnosis

A good history and physical examination, along with plain radiographic films, fully delineate the large majority of fractures to the mandible. When obtaining the history, one must determine the nature of the injury, including the weapon used (if any) and the force with which it was applied. Previous orthodontic work should be noted. Any history of previous mandibular or maxillary fractures should also be elicited. Questions regarding the neurologic status of the patient and questions pertinent to the status of the cervical spine should be answered. Not infrequently, a patient presenting with a mandible fracture has other associated facial fractures. Questions regarding pain, hearing, vision, and facial disharmony involving areas other than the mandible are pertinent.

A mandible fracture may cause the patient to complain of malocclusion (eg, "My teeth don't come together right."). Displacement of the maxillary or mandibular teeth of even less than 1 mm can cause severe patient complaints because of the sensitivity of the proprioceptive fibers in the periodontal membranes of the dentition as well as in the temporomandibular joint. The patient often complains of pain in the region of the fracture, most often from the disrupted periosteum surrounding the bone. Muscle spasm also plays a key role in contributing to the pain of mandible fractures, and difficult and painful opening of the jaw, or trismus, is a common complaint. Painful swallowing and sneezing may also occur. Numbness of the lower lip, from an avulsed or badly contused inferior alveolar nerve in any part of its course from the ramus to the parasymphyseal area, is commonly encountered.

Physical examination confirms the patient's complaint of malocclusion. If the fracture is in a tooth-bearing area, loose or missing teeth may be noted. Ecchymosis of the gingiva also indicates the presence of a fracture. Gross displacement of the fragments may be obvious (Figure 39–12). There may be a foul odor in the patient's breath from a combination of stagnant food, blood, and, frequently, alcohol. Bony crepitus and tenderness are also elicited on palpation. A thorough dental examination should be undertaken, including assessment of the presence or absence of teeth and the fractures of dental crowns. Dentures, even if broken, are preserved. The maxillomandibular occlusal relationships are documented. Soft tissue swelling over the point of contact is often found. If there is a missing tooth that is unaccounted for, aspiration is a possibility. In the unconscious patient, evidence of a recently lost unrecovered tooth necessitates a chest radiograph.

Radiographic evaluation can be used to confirm the physical findings and assess the severity of the fracture, including whether it is comminuted and whether there are actually fractured teeth

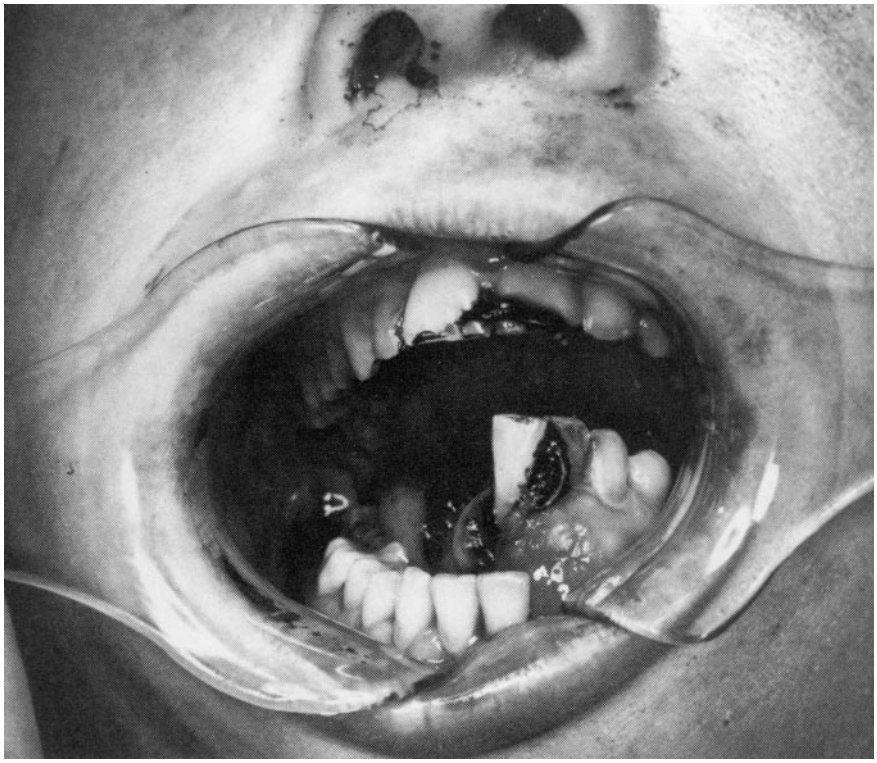


FIGURE 39–12. Patient with a grossly displaced fracture of the parasymphysis of the mandible.

involved. When there is one fracture in the mandibular arch, one must always be aware that there may be an occult fracture in the contralateral ramus or condyle. The panoramic radiograph, the so-called Panorex film, is excellent for viewing fractures of the ramus, angle, and body[7] (Figure 39–13). This type of radiograph is limited, however, in visualizing the symphyseal area as well as telescoping fractures of the condylar area. A modified Towne's view usually visualizes significant fractures of the condyles (Figure 39–14). If a strong suspicion of subcondylar fracture exists, a CT scan will delineate the degree of fragment overlap or displacement into the infratemporal fossa (Figure 39–15). If a parasymphyseal or symphyseal fracture is suggested clinically, but not visualized on plain radiographs, the CT scan will demonstrate it. Dental fractures are best diagnosed with dental occlusal films. If the patient sustains massive trauma, as in a motor vehicle accident or other collision, CT scans of other pertinent anatomic areas should be taken, including full facial and cervical spine films.[8] The scan is usually extended superiorly to rule out any intracranial injury.

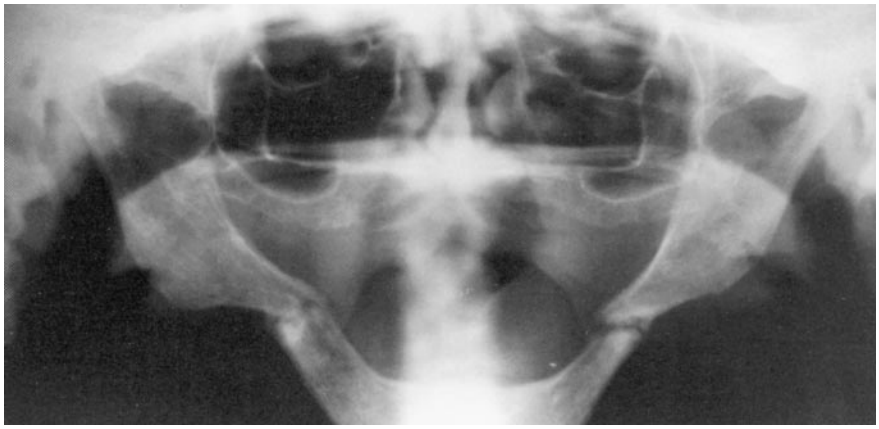


FIGURE 39–13. Panorex film showing a bilateral body fracture.

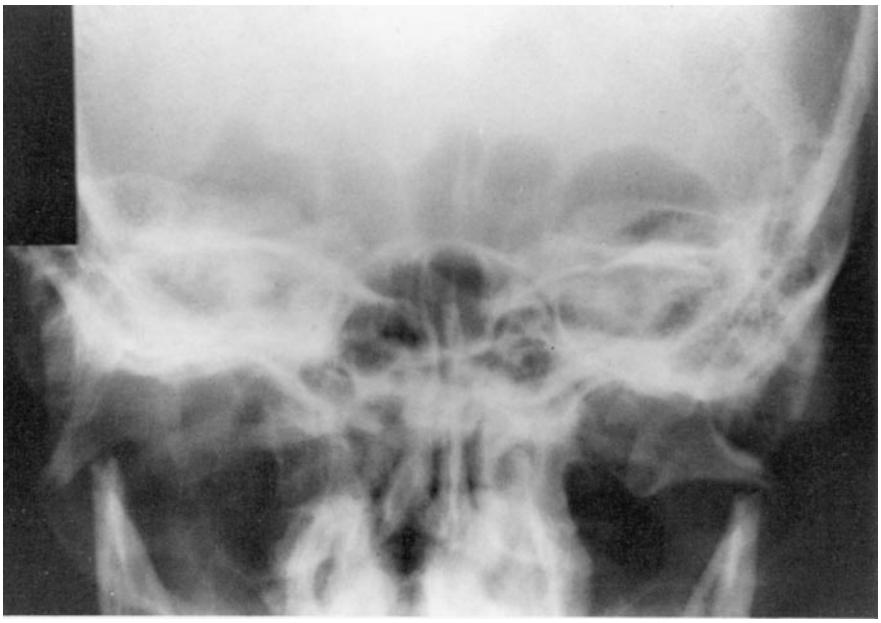


FIGURE 39–14. Modified Towne's view showing bilaterally displaced subcondylar fractures.

Treatment

After an adequate history and physical examination have been obtained and a radiographic assessment has confirmed the physical findings and ruled out any "occult" fracture, treatment is contingent on the following factors: (1) presence or absence of other serious trauma, either intracranial, intrathoracic, or intra-abdominal; clearance of the cervical spine by a qualified neurosurgeon or orthopedist; (2) adequate surgical consent from the patient; (3) a thorough dental evaluation, including preoperative documentation of occlusion; and (4) availability of bridgework or dentures that accompany the patient to the operating room.

Although many authorities recommended surgical intervention as early as possible to prevent infection, there is no conclusive evidence for this.[5,8–10] To decrease the incidence of infection, it is best to administer preoperative antibiotics. Surgery should be done as soon as possible with the previously mentioned factors taken into consideration. It must be emphasized that repairing a mandibular fracture is not, in the true sense of the word, an emergency but rather should be done as soon as possible when safe. Some fractures of the mandible do not require surgical intervention. Most commonly, nondisplaced ramus fractures, as well as some subcondylar fractures, when there is no complaint of malocclusion, should be treated with a soft diet. The

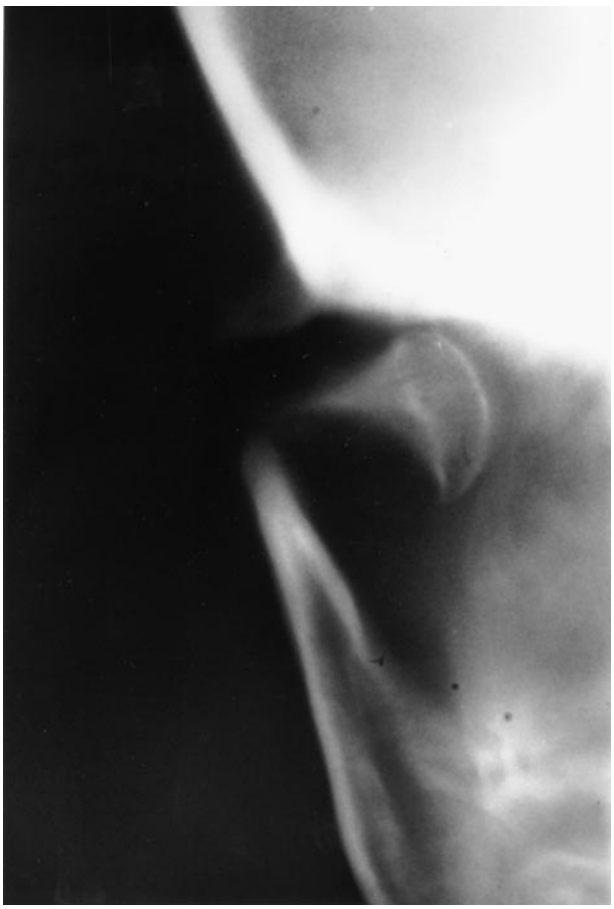


FIGURE 39–15. Linear tomogram demonstrating dislocation of the mandibular condyle into the infratemporal fossa.

remainder of mandible fractures,[5] especially when there is malocclusion, should be treated either by closed or open reduction or, in some cases, both.

The cornerstone of facial fracture repair is still intermaxillary fixation. This method of repair entails the ligation of the teeth of each arch to those that oppose it. According to Dingman and Natvig, ligation of teeth to each other dates back to the time of Hippocrates.[1] The ancients advised fixing the teeth to one another with gold wire or linen thread. The Greek physician Soranos of Ephesus first described the concept of support of the fractured jaw by a barrel-type bandage in the second century AD.[11] These first methods appeared to rely on healing resulting from the wiring of adjacent teeth in the same arch. William of Saliceto, who practiced in Bologna and Verona, first suggested interarch fixation for mandibular fractures in the thirteenth century. He used waxed, twisted silk threads and bound "the injured to the uninjured jaw." It remained for Gilmer, in 1887, to be the first to propose actual rigid interdental and intermaxillary fixation, which he accomplished with iron wire.[12]

Intermaxillary fixation began by separately wiring each tooth of the maxillary and mandibular arches and then connecting these wires one by one as the teeth were brought into occlusion. From this method, certain modifications were made and evolved into the method we know today. The most significant contribution was the development of the arch bar. Grunell Hammond, in 1871, may have described the first arch bar, a firm wire that went around the lingual and buccal surfaces of the teeth. Sauer, in 1882, used a circumferential gold wire with a spring attachment. These evolved into the Erich and Jelenko arch bars we recognize today. The Erich bar, the one most commonly used, is a malleable metal bar that easily adapts to the buccal surfaces of the teeth and has hooks to which intermaxillary wires or rubber bands can be attached. The Jelanko bar is rigid and difficult to accommodate to the dental arches. Only the application of eyelet wires and the Erich arch bar will be described.

Eyelet Wires The eyelet wire is best used for subcondylar fractures, greenstick fractures, and favorably aligned noncomminuted mandibular fractures in patients who are cooperative and compliant. The eyelet wire may also be used as temporary fixation until definitive fixation is done. The wire is constructed by taking a half-length of No. 26 gauge wire, bending it in half, and twisting in a small loop.

The idea of the eyelet wire is to capture two independent teeth on each side of a mandibular fracture and then fix these to two adjacent pairs of maxillary teeth (Figure 39–16). The ends of the wire are directed from the buccal side through the interdental space below the contact point and close to the gum between the pair of mandibular teeth distal to the fracture. In Figure 39–16, A, each wire is brought around the neck of each tooth, and the end of one is passed through the loop and twisted to its mate at a comfortable distance from the loop. The wire ligature now captures the two teeth (Figure 39–16, B). Care is taken to be sure that the wire is pushed below the embrasure of each tooth and cinched around the tooth neck. The pair of teeth mesial to the fracture line is now ligated in the same way. Similar wires are placed in the same fashion around the maxillary teeth that occlude with the opposing wired mandibular teeth. Pairs of mandibular and maxillary teeth are captured with eyelet wires on the opposite side of the arch for additional stabilization (Figure 39–16, C). The patient is placed in premorbid occlusion, and each loop on the mandibular side is individually wired to the loop above. The fracture is now fixed, and the wires are checked on a weekly basis. Occasionally, the interloop wires may loosen or break and may need tightening or replacement. The wire eyelet itself may also come loose and require tightening. Wires are removed at 3 weeks for subcondylar fractures and 6 weeks for others.

Arch Bars Erich arch bars generally come in a roll. The correct length of bar for each arch is measured by placing one end of the bar on the most distal tooth in the arch on which it is to be fixed. The bar is bent around to the space between the central incisors. This length is now doubled, and the bar is cut from the roll. Adjustments are made for missing molar teeth. If significant gaps of two or more teeth are present, the gap can be filled in with a pad of cold cure acrylic pressed into the bar.

Once trimmed to length, the bar is carefully contoured to the teeth by hand. Great care is taken to shape the distal ends to fit around the last teeth in the arch so that the bar will not dig into the cheek. Each bar is placed over the buccal surfaces of the teeth and wired into position with the hooks facing away from the occlusal surface.

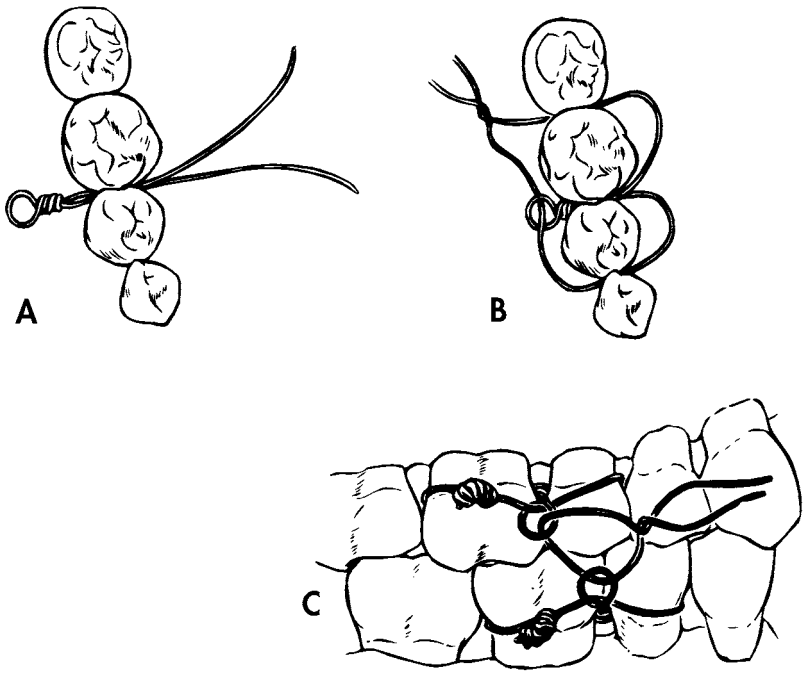


FIGURE 39–16. Technique of placing eyelet wires for intermaxillary fixation. Reproduced with permission from Bernstein L. Fractures of the mandible. Otolaryngol Clin North Am 1969;373–95.

The best teeth for securing the bar are the molars and premolars. The canine has the longest root but has a shape that is not conducive to retention of a wire. The incisor teeth, because of their peglike configuration, hold the wire even more poorly. If enough teeth are present in the arch, the four incisor teeth of each arch are left unligated. For the purposes of orientation, the application of the arch bar to the maxillary teeth will be described first. A 6-inch piece of prestretched 26-gauge wire is passed through the interdental space below the bar (Figure 39–17). Care is taken to prevent injury to the interdental papilla; however, in patients with periodontal disease, this may be unavoidable. The wire is placed around the tooth adjacent to the interdental space. The two ends are twisted together with forceps while a periosteal elevator holds the wire at the level of the neck of the tooth (Figure 39–18). The wire is twisted until tight. The resulting knot is turned into a tight loop and placed away from the lug on the arch bar. Its end is twisted away from the cheek. The molar and premolar teeth are all ligated to the arch bar using the ante-

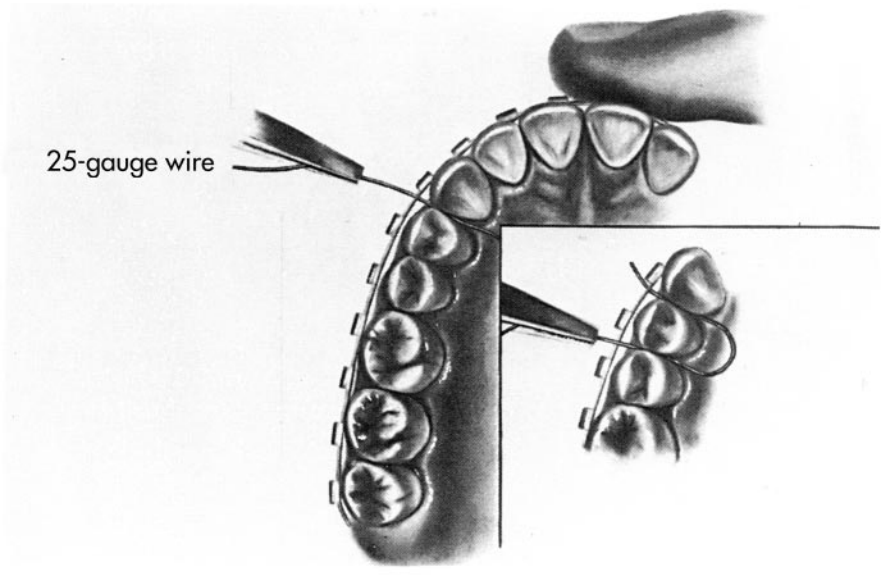


FIGURE 39–17. The wiring of an arch bar to the teeth (see text). Reproduced with permission from Chayra GA et al.[7]

rior wire above the bar and the posterior below formulation. All of the wires are twisted in a clockwise direction (see Figure 39–18). This convention helps in tightening loose wires during follow-up.

Special attention is spent in securing the canine tooth. Its long root embedded in thick bone makes it the key anchor for the bar. Because of its unfavorable shape, adaptation of the wire ligature enhances its holding ability. The wire is placed above the bar, both anteriorly and posteriorly to the tooth, and then looped around the anterior aspect of the bar (Figure 39–19).

Once bars have been placed on both upper and lower teeth, they are secured to one another either with wire or elastics (Figure 39–20). Dental wax may be placed on a protrusive edge that irritates the soft tissues. The patient is also given a stick of wax and instructed how to apply it to irritating wire or lugs. All patients are followed on a weekly basis to check for bar stability and tightness of the interarch wires. Loose wires are tightened and broken wires replaced. Additionally, careful attention is paid to dental hygiene and nutrition. Already carious dentition, a condition not uncommon in patients with facial fractures, is usually the reflection of past neglect. Careful and insistent instructions in proper brushing technique and use of a "water pick" type of device are essential features in the follow-up regimen. Dental caries can lead to the formation of dental abscess during the healing phase, which, in turn, results in osteomyelitis of the jaw.

Nutritional advice is important. The average weight loss in our patient population following intermaxillary fixation is 15 lb. A booklet on dental hygiene and nutrition is supplied to each patient, with instruction on brushing technique and diets that can be employed while in fixation. Of course, only food with a consistency that can be sucked in the free space around the back of the teeth can be used. Balanced high-calorie supplements such as Ensure-Plus® and Sustacal® can be used to augment caloric intake. These products come in both liquid and pudding forms and are quite palatable. When the patient is fixed into occlusion by either Erich arch bars or eyelet wires, it is important to supply the patient with wire cutters so that if he or she vomits, the interocclusal connection can be cut to prevent aspiration.

Gunning Splints In the edentulous patient, the patient's reduced mandible can be maintained by the use of Gunning splints. In 1861, T. P. Gunning of New York was, according to Dingman and Natvig,[5] the first to describe the use of intermaxillary splints, which he fabricated from vulcanite. To make the

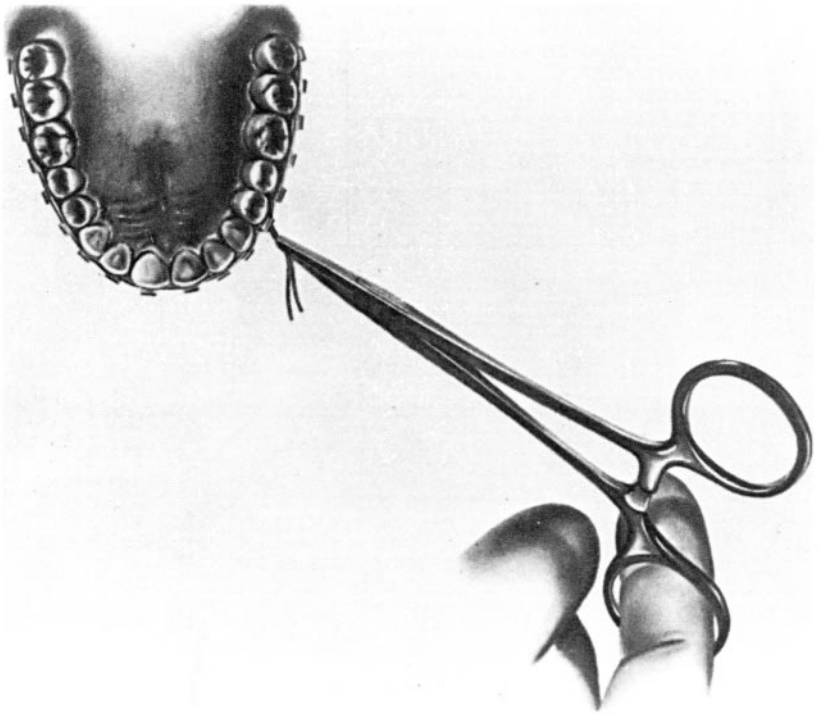


FIGURE 39–18. Wire twisting to lock the arch bar into position. Reproduced with permission from Chayra GA et al.[7]

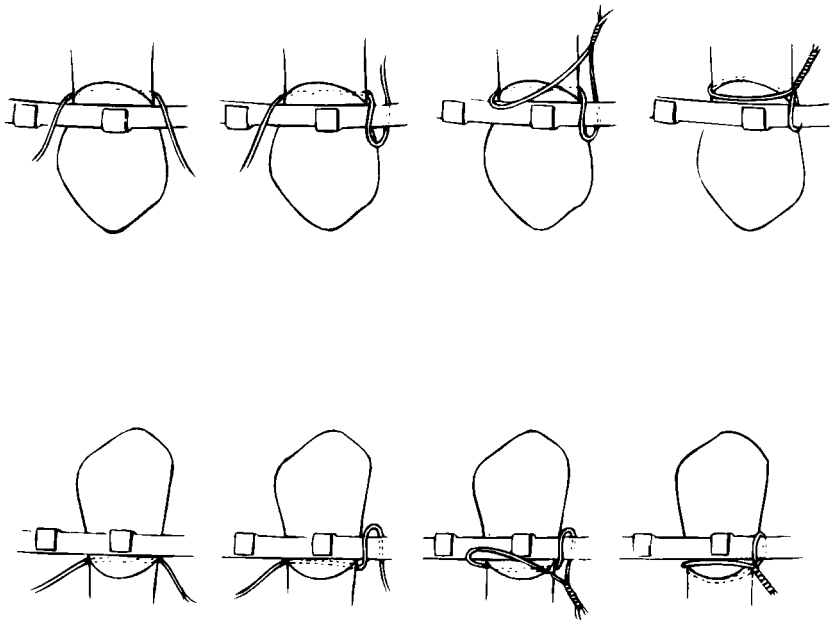


FIGURE 39–19. The special adaption for wiring a canine tooth. Reproduced with permission from Chayra GA et al.[7]

splint, a dental impression of the jaws is taken with an impression compound such as alginate and poured up in stone. The stone model of the mandible is cut at the fracture line and realigned in the normal anatomic position and fixed with sticky wax. An impression of the realigned model is taken, and a hot-cured acrylic stent is made. A stent is also made of the maxilla. A flange and corresponding groove are constructed so that the maxillary splint can fit into the mandibular splint in a lock and key type of articulation. Care is taken so that the normal preocclusal relationship is established between the mandible and the maxilla. An arch bar is imbedded into the splint prior to hardening so that the splints can be ligated together.

At the time of surgery, the mandibular fracture is reduced and the splint fixed to the jaw by at least four circum-mandibular wires. A pair of drop wires

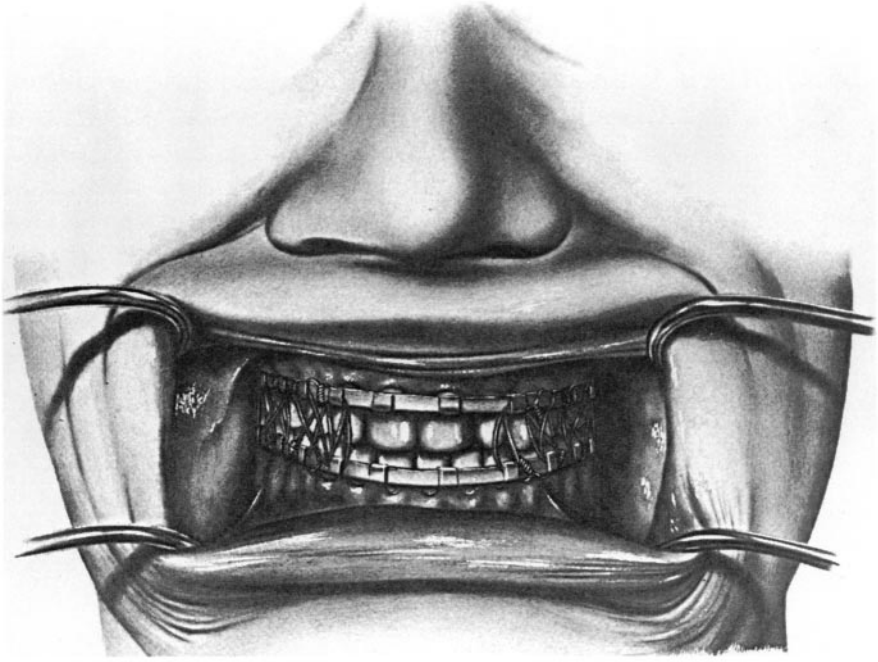


FIGURE 39–20. Maxillary and mandibular arch bars applied and intermaxillary wires in place. Reproduced with permission from Chayra GA et al.[7]

suspends the upper splint from the bone of the piriform apertures and from the zygomata by wires that encircle the arches. Once the flanges on each side of one splint are fitted into the slots of the corresponding splint, then the lugs of the maxillary and mandibular arch bars are wired together. Alternatively, the splint may be fixed with screws that pass through the splint and the unfractured alveolar ridge.

A cruder type of splint can be made in the operating room at the time of fracture reduction in the edentulous patient using cold cure acrylic. A lock and key arrangement is constructed, and arch bars are pressed into the lateral surfaces of the acrylic before it cures. This is a less precise splint and less comfortable for the patient. Any denture, even one that is broken, will serve well as a Gunning splint. Even a loose fitting denture will provide adequate stabilization of the fracture.

Fractures of the Parasymphyseal Area These fractures tend to occur in an oblique line, sometimes even approaching the midline from the mental foramen (Figure 39–21). The mental foramen is the most commonly involved site of a parasymphyseal fracture. A "pure" symphyseal fracture, that is, a fracture through the middle incisors, is rare. Either an ipsilateral or contralateral fracture of the angle, ramus, or condyle often accompanies unilateral parasymphyseal fractures. Parasymphyseal fractures may present with a crossbite on the involved side from the posterior pull of the mylohyoid. This is especially true when there is a concomitant ipsilateral fracture of the ascending portion of the mandible. Loose anterior dentition, ecchymotic mucosa, and grossly mobile fractured segments are the hallmarks of these fractures.

The presence of posteriorly displaced bilateral parasymphyseal fractures constitutes an airway emergency. An adequate airway should be established promptly by placing the fractured anterior mandibular segment in one's forefinger and thumb and projecting it anteriorly. Often an emergency tracheostomy or cricothyrotomy is indicated.

Treatment of a parasymphyseal fracture with only intermaxillary fixation in these cases is not usually adequate. It is often necessary to stabilize these fractures with interosseous wires, plate, or a lingual splint to prevent a permanent crossbite.[5,13–15] The approach to open reduction of these fractures can be either intraoral or external but is usually internal. Care must be taken to avoid avulsing the mental nerve by way of either approach.

Fractures of the Body and Angle of the Mandible Fractures of the mandibular body in patients with good dentition have the advantage of strong articulating mandibular and maxillary teeth. Intermaxillary fixation stabilizes many of these fractures sufficiently, alleviating the need for open reduction. The forces generated by the pterygomasseteric sling may distract these fractures, requiring open reduction. Once closed reduction is accomplished by means of Erich arch bars, a decision can be made whether the repair is stable. If not, an open reduction can be approached by means of a gingivobuccal incision intraorally.[16] Rarely, an extraoral approach will be necessary, and a modified Risdon incision is made with care taken to preserve the marginal mandibular branch of the facial nerve. Angle fractures may be treated the same way as body fractures. If the oblique line of fracture is in a favorable direction, closed reduction is adequate.

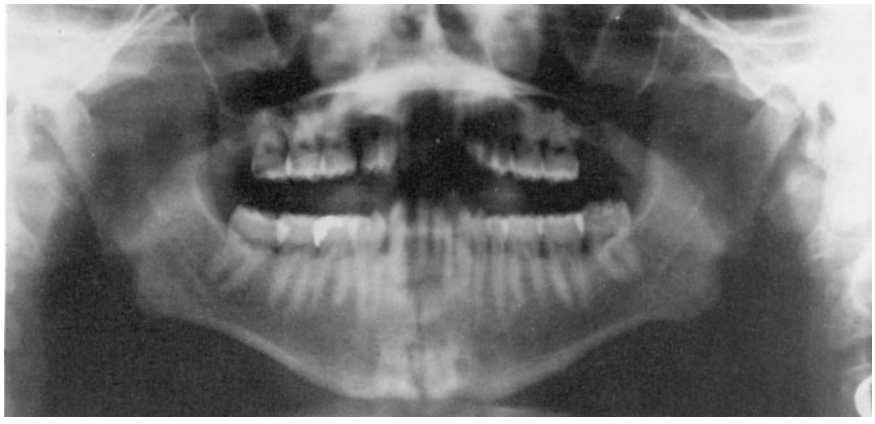


FIGURE 39–21. Parasymphyseal fracture.

When there is an "unfavorable" fracture, open reduction is often necessary.

Fractures of the Condylar Process and Ramus Subcondylar fractures are usually handled by closed reduction. Fractures of the mandibular ramus can be treated in a similar fashion to condylar fractures; that is, if there is displacement of the fractures and malocclusion results, closed reduction may be all that is needed. If there is a massive comminution, severe telescoping, or just displacement that is not adequately reduced by closed reduction, open reduction should be undertaken to avoid occlusal discrepancies. Severe condylar fractures are defined as (1) telescoping greater than 1.5 cm, (2) a condylar head displaced out of the condylar fossa, or (3) bilateral subcondylar fractures associated with a Le Fort or maxillary fracture.

The open approach is facilitated by an incision in the preauricular area extending into the temporal hairline, going deep to the temporalis fascia.[5,17] A flap of superficial temporalis fascia should be raised and the root of the zygoma followed down to the fracture site. The trunk of the facial nerve is avoided by staying on the zygomatic process of the temporal bone, tracking down to the glenoid fossa. With a gentle retraction of the superior parotid tissues, the fossa can be exposed and the fracture reduced by means of either a wire or a miniplate. Alternatives to repair of a condylar fracture include a Kirschner wire placed by way of the angle of the mandible projected superiorly toward the glenoid fossa.[18,19] External pin fixation with a Morris biphase and plating techniques have also been described.[20,21]

Open Reduction Before performing open reduction of a mandible fracture, it is always advisable, if possible, to place the patient in closed reduction to establish the occlusal relationship of the teeth. It is folly to attempt an open reduction of any sort when the maxillomandibular dental relationships have not yet been established.

Surgical Approaches Intraoral approaches have the advantages of scar camouflage and a more direct approach to fracture reduction and fixation. These approaches require more elaborate instrumentation to allow precise fracture fixation. Intraoral incisions allow less direct visualization of fractures; adequate exposure and accurate reduction through this approach require an experienced surgeon. Extraoral approaches have the advantage of increased exposure and increased visualization in the region of the posterior part of the body, angle, and ramus. External exposure requires a cervical incision; however, this involves potential risk to the marginal mandibular branch of the facial nerve.

The incision and approach chosen for a given fracture of the mandible depend on several factors. These include the location and type of the fracture, the available instrumentation and technology, and, most important, the surgeon's comfort with the given approach. The choice for fractures of the symphyseal and parasymphyseal region is intraoral. Linear posterior fractures may be approached intraorally, whereas more comminuted fractures or fractures with significant bone loss usually require an extraoral approach. The approach chosen should allow adequate exposure to diagnose, reduce, and immobilize the given fracture.

The symphyseal and parasymphyseal regions of the mandible are easily approached through either an intraoral or an extraoral route. The intraoral incision is made from canine tooth to canine tooth, leaving an adequate mucosal cuff for closure of the incision (Figure 39–22). Subperiosteal dissection is made, identifying and preserving the mental nerves. The symphyseal and parasymphyseal regions are also easily approached through an external submental incision oriented in relaxed skin tension lines. Reapproximation of the mentalis muscle must be carefully performed to prevent postoperative ptosis of the soft tissues of the chin.

The approach to the mandibular fracture posterior to the mental foramina depends on the surgeon's experience and the degree of the fracture.[12] Intraoral approaches to the body and angle are best performed by making a gingivobuccal incision immediately adjacent to the fracture. Repair of posterior mandibular fractures requires surgical experience and advanced technology to achieve fracture reduction fixation. Fracture fixation is achieved by a transbuccal system placed through transcutaneous facial stab incisions (Figure 39–23). The extraoral approach to fractures of the mandibular body, angle, or ramus is made through a transcervical incision two fingerbreadths below the angle of the mandible. Care is taken to elevate the marginal mandibular branch of the facial nerve to prevent injury. This approach is preferred when the patient has signifi-

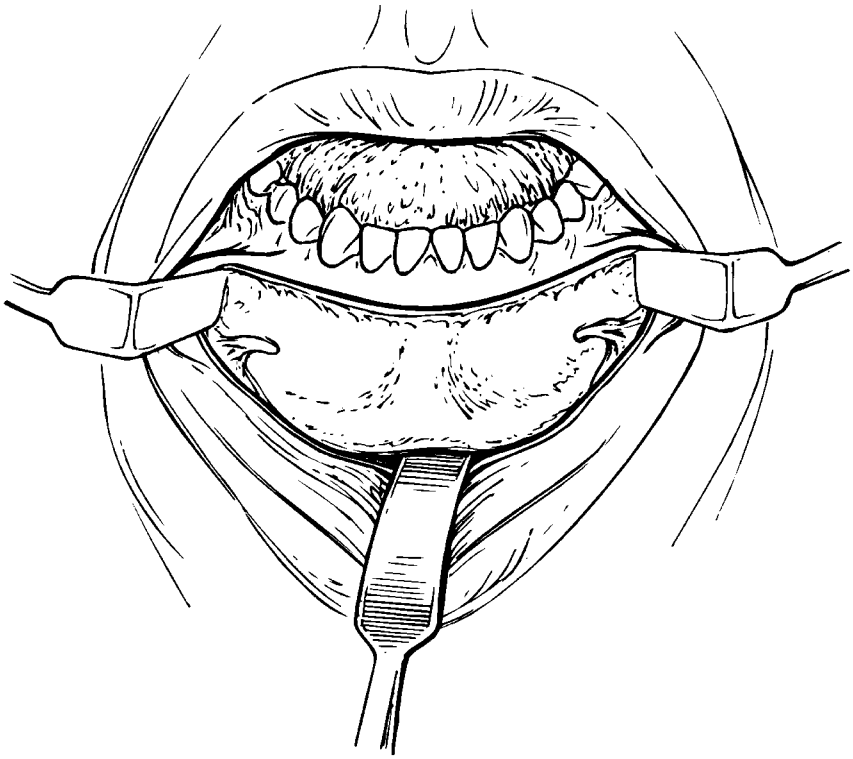


FIGURE 39–22. Schematic diagram of intraoral approach to symphyseal and parasymphyseal fractures. The dissection is subperiosteal and isolates and preserves the neurovascular pedicles from the mental foramina.

cant comminution of the fracture or bone loss. In each of these cases, a larger reconstruction plate may be required for repair. The external approach allows greater exposure for placement of large reconstruction plates.

As instrumentation and technology have progressed, treatment of mandibular fractures has evolved, but the goals have not changed. These goals include anatomic and functional stability of the mandible. Reduction and stabilization of any mandibular fracture should result in a pain-free function of the mandible without any eventual changes in the temporomandibular joint.

The strength of any bony fixation must be adequate to overcome any forces that will act on the repaired bone during function.[22] As previously mentioned, former techniques used wire fixation with interosseous wiring.[13,23] This method allows the bone to heal indirectly. Wire fixation is rarely used today. The only theoretical advantage of wire fixation for repair of mandibular fractures is the possibility of increased flexibility in cases with significant bone loss or comminution. In these situations, if rigid fixation is applied without proper restoration of premorbid occlusion, the result can be either poor occlusion or eventual change in the temporomandibular joint. If interosseous wiring is used, intermaxillary fixation should also be used for approximately 6 weeks for stable bone repair.

The rationale for the use of rigid internal fixation for repair of mandibular fractures is well documented.[13,24] With interosseous wire fixation, there is a clinically proven decrease in mouth opening seen 6 months after release of intermaxillary fixation.[15–18] Additionally, experimental and clinical evidence against wire fixation includes muscle atrophy[18] and histologic changes in the temporomandibular joint[19] after prolonged periods of intermaxillary fixation. Although an increased rate of infection has not been conclusively shown with intermaxillary fixation and interosseous wiring, the increased bone movement with nonrigid fixation makes this a theoretical consideration. As knowledge and technology have progressed, rigid internal fixation has become the standard in most centers for treatment of mandibular fractures. This allows the direct form of bone healing. Use of this type of bone repair requires surgical experience, advanced technology, and patient compliance. Direct bone healing can be achieved only with rigid fixation. If properly applied, stable rigid fixation of the mandible obviates the need for intermaxillary fixation.

Rigid fixation of the mandible can be performed by a variety of methods. The fracture is usu-

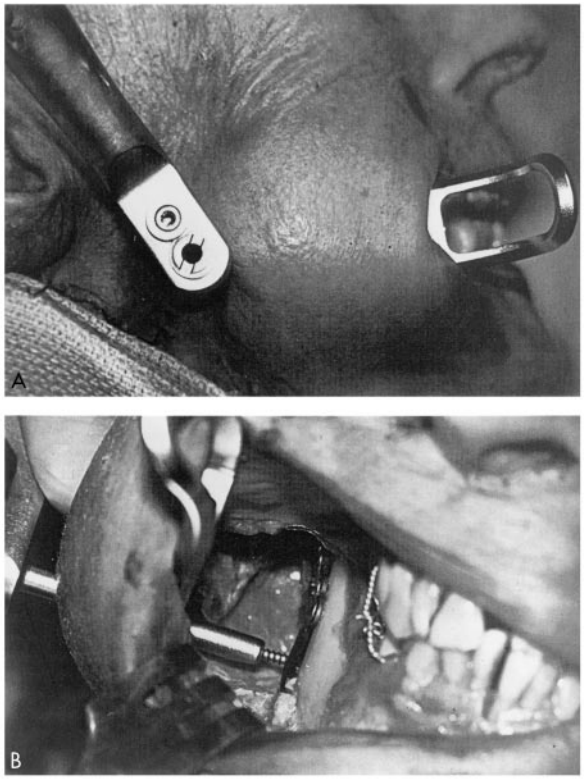


FIGURE 39–23. *A*, Lateral view of a cadaver showing use of the transcutaneous-transbuccal system approach for posterior fractures of the mandible. *A*, Stab wound incision has been used to insert the transbuccal system. *B*, Intraoral view showing use of the transbuccal system to place a plate and screws for fixation of a fracture of the angle of the mandible.

ally first reduced and the teeth put into premorbid occlusion by placing the patient in intermaxillary fixation. The fractures are then directly approached (with intermaxillary fixation in place), and anatomic fragment reduction is obtained. Rigid fixation is then performed. This can be accomplished with an inferior mandibular fracture plate (2.4 to 2.7 mm) and a superior monocortical two-hole tension band. The intraoral incision is retracted with a special intraoral retractor that is fitted to an extraoral device that will maintain a trochar passed through a stab incision in the cheek over the fracture site. Figure 39–24 illustrates the use of the device developed by Stryker-Leibinger (Kalamazoo, Michigan) for this purpose. The hollow trochar enables the passage of a plate-grasping device that itself is hollow and permits the passage of the drill bit. The plate is contoured to conform to the surface of the mandibular bone at the fracture site. With the plate in position, the screw holes are drilled, and the bicortical screws are passed. Two screws, as a minimum, are placed at each side of the fracture line.

In 1973, Michelet et al introduced the use of small monocortical plates for the fixation of mandibular fractures.[25] Champy et al popularized the use of these plates and emphasized the importance of lines of osteosynthesis (Figure 39–25).[26] The lines of natural compression forces at the angle and distal body of the mandible are exerted at the inferior border of the mandible, whereas the lines of tension or distraction occur at the superior border at

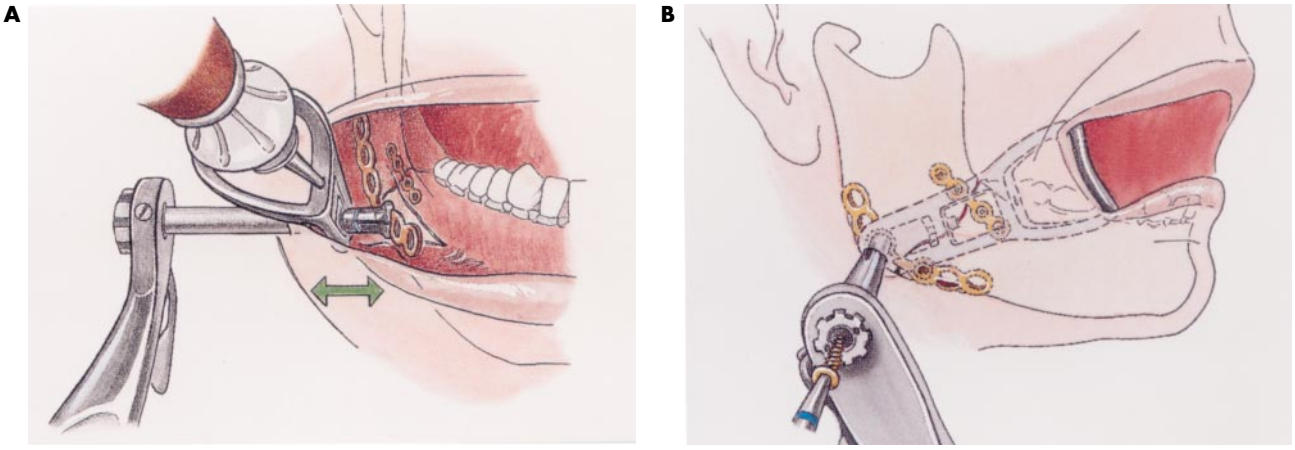


FIGURE 39–24. Transbuccal system for intraoral plating. *A*, Plate holder passed through a stab wound in the cheek. The buccal retractor is being secured to the plate holder. *B*, The drill hole has been made through the plate holder, and now the screw will be placed to fix the plate to the mandible.

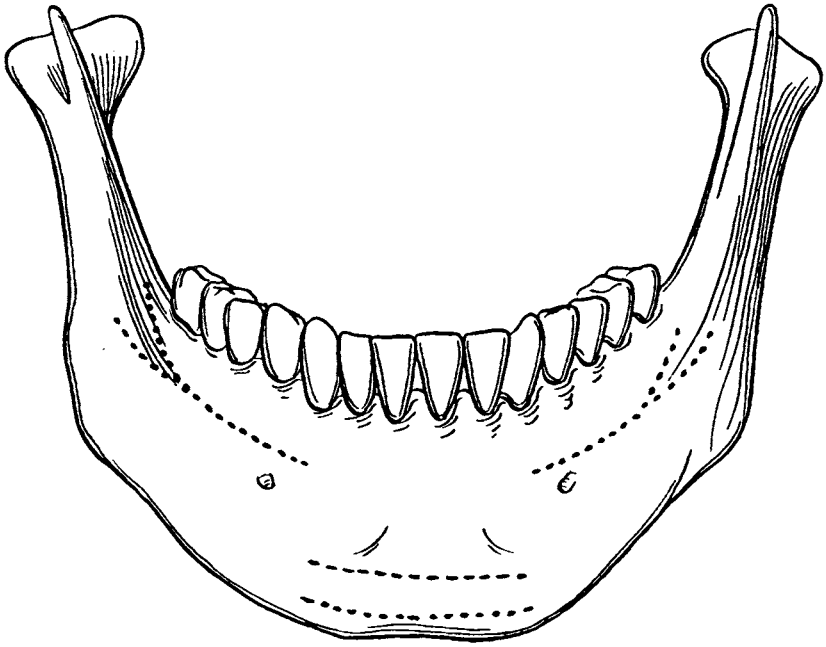


FIGURE 39–25. Lines of osteosynthesis on mandible. Reproduced with permission from Larsen OD and Nielsen A.[24]

the so-called lines of osteosynthesis (Figure 39–26). At the angle, only one plate is needed, but two plates are required at the parasymphyseal area because of the two lines of osteosynthesis. The plates are 0.9 mm thick and have either 4 to 6 holes or 8 to 16 holes. The diameter of the screw holes is 2.1 mm, and the holes are beveled at 30 degrees. Their rigidity and tensile strength are well within the forces of mastication and other forces normally encountered in mandibular activity. In angle fractures, the plate is placed at the external oblique line and in the symphyseal area, one at the inferior line and one in the superior line of osteosynthesis (Figure 39–27). The plates are placed in such a fashion to avoid contact with the tooth roots, although such contact is of no consequence in most instances. Care must be taken to avoid overtightening of the screws as this will produce microfractures and destabilize the fixation (Figure 39–28).[22,27]

Another means to achieve stable rigid fixation is the placement of lag screws.[20] This method employs a basic principle of woodworking and applies compression across the fracture line with minimal direct bone healing. This technique requires adequate exposure and subperiosteal undermining to allow placement of long screws that engage sufficient bone for fixation.

Lag-screw fixation requires that there will be sufficient obliquity to the fracture line to allow at least two, preferably three, lag screws to be placed at a significant distance from one another that will catch both mandibular cortices, thus stabilizing the reduction. The exception to this is the mandibular angle, where one screw is sufficient to achieve ade-

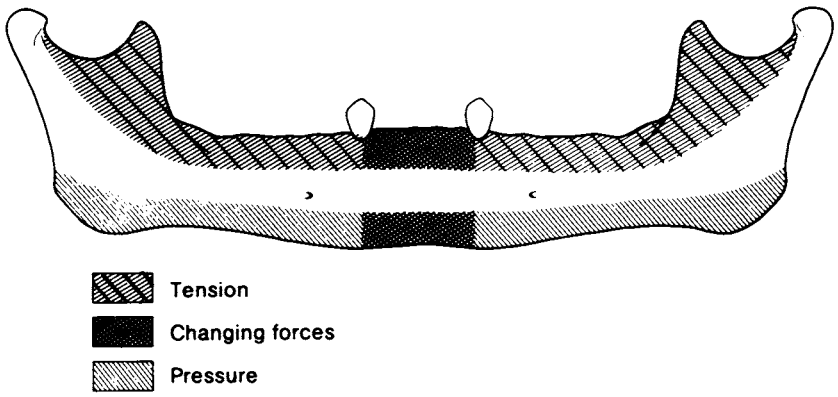


FIGURE 39–26. Forces of compression at the inferior border of the mandible and forces of tension at the superior border with a mandibular fracture.

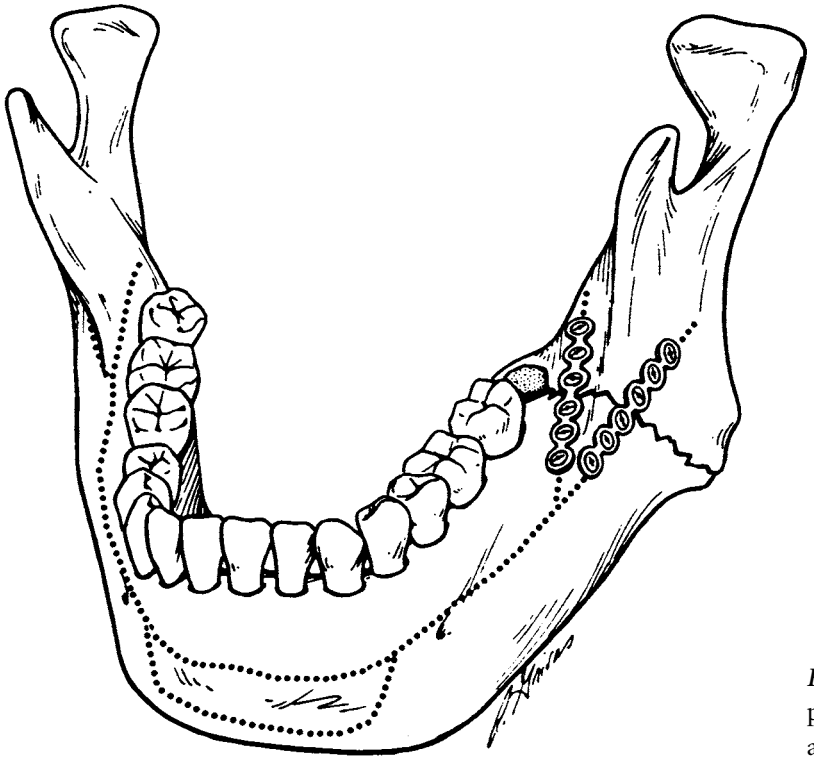


FIGURE 39–27. Champy miniplates placed on the mandible to secure an angle fracture of the mandible.

quate fixation. This technique is not applicable to all mandibular fractures and takes a degree of understanding of the dynamics of mandibular function and skill in inserting the screws appropriately. Through an intraoral approach, the fracture line is exposed, and the fracture is reduced and then maintained by arch bars or eyelet wires. For anteriorly located fractures in the parasymphyseal and mesial body, an intraoral approach can be done. For those fractures near the angle, the transbuccal approach is employed. A trochar is introduced through a stab incision in the cheek, and the penetrating point is removed prior to introducing the drill bit (Figure 39–29). A drill guide is placed

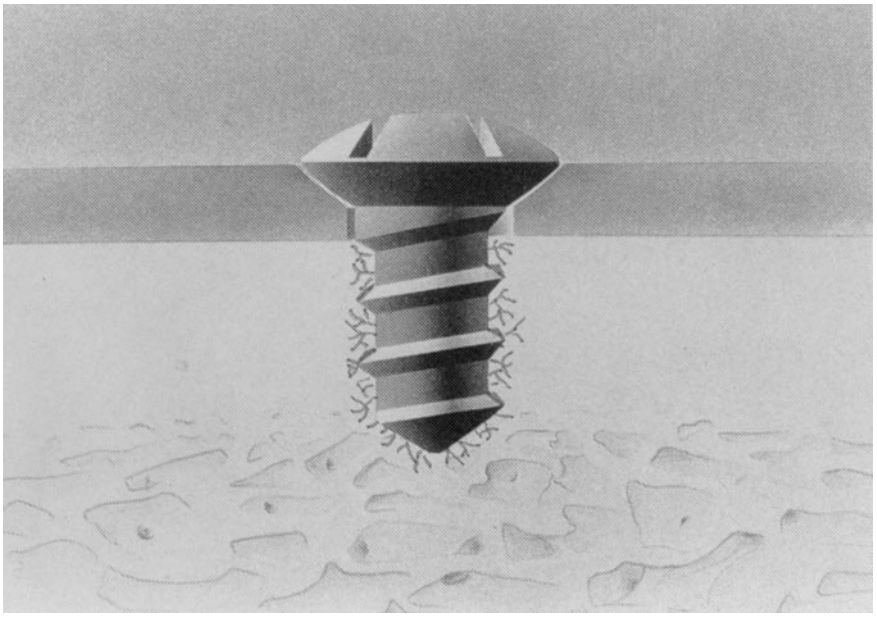


FIGURE 39–28. Overtightening of the screws in Champy miniplate fixation may cause microfractures. Reproduced with permission from Glineburg RW et al.[27]

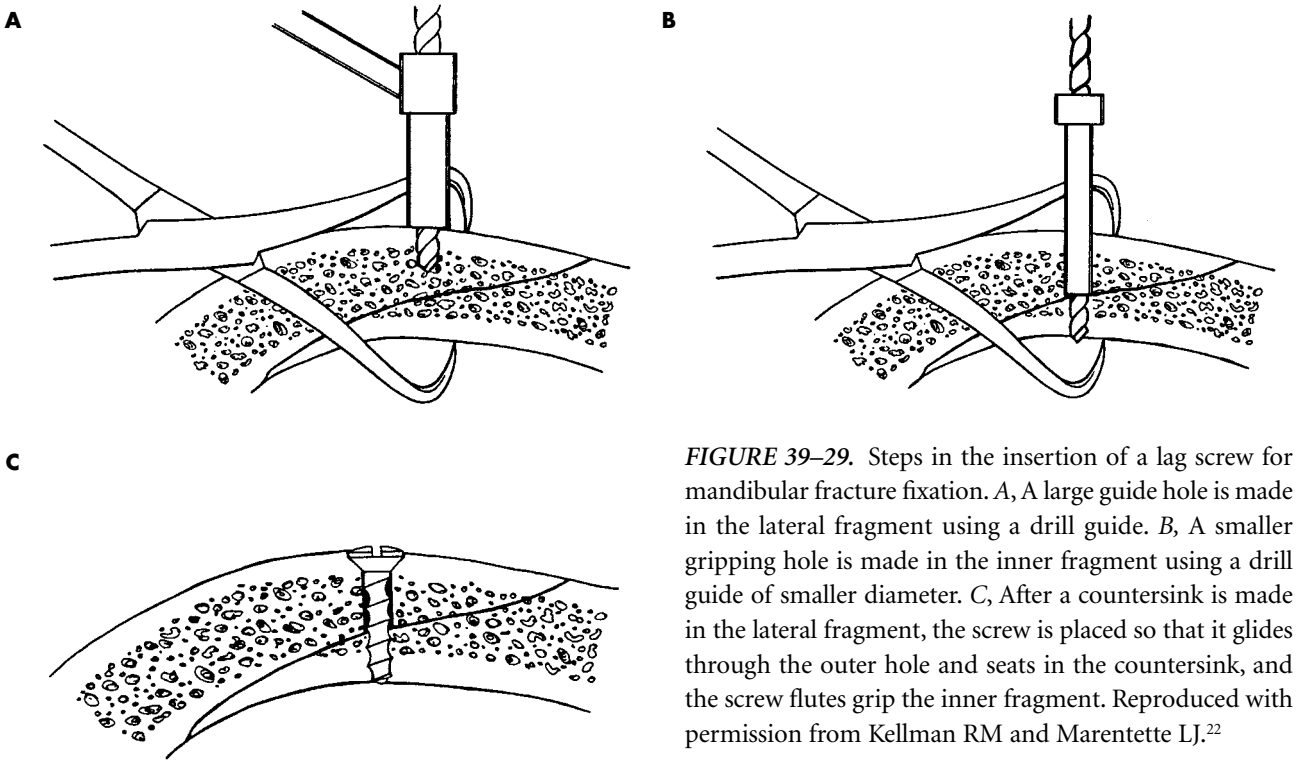
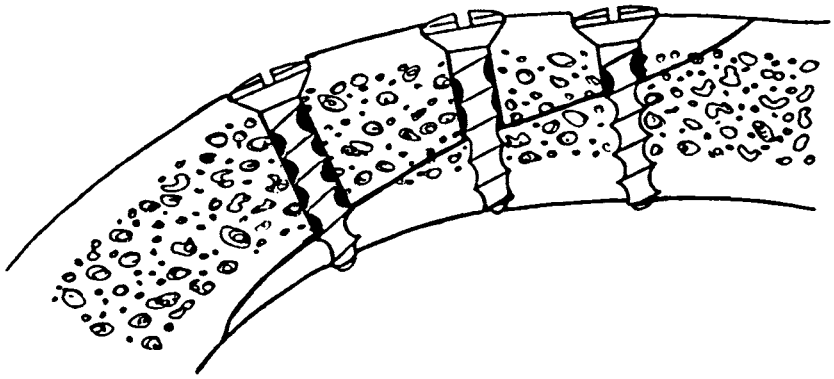


FIGURE 39–29. Steps in the insertion of a lag screw for mandibular fracture fixation. *A*, A large guide hole is made in the lateral fragment using a drill guide. *B,* A smaller gripping hole is made in the inner fragment using a drill guide of smaller diameter. *C*, After a countersink is made in the lateral fragment, the screw is placed so that it glides through the outer hole and seats in the countersink, and the screw flutes grip the inner fragment. Reproduced with permission from Kellman RM and Marentette LJ.[22]

that will produce the "glide hole" for the portion of the screw that will pass through the buccal cortex. Once the fracture line is encountered, then a second guide is inserted through the hole just made to make the smaller guide hole in the lingual fragment. This screw hole must begin precisely in the center of the glide hole and be carried through the lingual cortex. A countersink is done at the start of the glide hole to accommodate the head of the screw. The screw is slid through the glide hole, and the threads then engage the thread hole. The fracture is compressed as the screw head enters the countersink.

In those fractures in the parasymphyseal area, the exposure can be done intraorally. The elevation of the mucoperiosteal flap must be done with an eye to avoid injuring the mental nerves. The gliding hole and the threading hole are placed in the same fashion as those in the area of the angle. The countersink is done, and the lag screw is placed to ensure stability. Two or three screws are placed (Figure 39–30). Screws are now available for which the screw flutes are wider than the diameter of the glide portion of the screw. This then requires that only one screw hole is drilled, and the passage of the thread portion of the screw pulls the gliding part through, giving tighter grip to the latter while achieving the necessary compression.

FIGURE 39–30. Illustration of the importance of placing two or three lag screws to ensure rigid fixation. Reproduced with permission from Kellman RM and Marentette LJ.[22]



If rigid fixation is stable, the patient may be taken out of intermaxillary fixation and allowed to function. Of course, this decision is based on the stability of the rigid fixation and on the individual patient's compliance. If either the fixation stability or patient compliance is in question, the patient may be left in intermaxillary fixation for up to 6 weeks.

Dental Injury Injury to the dentition is a common accompaniment to fractures of the mandible and maxilla. A chipping of a tooth that does not extend to the dentin can be managed by the dentist by grinding the occlusal surface smooth. Such injuries are usually unaccompanied by pain. Larger fractures of the tooth substance expose the dentin or pulp and are painful. It is important to dress the tooth immediately with a dental compound that not only protects the tooth but also alleviates discomfort. A mixture of zinc oxide and eugenol done with a metal spatula or a glass plate forms a sticky compound with the consistency of putty that hardens when wet by the saliva. Eugenol helps to soothe the pain. Other proprietary compounds such as Dycal may be used. Later, the dentist can restore the tooth with a crown. In the case of fracturing of a slab of tooth face in a vertical direction, especially from the incisors, the tooth can be repaired by bonding.

Avulsion of the crown at the gum line is managed either by the extraction of the root or by application of a dressing until the tooth can later be restored with a crown. If the dentition is neglected and carious, then extraction is the best solution.

Dental fractures that occur through the root present a special problem. Root fractures seriously jeopardize the viability of the tooth. If the soft tissue of the pulp is sufficiently traumatized, the tooth will die. Once the pulp undergoes necrosis, the stage is set for the development of a root abscess. This may occur a short time after the injury or may be delayed for years. If such an abscess were to occur during intermaxillary fixation, an osteomyelitis and even a nonunion could occur.

The management of a tooth in the line of fracture is controversial. Some guidelines do exist that may be of help. If the tooth is carious and the root is fractured, then the tooth should be removed. If a carious tooth has a fracture through the periodontal ligament, but the root is intact, extraction should be seriously entertained. If the tooth is healthy, the vital consideration is the importance of the tooth in the stability of the fracture. This is especially cogent in the circumstance of the erupted third molar tooth that lies in an unfavorably aligned fracture of the mandibular angle. In such a fracture, the line extends obliquely from anterior to posterior as it proceeds from the occlusal surface to the inferior border of the mandible. If the molar is in the distal fragment, its occlusion with the maxillary third molar above it will preclude its displacement. If this molar is extracted, the pull of the medial pterygoid and masseter muscles will displace the distal fragment upward. The wisdom tooth, being the most poorly calcified in the mouth, is the most susceptible of all to caries. One should probably remove a carious or impacted third molar in such a circumstance and achieve fixation by an open technique.

Any dental trauma can cause pulp necrosis, even without a root fracture. It is important to follow up patients after facial fractures with vitality testing of the teeth. Once viability is lost, then a root canal procedure should be done to prevent abscess and tooth loss.

The final consideration is that of avulsed teeth. Such teeth should be immediately replaced in the avulsion socket and fixed by wires to the adjacent teeth for splinting. Many such teeth later become nonvital and require root canal procedures. Most are retained, however, and their preservation helps greatly in the maintenance of dental arch integrity.

Postoperative Care Drains are usually placed in open reduction wounds of the body, angle, and ramus. These drains are usually removed on the first postoperative day. Local wound care to the suture line is best done with peroxide cleaning and antibiotic ointment to avoid infection. Intraoral care should begin as soon as possible and consist of a water pick to the dentition but not the suture line and/or gargle of a solution of mouthwash and peroxide. The lingual aspect of the dentition is inaccessible in the patient with closed reduction in intermaxillary fixation. The effervescence of the peroxide will help to débride the areas of the gum, gingiva, and teeth of the lingual surface.[14] These patients often complain of pain, especially in the first 2 postoperative weeks because of the spasm of the pterygoid and masseter muscles while the patient is in intermaxillary fixation. Codeine and acetaminophen with codeine syrup are useful analgesics during this immediate postoperative phase.

Persistence pain beyond 2 weeks requires nonsurgical examination of the operative site as well as a radiographic study to determine whether there has been a shift in the fracture site.

The increased metabolic demands of healing and the difficulty of eating while in intermaxillary fixation make adequate nutrition a common problem. Liquid diets must be supplemented with a protein powder as well as whole milk. It is not uncommon for patients to sustain severe weight loss during the ensuing 6 weeks of intermaxillary fixation. Dietary counseling is often helpful. If the patient remains in intermaxillary fixation, weekly checkups to tighten the intermaxillary wires are absolutely necessary. If the wire ligatures are used for intermaxillary fixation, the wire stretches, loosening the fixation. If there is not close contact with the maxillary and mandibular teeth, a lingual version may develop. This results in a severe malunion if not corrected before fibro-osseous union occurs. Tightening of the wires every week prevents this from happening. Dental checkups, when possible, are also necessary in the case of impending dental abscess.

Intermaxillary fixation may be removed 5 to 6 weeks postoperatively, providing that there is no evidence of a nonunion. Nonunion can best be assessed by palpation at the fracture site. If there is pain in the mandible and not the teeth at the operative site, a nonunion or fibrous union should be strongly suspected, especially when accompanied by mobility. Another technique is the "torque test," which involves placing a standard wooden tongue blade on the side of the fracture and having the patient bite down and break the tongue blade.[15] If there is no bone pain with this procedure, a good union has been achieved, and the arch bars may be removed. The patient is still cautioned to remain on a soft diet for 2 to 3 additional weeks.

When arch bars are removed, a dental consultation is needed to inspect the teeth for occlusal discrepancy. Orthodontia may provide the necessary correction. Major malunions may require osteotomy of the involved segment and repositioning with further intermaxillary fixation. With an accurate initial assessment, careful surgical technique, and good postoperative care, the need for further adjustments of occlusion should be infrequent.

Special Considerations: The Edentulous Mandible The fractured edentulous mandible presents specific management problems for the surgeon. With the absence of teeth, the alveolar cortical bone undergoes total resorption. The superior portion of the mandible is greatly attenuated. The overall height of the mandible, therefore, is much less than that of a tooth-bearing mandible. Maintaining correct mandibular height, however, especially in the case of angle, ramus, and condyle fractures, is essential to avoid possible temporomandibular joint problems.

Ideally, in a completely edentulous patient with mandible fractures, the dentures can simply be wired into place. In the maxilla, screws used in plating techniques or Kirschner wires are excellent means of fixing the upper denture to the palate. The lower denture is affixed using circum-mandibular wires. The patient should be kept in some form of intermaxillary fixation for at least 6 weeks.

Another form of fixation, although less acceptable, is a cap splint. This can be especially helpful in a fracture involving the body or symphyseal area. The cap splint is formed by placing acrylic over the ridge of the mandible with the fracture and affixing it to the mandible, using circum-mandibular wires. When dentures are not available or the fracture is in an area that cannot be splinted but intermaxillary fixation is necessary, dental impressions can be made as a template for Gunning splints. Gunning splints can be affixed to the edentulous upper and lower jaw by means of the same techniques used to affix dentures.

Open reduction is necessary in edentulous mandibles with unfavorable "fracture lines" and in severely comminuted fractures. Interosseous wiring can be accomplished in the same fashion as with tooth-bearing mandibles. Intermaxillary fixation, however, must be used in this situation. Compression plating is also convenient in these cases by way of either the external or intraoral approach. The advantage of using a compression plate is that rigid fixation is achieved that does not require intermaxillary fixation. In elderly patients, however, especially those with osteoporosis of the mandible, placement of a secure plate may be difficult and may loosen with time.

Pediatric Mandibular Fractures A mandibular fracture in a patient younger than 15 years requires a knowledge of the state of both permanent and deciduous dentition. The tooth buds of the permanent teeth are located below the roots of the deciduous teeth. Surgical trauma to the tooth buds may retard or completely negate the buds' ability to erupt later. Pri-

mary teeth are fully erupted usually by the second or third year of life. The central incisors are the first teeth to appear, followed by the lateral incisors, first molars, cuspids, and second molars. The shedding of these teeth occurs in the same order and is usually complete by 12 years of age. Therefore, in the first 12 years of life, there are tooth buds of permanent teeth or partially erupted permanent teeth that may be endangered by surgery. The presence of these tooth buds also makes the tooth-bearing areas of the mandible in a child more prone to fracture than in an adult. They, in effect, create a more porous bony surface with less structural support.

In the tooth-bearing areas of a mandibular fracture in a pediatric patient, an overlapping cap splint provides adequate, stable reduction to promote anatomic bone healing. Because of the ability of children to heal rapidly, the splint is required only for approximately 4 weeks. In the angle area, if the fracture is in an unfavorable line, open reduction may be attempted. Radiographic documentation of the proximity of the erupting of the third molar bud and second molar tooth bud should always be confirmed. If the fracture line is close to a tooth bud, closed reduction may be necessary. Arch bars can be used in a pediatric patient with wire ligatures; however, they may cause premature shedding of the deciduous teeth. Generally, intermaxillary fixation is not necessary in a child for safe healing of a fracture of the symphysis or body. Condylar fractures in the pediatric population have been seen to remodel completely without surgical intervention; therefore, no treatment is recommended.

Complications in Pediatric Fractures The tooth buds of an infant or young child are especially susceptible to damage that may cause deformities of the erupted mature tooth later.[28] A long-term follow-up study of 28 children with mandibular fractures was done by McGuirt and Salisbury.[29] The mean length of follow-up was 10½ years. The children were divided into condylar only, noncondylar, and condylar fractures with other facial fractures. Forty-six percent of the patients had temporomandibular joint compromise. Five patients were concerned about their external facial asymmetry. A medical/dental team reviewed abnormalities in 47% of the patients radiographically. Sixteen percent of the overall group were thought to be both clinically and radiographically abnormal enough to require interventional therapy. Therefore, one can expect nearly one of five children with a mandible fracture to have some sort of difficulty that may require later intervention. It is recommended that these patients be followed closely postoperatively by a dentist and/or orthodontist, and the use of braces and elastics may be necessary to promote growth and development of the fracture area.[29] The remodeling of untreated subcondylar fractures is well documented.[30,31]

Severe Mandibular Fractures If the trauma is from an explosion or a gunshot wound, or if the blunt trauma is excessively severe, there can be multiple missing pieces of mandible.[32,33] In these cases, open reduction and exploration are mandatory if the patient's condition permits. The bony fragments that do not have periosteal attachments should be removed and the area débrided and cleaned aggressively. Often in these fractures, there are portions of mandible that are pedicled on small pieces of periosteum. If less than 25% of the bone surface area is attached to periosteum, most surgeons would agree that débridement is necessary to prevent this from becoming a sequestrum.

Gunshot wounds require soft tissue débridement, especially high-velocity wounds in which there may be a large temporary cavity. A discussion of ballistics is beyond the scope of this chapter. Let it suffice to say that high-velocity weapons cause extensive soft tissue damage from both the impact of the weapon and the secondary missile effect of the mobilized bone through soft tissues. The use of external pin fixation and the application of the biphase appliance of the mandibular segments alleviate the possibility of infection that would otherwise be incurred by placement of a foreign body such as a plate in a potentially grossly infected wound.

If the wound can be cleaned adequately and soft tissue coverage achieved, a long mandibular plate can be used to approximate the (usually) multiple segments. Once healing has taken place and the infection is cleared, in 6 months to 1 year this area can be reconstructed using autogenous bone from a variety of sources or a free flap (Figure 39–31).[32]

ZYGOMATIC FRACTURES

The zygoma is the third most frequently fractured facial bone, with 85% of fractures occurring in men.

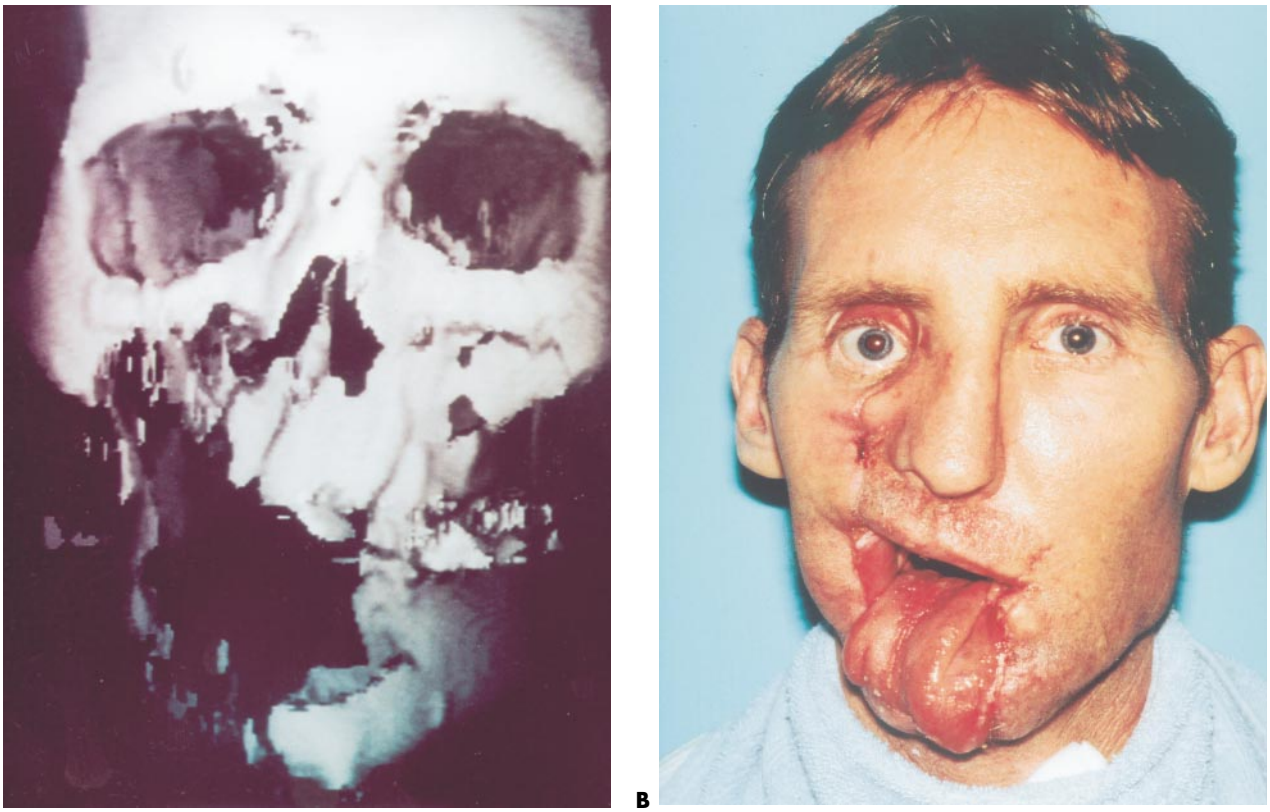


FIGURE 39–31. *A*, Three-dimensional computed tomogrpahic reconstruction of patient with a large defect secondary to a self-inflicted high-power rifle wound to the face. *B*, Preoperative patient has loss of mandible and soft tissue of the chin and neck. *C*, Reconstruction with fibular free flap.

Blunt trauma, owing to motor vehicle accidents or sports, makes up the great majority of zygomatic fractures. Fractures of this region are often called trimalar fractures, meaning fractures of the zygomatic frontal, zygomatic maxillary, and zygomatic temporal suture lines (Figure 39–32). Indeed, the fracture of the zygomatic sphenoid suture makes these fractures quadramalar.[6] The zygoma can withstand severe force, but as the most lateral extension of the facial skeleton, it is prone to injury in both the anteroposterior direction and the lateral and transverse directions. The attachments of the temporalis superiorly and the masseter muscle inferiorly tend to negate each other. However, the movement of the jaw with the masseter muscle does tend to distract the segments downward and medially.

The arch of the zygoma, which has contributions from both the zygomatic bone and the temporal bone, lies over the coronoid process. A depressed

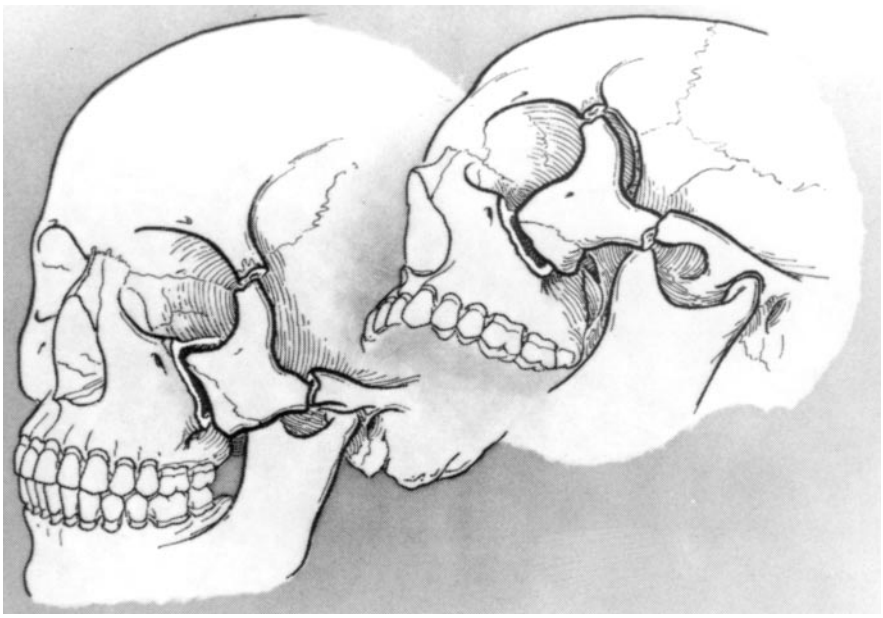


FIGURE 39–32. Classic tripod fracture of the zygoma. Reproduced with permission from Donald PJ. Zygomatic fractures. In: English GM, editor. Otolaryngology. Philadelphia: JB Lippincott; 1990.

fracture of the arch pushing into the temporal fossa often results in restriction in movement of the jaw because of impingement of the arch on the coronoid process. The fractures of the zygoma tend to occur at the articulations with the aforementioned bones. The body of the zygoma, which makes up the inferolateral orbital wall and is often the point of impact from trauma, is rarely fractured because it is the thickest part of the bone. Comminuted fractures of the zygomatic body are a difficult problem in surgical management.

Surgical Anatomy

The zygoma has four suture lines connecting it with the temporal, frontal, maxillary, and sphenoid bones. In addition to the temporalis and masseter muscles, the lesser and greater zygomatic muscles also insert on its surface. The orbital process of the zygoma makes up the anterolateral portion of the infraorbital foramen in the floor of the orbit. It is not uncommon, with a severe zygoma fracture, for there to be a herniation of orbital contents through the fracture of the maxillary zygomatic suture line. The infraorbital nerve exits from the infraorbital foramen at the articulation of the zygoma and maxilla. Damage to this nerve causes hypesthesia of the cheek on the affected side as well as the lateral aspect of the nose. From the body of the zygoma exit two sensory nerves, the zygomatic frontal and zygomatic temporal, which generally are clinically insignificant. When discussing a fracture of the zygoma, one must realize the importance of the intimate association to the lateral canthal tendon and the suspensory ligaments of the globe. There is some downward displacement in many fractures of the zygoma because of the previously mentioned traction of the masseter; therefore, the globe position can change. The entire globe can be pulled down because of downward displacement of the suspensory ligament of Lockwood, which attaches to Whitnall's tubercle, located on the lateral aspect of the orbital process of the zygoma. Also associated with the floor of the orbit are the inferior oblique and inferior rectus muscles, which are most often entrapped when the floor of the orbit is badly fractured (blowout fracture). Impalement of the muscles on a fracture fragment puts the muscle into spasm, which limits upward gaze. True entrapment of these muscles, which is rare, must be corrected surgically.

Diagnosis

History and physical examination are of primary importance in evaluating a zygomatic fracture. The patient gives a history of having received a blow from a fist, a ball, or other blunt object or has been involved in an automobile or motorcycle accident. He or she complains of localized pain and often has numbness over the ipsilateral cheek. If there is significant herniation of fat into the maxillary sinus,

diplopia ensues from herniation of orbital fat or entrapment of the inferior rectus and/or inferior oblique muscles. With no herniation of fat or evidence of an orbital floor dehiscence, inferior displacement of the zygomatic compound distracts Whitnall's tubercle and therefore the ligament of Lockwood, causing the same problem of diplopia.[6]

Trismus can result from impingement of the zygomatic arch on the coronoid process of the mandible. Often there is substantial masseter and temporalis spasm from contusion to that area, which may produce trismus even if the arch is intact.

The patient with a zygomatic fracture may also have epistaxis from bleeding into the maxillary sinus, the blood exiting by way of the ostium of the maxillary sinus and nose. On inspection, the patient has severe periorbital ecchymosis and swelling as well as a lateral subconjunctival hemorrhage secondary to tearing of vessels of the canthi. Blood in the anterior chamber (hyphema) indicates severe damage to the globe, and an emergency ophthalmologic consultation is mandatory.[34] Hemorrhage restricted to the subconjunctiva does not imply serious damage to the globe per se. Enophthalmos is common secondary to fat and muscle herniation into the maxillary sinus. Because concomitant intraocular injury is common, an ophthalmologic consultation should be obtained in all cases.

There is palpable facial skeletal asymmetry over the malar eminence and zygomatic arch if the swelling is not too severe. One may place a gloved finger into the oral cavity to feel the zygomatic buttress and appreciate crepitus and swelling. This movement is exquisitely tender to a patient with a fracture. There is trismus in some of these patients, defined as the inability to open the mouth more than 3 cm. In the absence of severe swelling, a fracture of the arch and the infraorbital rim can be palpated. The patient may be unable to raise the affected eye from the equator and experiences diplopia (Figure 39–33). A forced duction test should be done on all patients with diplopia. A positive test indicates entrapment of the inferior rectus or inferior oblique muscle. Table 39–3 provides a listing of signs and symptoms of zygomatic fractures.

Zygomatic fractures can be classified according to their severity. An isolated fracture of the zygomatic arch may occur with no orbital involvement. However, most fractures involve the four suture lines previously alluded to. Fracture of the zygomaticomaxillary suture line and orbital floor without involvement of the frontal zygomatic or frontal tem-

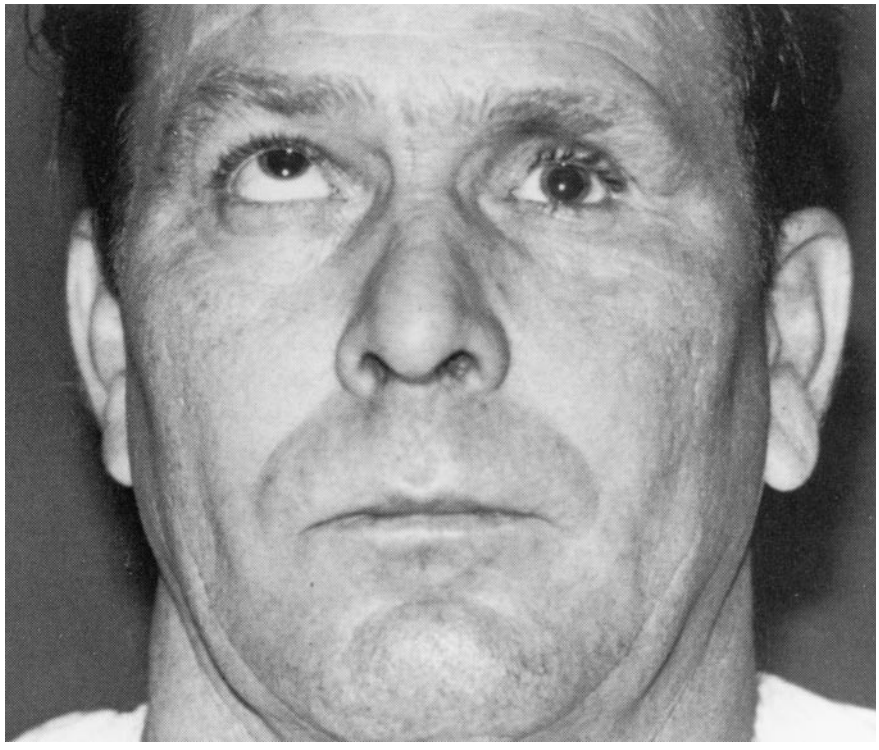


FIGURE 39–33. Patient with diplopia and restriction of upward gaze.

TABLE 39–3. Symptoms and Signs of Zygomatic Fractures

Symptoms	*Signs*
Pain	Infraorbital tenderness
Double vision	Lack of conjugate gaze
Numbness	Hypoesthesia: cheek, upper teeth
Epistaxis	Malar flattening
	Hypo-ophthalmus/enophthalmos
Trismus	Subconjunctival hemorrhage
Cosmetic deformity	

poral suture line constitutes an impure blowout fracture. A pure blowout fracture refers to a dehiscence of the orbital floor with an intact orbital rim.[35] A severely comminuted body fracture requires extensive exposure and is managed differently than a tripod fracture. The treatment of these various fractures differs, and diagnosis should be confirmed with radiographs.

Standard facial films including Waters', Caldwell's, Towne's, and submental vertex views are usually obtained on these patients. This battery of films helps rule out associated facial fractures. The Waters' and submental vertex views are of great help in diagnosis. The Waters' view can reveal an orbital floor dehiscence by the classic teardrop sign indicating herniation of orbital contents into the maxillary sinus. There is often blood in the maxillary sinus, however, which may obscure this sign. Also, the orbital rim, frontozygomatic suture line, and body of the zygoma are well visualized with the Waters' view. The submental vertex view is excellent for evaluating the zygomatic arch. Although most zygomatic fractures are readily diagnosed on plain radiographs, the CT scan affords a more accurate rendition of the fracture details. It is especially good at depicting the degree and severity of orbital floor fractures (Figure 39–34) and in detecting fractures of the orbital apex. Subtle unilateral associated fractures of the palate will be detected as well. Three-dimensional CT scanning provides a dramatic view of the fracture but adds little to the therapeutic decision-making process.

Treatment

The decision to repair a zygomatic fracture should be based on the goals one hopes to attain by such surgical intervention. The only strict indications for surgery are relief of trismus and correction of diplopia from an inferior displacement and entrapment of the inferior rectus and inferior oblique muscles. Most often, the patient desires repair for cosmetic reasons. If there is ipsilateral numbness after a fracture, this improves, in the large majority of cases, with no surgical intervention. Therefore, numbness of the cheek should not be the only indication for surgery. When obtaining the patient's consent for surgery, one must be careful to give counsel about the proximity of the globe to the surgical site and alert the patient to the real risk of retrobulbar hematoma and retinal tear from retraction of the globe.[34] These complications are extremely rare with careful surgical technique. More important in preoperative counseling, if an infraorbital rim approach is needed, is to inform the patient of the possibilities of prolonged lower lid edema and the possibility of ectropion. The onset of visual field changes,

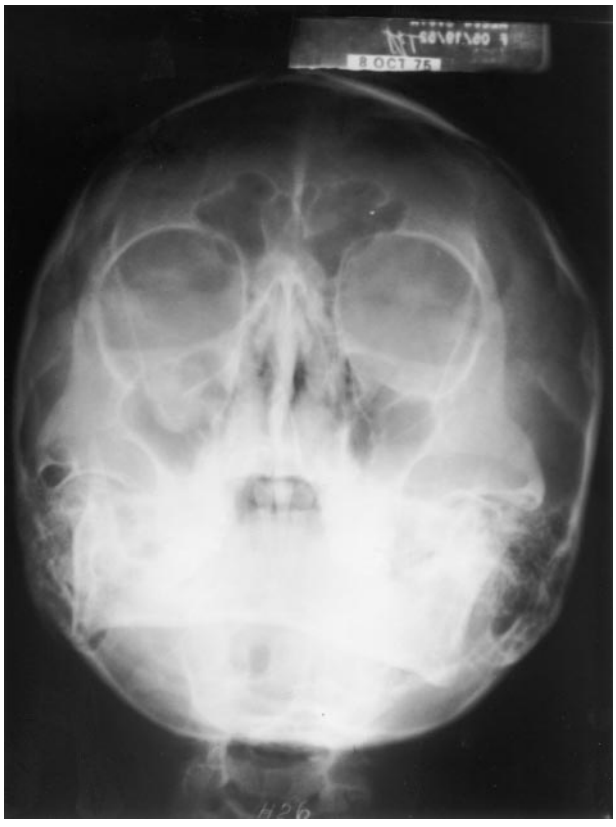


FIGURE 39–34. Waters' view demonstrating orbital floor fracture and prolapse of orbital contents into the maxillary sinus.

diplopia, or a change in the patient's visual acuity requires an ophthalmologic consultation.

One of the possible occult injuries of a zygomatico-orbital fracture is a retinal tear. Traction of the globe during surgery to repair a zygoma may extend a small, insignificant retinal tear, creating a large visual field defect. This can be avoided by limiting retraction of the globe to short periods of time and using extreme caution.

Serious intraocular injury mandates delay of repair until the globe has stabilized. Otherwise, the timing of the repair of a zygomatic fracture depends on the degree of soft tissue swelling. If swelling is severe, waiting for 5 to 10 days is acceptable, and moving the zygoma will not be a significant problem during that time. After 10 to 14 days, the zygoma may form a fibrous union, making mobilization of the fracture difficult. A good rule of thumb is to reduce the fracture as soon as the swelling has subsided enough to compare both zygomas intraoperatively.

Surgical Technique Multiple techniques have been used to repair a zygoma fracture. The four fracture lines imply that there should be points of fixation to stabilize the fracture, both laterally and inferiorly. The "classic" technique of placing a plate or wire in the frontal zygomatic fracture line and the zygomaticomaxillary suture line at the infraorbital rim provides adequate stabilization for a large majority of fractures. Some fractures treated in this manner are not adequately stabilized and require antral packing or an intraoral approach to reduce the inferior aspect of the fracture. One-point fixation techniques have been described using either a Kirschner wire[36] or, more recently, a rigid miniplate.[37–39]

Fractures of the zygomatic arch (Figure 39–35) with a significant cosmetic defect or trismus can be elevated by means of a Gillies incision in the temporal hairline (Figure 39–36, A) and reduced directly. The temporal branch of the facial nerve is avoided by placing the elevator deep to the superficial fascia of the temporalis muscle. The Boies elevator is placed under the arch, and the fragments are levered into position (Figure 39–36, B). A plate is rarely ever required for fixation because of the splinting action of the underlying temporalis muscle.

For the open technique, a supraorbital brow incision is made within the brow line or as an

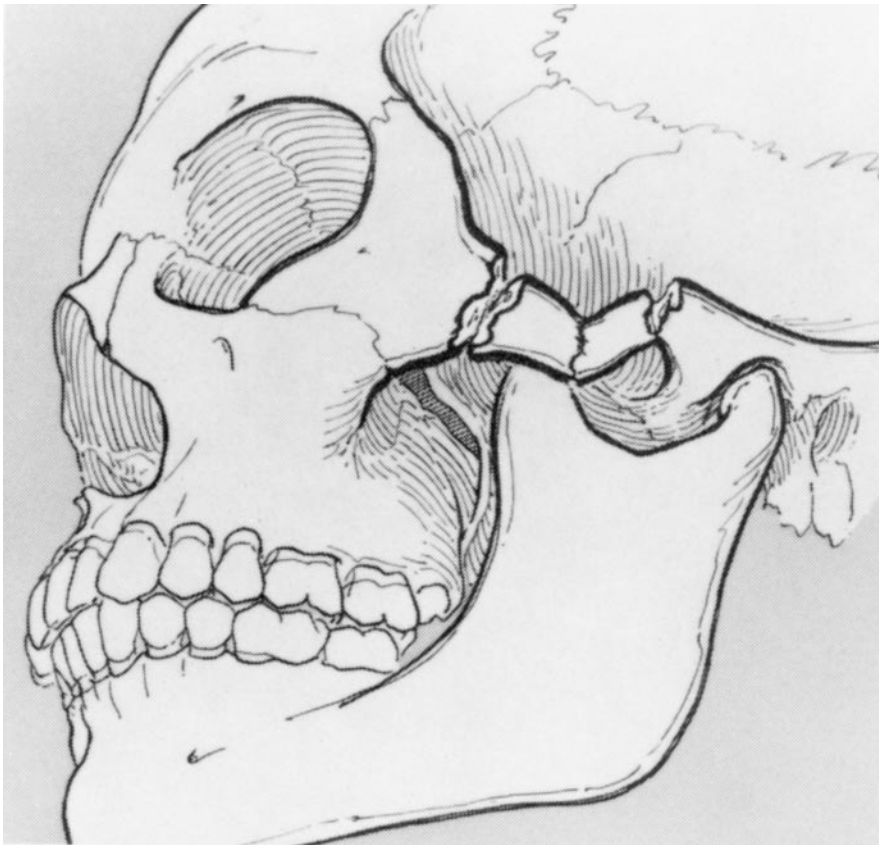


FIGURE 39–35. Depressed fracture of the zygomatic arch. Reproduced with permission from Donald PJ. Zygomatic fractures. In: English GM, editor. Otolaryngology. Philadelphia: JB Lippincott; in press.

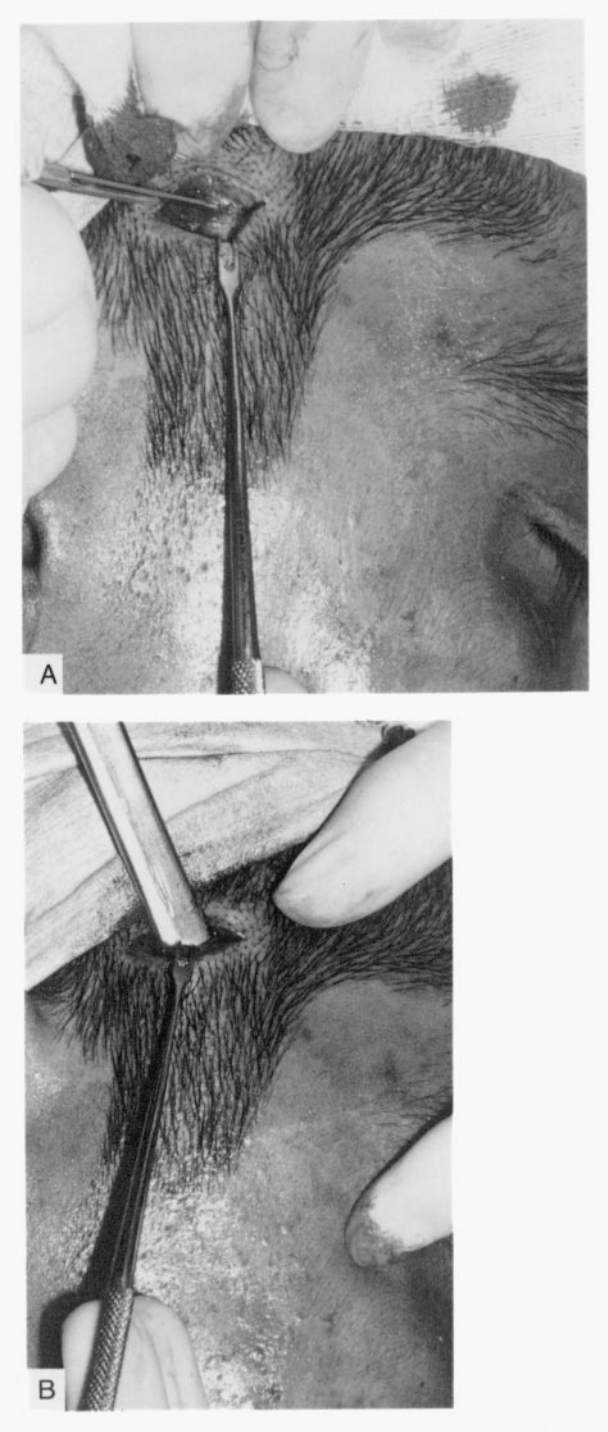


FIGURE 39–36. The Gillies surgical technique. *A*, Incision in the hairline for reduction of a zygomatic arch fracture. *B*, Boies elevator placed under the fractured arch to lift fractured segments into position. Reproduced with permission from Donald PJ. Zygomatic fractures. In: English GM, editor. Otolaryngology. Philadelphia: JB Lippincott; in press.

extended upper blepharoplasty incision (Figure 39–37), and the frontozygomatic suture line is exposed. The approach to the infraorbital incision varies depending on the surgeon's discretion. Choices include a transconjunctival approach (Figure 39–38) or an approach through the skin of the lower lid at various distances from the gray line of the lower lid. The placement of the subciliary incision should be along the crease and should be at least 3 to 4 mm below the ciliary line to avoid ectropion and prevent lid edema. The author advocates the step technique (Figure 39–39), in which a skin flap is elevated, leaving the orbicularis oculi muscle intact to prevent fat herniation from obstructing visualization of the fracture. Once the step incision of approximately 4 cm in length is made, it is carried down to the maxilla. The periosteum is entered below the infraorbital rim to avoid interruption of the septum orbitale that may cause later scarring and ectropion. The entire orbicularis oculi muscle and orbital septum are retracted superiorly.

These approaches provide good visualization of the rim and orbital floor while, at the same time, decreasing the chance of trauma to the infraorbital nerve. If there is a significant dehiscence of the orbital floor (greater than 1.5 cm) or entrapment of the inferior rectus or inferior oblique muscle, a graft of malar bone, nasal septum, Gelfilm, or irradiated cartilage should be used to repair the defect. The author recommends homografts to reconstruct the floor and feels that the danger of migration or extrusion with Silastic sheeting prohibits its use.[40–42] Titanium mesh has been added to the armamentarium of alloplasts to reconstruct the orbital floor. Initial results are encouraging, but the implant will need to stand the test of time.

A Boise elevator is then placed into the infratemporal fossa from the brow incision, beneath the arch, and the fracture is elevated to align the infraorbital rim and elevate the arch. Once the fracture has been put in proper position, a miniplate is bent to adapt to the reduced lateral orbital rim and is held in position to drill the screw holes. Holes to accommodate the screws for the miniplate that will be used to fix the frontozygomatic fracture are drilled on each side of the fracture line. Usually, two holes will be sufficient, but, on occasion, three holes are placed. Most plating systems now have self-taping screws so that tape is unnecessary. Some systems have self-drilling screws, however, and a small

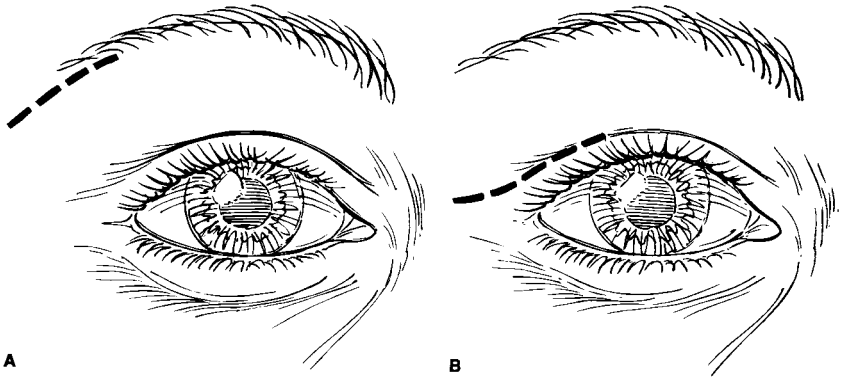


FIGURE 39–37. A, Schematic diagram of a lateral brow incision for approach to the frontozygomatic suture line. *B*, Schematic diagram of a right extended upper blepharoplasty incision for approach to the frontozygomatic suture line.

amount of pressure is initially required to get the screw started. This is not usually a problem in the lateral orbital wall but may be impossible in the infraorbital rim. A microplate is adapted to the infraorbital rim, and two drill holes are placed on either side of the fracture line (Figure 39–40). The screws are placed in the infraorbital and lateral orbital rims, providing the necessary two-point fixation. Some individuals add an additional plate at the inferior aspect of the zygomatic buttress. This will require an additional incision in the gingival buccal sulcus such as an extended Caldwell-Luc incision. A subperiosteal elevation will expose the underside of the zygomatic buttress. In most trimalar fractures, this step is entirely unnecessary.

Compound comminuted fractures of the body (Figure 39–41, A) require a direct approach by way of either a coronal scalp incision or extending the infraorbital incision superiorly and the upper brow incision laterally. Either a rigid miniplate or multiple interosseous wires can be used to fix multiple comminuted body fractures. When interosseous wiring is used for the compound body fracture, it is sometimes necessary to maintain elevation of the zygoma as it tends to prolapse. This can be done with an external pin fixation device such as a Morris biphase,

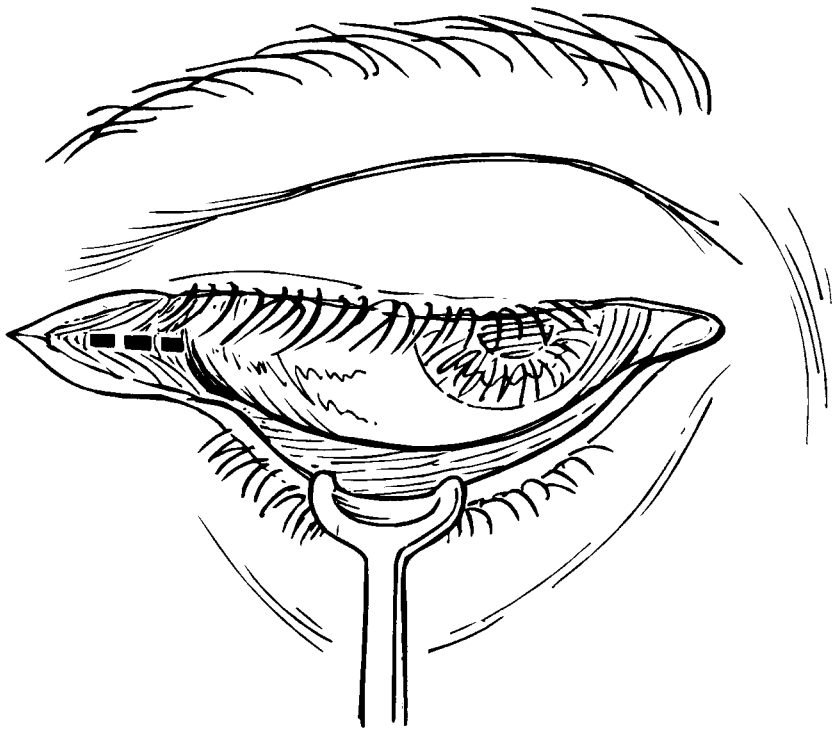


FIGURE 39–38. Schematic diagram of a transconjunctival approach to the infraorbital rim and orbital floor. It shows the transconjunctival incision and the lateral canthotomy already performed. The *dotted line* indicates the area for inferior cantholysis to improve exposure.

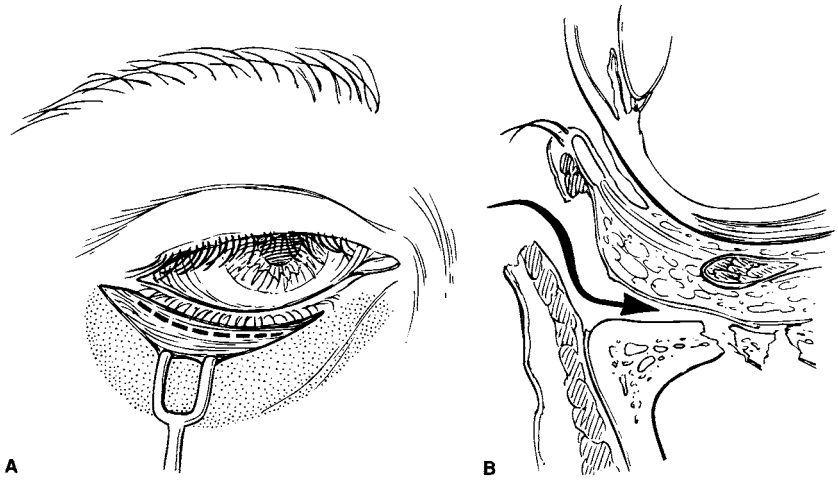


FIGURE 39–39. Stair-step incision through the infraciliary area, done to avoid the ectropion. *A*, Frontal view. *B*, Lateral view. Reproduced with permission from Foster CA and Sherman JE.[10]

stabilizing it superiorly to the frontal bone. Properly placed rigid miniplate fixation eliminates the need for an external device and is the method of choice (Figure 39–41, B).

Postoperative Care After surgery, if the fracture is unstable, a head dressing should be applied to hold a zygoma guard in place (Figure 39–42). This will not necessarily protect the zygomatic arch from collapsing with some pressure but will remind the patient not to roll on that side of the face while sleeping. This protective device is best retained for at least 2 weeks but can be removed when the patient is not sleeping. The zygoma guard is necessary in patients who have had interosseus wiring of a trimalar fracture, a reduced but not rigidly fixed arch fracture, or an unstable fracture with multiple comminution. Most fractures fixed with a rigid plating system will not need the guard.

Routine wound care is administered in the postoperative period, and the sutures should be

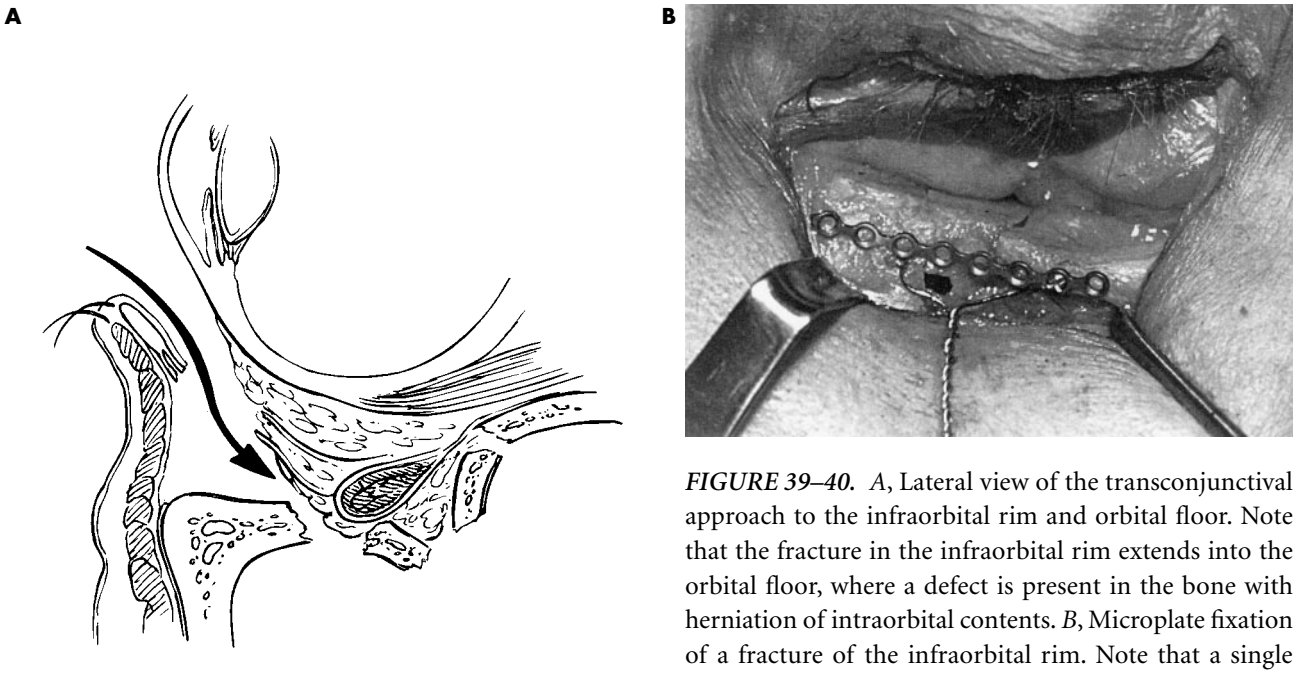


FIGURE 39–40. *A*, Lateral view of the transconjunctival approach to the infraorbital rim and orbital floor. Note that the fracture in the infraorbital rim extends into the orbital floor, where a defect is present in the bone with herniation of intraorbital contents. *B*, Microplate fixation of a fracture of the infraorbital rim. Note that a single interosseous wire has been placed to assist in reduction of the fracture while plating is performed.

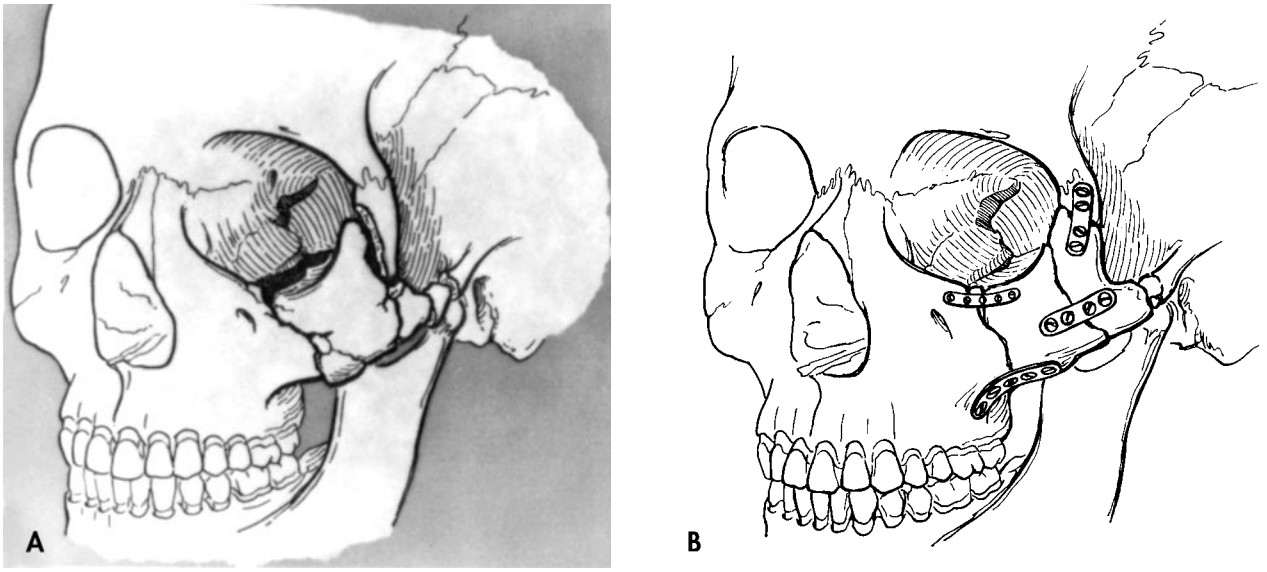


FIGURE 39–41. *A*, Complex body fracture of zygoma. *B*, At least three-point stabilization is required for adequate fixation. Reproduced with permission from Donald PJ. Zygomatic fractures. In: English GM, editor. Otolaryngology. Philadelphia: JB Lippincott; in press.

removed within 3 to 5 days. Diplopia from surgery may not resolve for up to 2 months. If there is postoperative diplopia and a thorough exploration of the orbital floor revealed no significant blowout fracture, the patient may be managed expectantly, with a good prognosis. Unrepaired zygoma fractures often lead to cosmetic deformity. Because of the pull of the masseter muscle, the zygoma is distracted in a downward and medial direction. This type of deformity causes lack of cheekbone prominence and an increase in orbital volume.

Enophthalmos is the result of failure to reduce an orbital floor fracture properly, with subsequent atrophy of herniated orbital fat. The actual orbital volume is increased if the zygoma heals in a rotated position. If orbital volume is increased appreciably, the remaining orbital contents are insufficient to maintain normal anterior protrusion of the globe, resulting in the "sunken eye" appearance of enophthalmos.[43] In rare instances, posteromedial displacement creates a smaller orbital compartment and a concomitant exophthalmos.[44]

In the nonseeing eye, the problems of enophthalmos can be corrected with minimal risk. Allografts or autogenous material may be placed in the orbit to increase orbital volume. Bony deformities are treated with onlay grafts or osteotomies of the zygoma with interposed calvarial bone grafts.

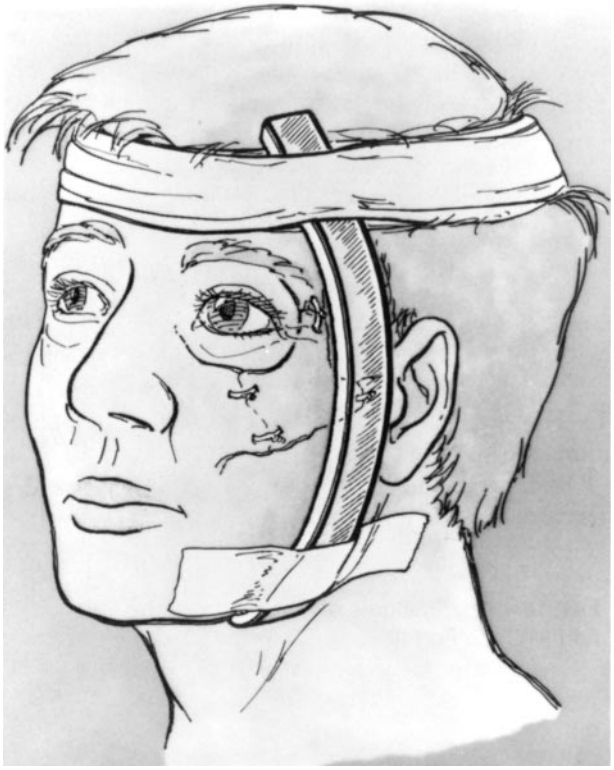


FIGURE 39–42. Guard placed to prevent displacement of the fracture reduction.

Enophthalmos in a person with vision can be safely corrected with calvarial bone grafts to the orbital floor. The use of other material such as titanium mesh or mesh covered in hydroxyapatite bone cement has also been advocated. Danger to the optic nerve and of extrusion or migration of alloplastic material[32] makes this procedure potentially hazardous.

An unreduced fracture of the zygomatic arch or one that is incompletely elevated into position may form a bony union with the coronoid process of the mandible. This is a very rare complication but will produce severe trismus that can be eliminated only by an open osteotomy and rigid fixation of the arch. Removal of the bony connection between the arch and the coronoid is essential, and placement of a Silastic sheet between the two bony structures may be necessary to prevent a relapse. The Silastic sheet may have to be removed at a future time, but if it is not problematic, it may remain indefinitely.

MAXILLARY (LE FORT) FRACTURES

Causes

Fractures of the midface, commonly referred to as Le Fort fractures, occur with much less frequency than fractures of the mandible, zygoma, or nose. Most often, they are the result of blunt trauma from accidents with automobiles, motorcycles, snowmobiles, or boats. The force required to fracture the maxilla and pterygoid plates of the sphenoid, which are the two fractured bones common to all Le Fort fractures, is considerable. Because of the alignment of the buttresses of the midface, which protect this area from vertical displacement, most of these fractures are caused by horizontal forces from the lateral, oblique, or anterior direction. The cranium and the orbit are intimately associated with these injuries and should be addressed with a high index of suspicion in all patients with midface injuries.

In the early 1900s, Rene Le Fort described the common lines of fractures associated with severe blunt trauma to cadaver heads.[45] Although his description is not characteristic of many compound midfacial fractures, it does serve to stage the degree of severity of the fracture and also alerts the physician to concomitant antecedent injuries that are peculiar to the three types of fractures. It is much more accurate to describe an injury in such terms as compound palatal zygomaticomaxillary or maxillonasal. Because of the number of bones in the face that can be associated with large compound midfacial injuries, however, and because many of them are not surgically accessible by any means, other than that associated with those injuries, the classic descriptions of Le Fort become useful indices for patient management and as tools for communication between physicians of different specialties.

Anatomy

The maxillae are the large paired bones of the midface that are approximately box shaped and almost fully pneumatized by their own sinuses. The lateral wall is part of the infratemporal fossa and articulates with the zygoma superiorly and the palatine bone posteriorly. The anterior wall has a lateral articulation with the nasal bones by way of the frontal process of the maxilla, which extends along the piriform aperture up toward the bony orbit to form the anterior crest of the lacrimal fossa. The medial wall is intimately associated superiorly with the ethmoid complex as well as the inferior turbinate. The posterior wall is relatively more substantial than the anterior wall "protecting" the pterygopalatine fossa and its nerves and vessels. Posteriorly, the maxilla also communicates with the sphenoid bone, both superiorly with the inferior aspect of the greater wing of the sphenoid and posterolaterally with the pterygoid plates. The vomer, ethmoid complex, palatine bones, zygoma, and nasal bones should also be considered part of the midface. One or all of these bones are often fractured in Le Fort fractures, although they may not all need direct treatment.

Le Fort's Classification Le Fort's classification refers to the lines of the three most common types of midface fractures (linae minores resistentiae). The Le Fort I fracture (Figure 39–43, A) is a lower palatal fracture that is also called Guérin's fracture. Le Fort II fractures (Figure 39–43, B) are also referred to as pyramidal fractures. Le Fort III fractures (Figure 39–43, C) are also known as craniofacial dysjunctions.

The Le Fort I fracture involves the floor of the nose, lower third of the maxilla, palate, and pterygoid plates, usually in one segment. Le Fort II fracture occurs across the nasal bony superstructure and the frontal process of the maxilla, down across

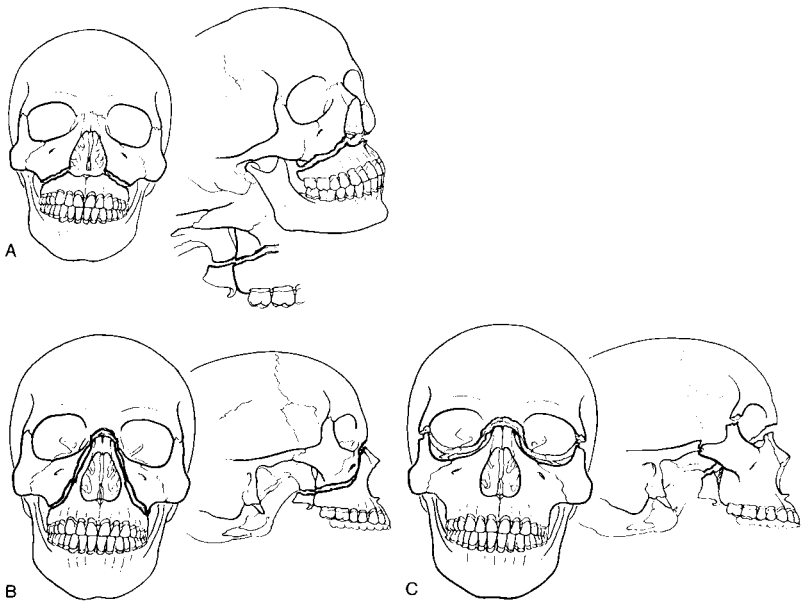


FIGURE 39–43. Le Fort fractures of the maxilla. *A*, Le Fort I. *B*, Le Fort II. *C*, Le Fort III.

the face of the anterior maxilla, and across the orbital floor, including the infraorbital foramen. The fracture line goes through the lateral wall of the maxilla extending to the pterygoid plates, usually higher than the Le Fort I. Le Fort III fractures are lateral fractures extending across the orbital floor into the lateral orbit through the zygoma. They can include the frontal zygomatic suture line. If the zygoma moves during palpation of the palate, a Le Fort III fracture is present.

Le Fort fractures are commonly found in combinations. A Le Fort I fracture may have an associated Le Fort II or III on the side of impact. A Le Fort II fracture may be a Le Fort III fracture on the side of the impact. The exact classification of the fractures on each side is important because of the differences of treatment associated with the different types. One should refer to the fractures based on the type of fracture on each side, for instance, left Le Fort III or right Le Fort I.

Further Considerations

Structural Pillars of the Midface Le Fort's classification was surprisingly complete in its description of the midfacial fractures, although there are certain limitations to his original descriptions. The classification of Le Fort fractures is based on the most superior fracture line involved in the fractures and does not take into account the fact that many of the fractures of the midface are comminuted and that within the confines of the midfacial skeleton, these multiple fractures may take on many configurations.[45–47] The concept of the structural pillars of the midface (Figure 39–44) helps to elucidate the key areas of fracture and therefore the subsequent stabilization of the midface as it relates to both potential cosmetic deformities and functional problems. An understanding of these structural pillars not only aids in assessment of the severity of the fracture but also provides an excellent strategic framework for surgical intervention.

There are three main buttresses of the midfacial cranial complex.[47] Going from anterior to posterior, the anteriormost buttress is the nasal maxillary buttress extending from the maxillary alveolus along the piriform aperture up to the medial side of the orbit through the anterior lacrimal crest, the frontal process of the maxilla to the superior orbital rim, and ending at the frontal bone of the cranium.

The zygomatic buttress extends from the maxillary alveolus above the anterior molar teeth to the zygomatic process of the frontal bone with a vertical component consisting of the zygomatic arch articulating with the zygomatic process of the temporal bone at the base of the skull.

The buttress of the pterygomaxillary alveolus articulates with the base of the skull through the orbital process of the palatine bone, articulating pos-

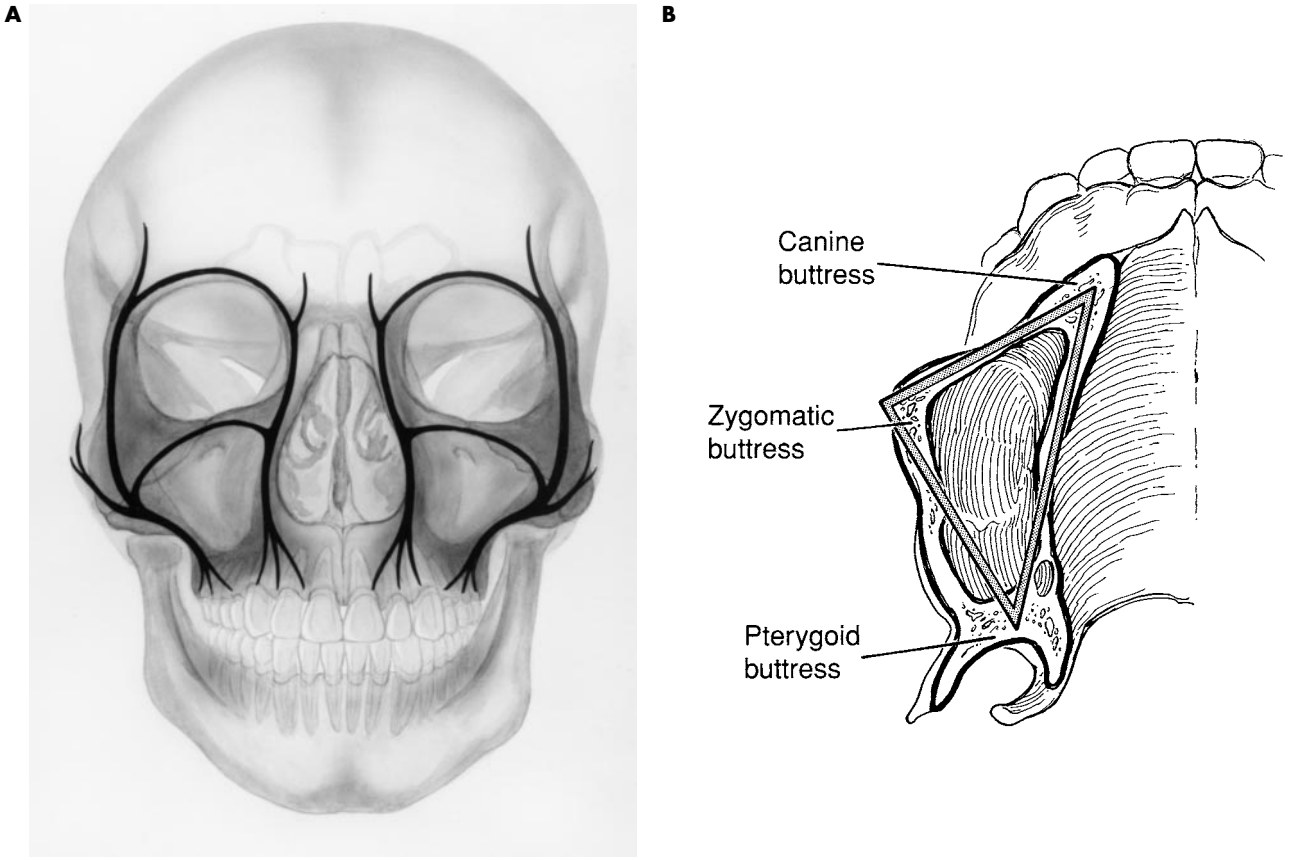


FIGURE 39–44. *A* and *B*, Biomechanical stress curves of the middle third of the face.

teriorly with the sphenoid bone. The pterygoid portion of this vertical buttress consists of the pterygoid plates extending inferiorly and anteriorly toward the maxilla. The horizontal system of buttresses divides the oral, nasal, and orbital regions of the face.

If one considered these pillars in midfacial fractures, therapy should be directed toward the stabilization of as many of these buttresses as possible. Probably the two most important areas of stabilization are the nasal maxillary buttress and the zygomatic buttress. The more posterior vertical pterygomaxillary buttress is more inaccessible surgically, although the most inferior portion of it can be stabilized if necessary. In severely comminuted Le Fort fractures, the surgeon must appreciate the advantage of open reduction and internal fixation of multiple fragments in their anatomic positions, especially those relating to this buttress complex. Without proper reduction of the midface fractures or proper stabilization of concomitant subcondylar mandibular fractures, there is always the possibility of midfacial telescoping, which results in a shortened midface.[46–48]

Diagnosis

Eliciting an accurate history on a patient with severe midfacial injury is often difficult or impossible because of associated intracranial, abdominal, or intrathoracic injuries. The motor vehicle accident victim may also be inebriated. If the patient is conscious, however, and does not have any life-threatening injuries requiring emergency surgical intervention, a thorough history is helpful in ascertaining the extent and severity of facial injury.

Because the maxilla forms the anterior floor of the orbit, there may be herniation of the orbital soft tissues into the maxillary sinus, similar to that seen in a blowout fracture. Diplopia may be a complaint in Le Fort II or III fractures. Complaints of blurred vision or a change in visual acuity mandate

an ophthalmologic evaluation. Epiphora may also be a component of compound fractures involving the lacrimal duct and/or the inferior medial orbital area. Difficulty in breathing is a common problem from congestion associated with edema from the fractures as well as physical derangement of the nasal bones and septal structures. Bleeding from the maxillary or ethmoid ostia or from lacerations within the nasal cavity may cause severe nasal obstruction. Patient complaints of a salty taste in the mouth should alert the physician of the possibility of a cerebrospinal fluid leak. Most commonly, these are found in the cribriform plate or the roof of the ethmoid sinuses.[5,6] Displacement of the maxilla causes the patient to complain of malocclusion. Although there may be either buccal or lingual version, the most common occlusal abnormality is an open-bite deformity. The open-bite deformity is caused by the powerful forces of the pterygoid muscles distracting the posterior part of the maxilla inferiorly.[6] A blow directed anteriorly causes displacement of the maxilla posteriorly, resulting in a so-called "dish face" deformity and a Class III malocclusion. A more detailed discussion of dental occlusion is given earlier in this chapter.

Airway compromise is often the presenting symptom. The posteroinferior displacement of the fractured maxilla by the pterygoid muscles tends to decrease the airway. This, coupled with attendant swelling and hemorrhage, may produce airway obstruction.

On physical examination, the patient may have periorbital ecchymosis, massive tissue swelling, or subconjunctival hemorrhage if the infraorbital rim is involved. Dental examination confirms the previously mentioned findings relating to open bite, lingual/buccal version, and Class III malocclusion. Bony crepitus of the midface, especially in severely comminuted injuries, is common. Also, extensive fractures in the presence of a large communication with the antrum and the nose cause subcutaneous emphysema.

Nasal or pharyngeal hemorrhage may be massive. This is important because most patients are evaluated in the supine position. Nasal and pharyngeal bleeding drains toward the posterior part of the pharynx and may go unnoticed. Palpation reveals a mobile palate. If the palate does not move even though there appear to be gross occlusal differences and/or obvious fractures, it is because the midface has impacted into an immobile position.[49] To move the palate in these cases may require considerable force.

Imaging

Radiologic diagnosis of Le Fort fractures is a very important adjunct to their treatment. Many midfacial fractures can be assessed with sinus radiographs, but the precision of the CT scan is a huge advantage in delineating the extent and severity of midfacial fractures. Central injuries can be ruled out by extending the scan through the head; any nasal-frontal or frontal sinus components can be delineated, and fractures of the sphenoid and orbital apex are excellently portrayed. Scans in both the axial and coronal planes are necessary for the most complete analysis. Magnetic resonance imaging is of little added assistance except in the case of cerebral trauma or compromise of the optic nerves. Three-dimensional reconstruction, although spectacular, offers little additional information that will guide therapy.

Treatment

As in all patients involved in trauma, the fundamental precepts of trauma care apply. The airway is of great importance in these patients. A displaced Le Fort fracture can compromise the airway, especially if it is associated with concomitant massive swelling of the tongue and oropharynx.[5,6] A cricothyroidotomy may be required urgently. Endotracheal intubation should be avoided because of the problems of poor visualization, the possibility of aggravating a cervical spine injury, and the possibility of causing injury to the central nervous system from the endotracheal tube. Nasotracheal intubation is even worse because of the danger of intracranial intubation through a fracture line in the skull base. Tracheostomy preferably should not be done in the emergency room but performed electively in the operating room. All other injuries should be ruled out, including intra-abdominal, intrathoracic, and intracranial injuries. The cervical spine should undergo radiographic evaluation.

If the patient is to undergo immediate general anesthesia for a related injury and is hemodynamically stable, immediate repair can be undertaken if good radiographic documentation of the injury has been obtained. Often, if a patient is rushed to the operating room because of a life-threatening injury,

proper radiographs may not have been obtained, and if a Le Fort fracture is suspected, the patient should be placed temporarily into intermaxillary fixation and undergo tracheostomy. In this emergency situation, the patient can be more fully evaluated radiographically at a later time to ascertain the extent of the midfacial fractures. Definitive repair of all fractures may be delayed for up to 10 days before bony union may make reduction difficult. Thorough preoperative planning to repair these fractures includes ophthalmologic examination, dental evaluation, and clearance by neurosurgeons if there are concomitant intracranial injuries. Repair should also be delayed until the status of the cervical spine is determined.

Surgical Approaches Incisions used to approach midfacial fractures vary according to the specific location of the fracture. Isolated Le Fort I fractures are usually approached by an extended sublabial incision (Figures 39–45 and 39–46). This incision allows exposure of the zygomaticomaxillary buttresses and piriform apertures bilaterally. If the fracture involves the infraorbital rim (Le Fort II), a transconjunctival-lateral canthotomy or subciliary incision is usually used to expose these fracture sites. If a Le Fort II fracture is more extensive and greater exposure to the nasoethmoid complex is required, an external Lynch incision or extended coronal incision may be used. The frontozygomatic suture line may be approached by a coronal, brow, or extended upper blepharoplasty (supratarsal) incision. The coronal incision has the advantage of providing exposure to the zygomatic arch and the nasoethmoid region. This obviates the need for a Lynch incision for nasoethmoid exposure.

The incisions used depend on the location and extent of the fractures. Most Le Fort fractures do not manifest as the original description but are combinations of complex fractures of the midface and require knowledge of multiple surgical approaches.

In patients with dentition, arch bars and intermaxillary fixation are initially applied to re-establish pretrauma occlusion. In edentulous patients, splint or denture containing an arch bar is fixed to the mandible or maxilla to re-establish appropriate skeletal relationships. This is performed with circum-mandibular wires or drop wires from the piriform rim or zygoma. The maxillary (palatal) splint may also be fixed to the palate with two transpalatal screws. In any case, splints or dentures are placed to re-establish the occlusal and skeletal relationship.

In patients in whom midfacial fractures are displaced, disimpaction of fractures may be required before placement of intermaxillary fixation. Disimpaction is best achieved by use of the Rowe-Killey disimpaction forceps (Figure 39–47). The straighter blade is placed in the nasal cavity along the nasal floor and the more curved blade over the alveolar ridge and along the palate. With the handles firmly grasped, a downward and anterior pull will disimpact the maxilla and restore it into its normal relationship with the mandible and skull base. After the occlusal relationship is re-established, all fracture

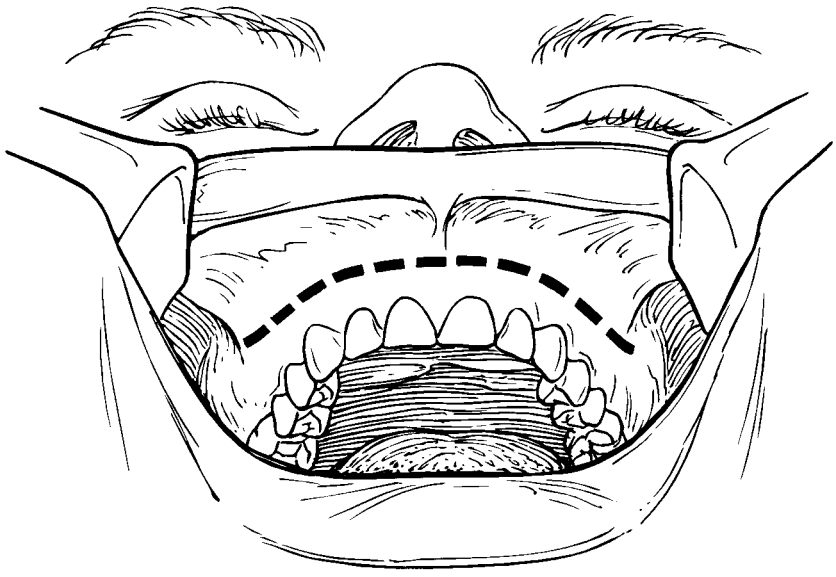


FIGURE 39–45. Schematic diagram of a sublabial incision for approach to midfacial fractures.

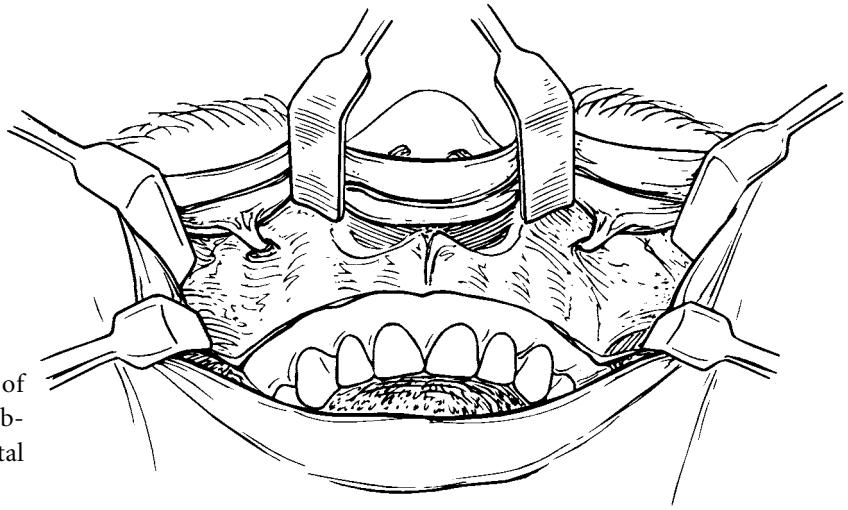


FIGURE 39–46. Schematic diagram of subperiosteal dissection through a sublabial incision exposing the infraorbital nerves.

sites are fully exposed. The facial skeleton should then be reconstituted in three dimensions: height, width, and depth.[46,47] Careful attention should be paid to re-establishing all important facial buttresses. In particular, the vertical zygomaticomaxillary and nasomaxillary buttresses should be carefully reduced and fixated to re-establish vertical facial height. All rigid fixation should be applied with sufficient stability to counteract any forces that could disrupt bone repair during healing.

For a Le Fort I fracture, a two-point stabilization at the nasomaxillary and zygomaticofacial buttresses is established on each side. Titanium low-profile miniplates in an "L," "X," or square configuration are placed on the anterior buttresses and usually an "L"-shaped plate on the undercurve of each zygoma onto the maxilla bilaterally (Figure 39–48). More recently, low-profile absorbable plates of polymers composed of polyglycolic or polylactic acid or a combination have been employed in place of titanium plates. These plates are reputed to retain sufficient strength to maintain fixation over the critical period of healing and are then absorbed. This is especially an advantage in children in whom the

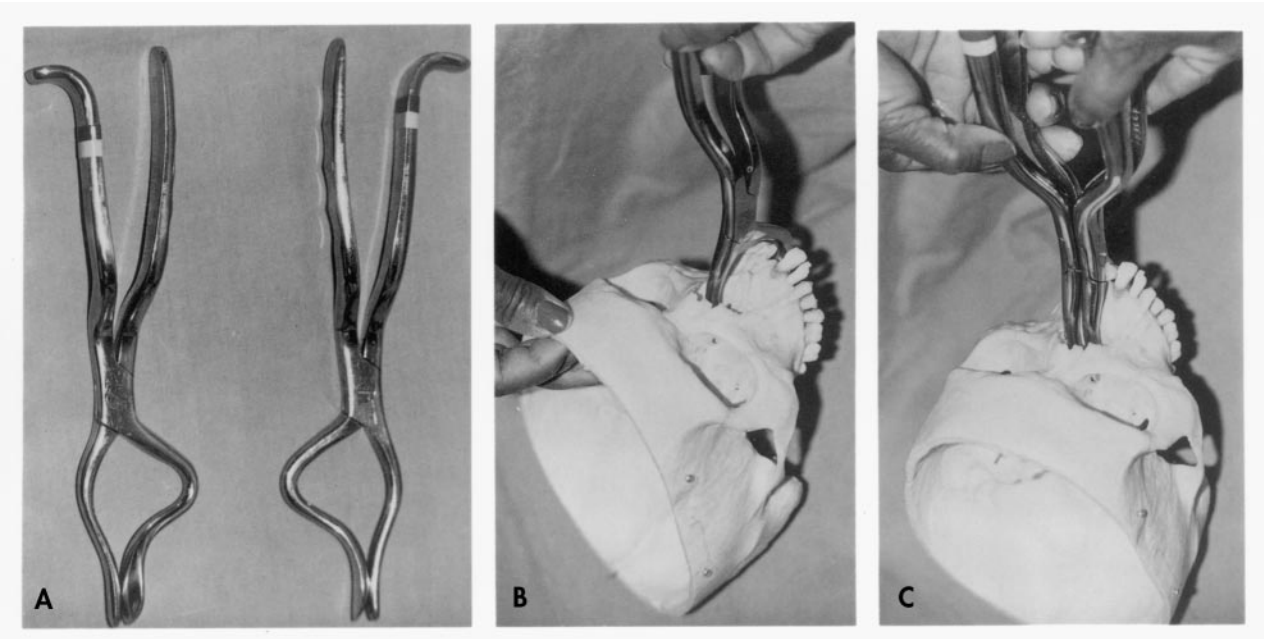


FIGURE 39–47. Maxillary disimpaction using Rowe-Killey forceps. *A*, Pair forceps. *B*, Forceps applied on one side. *C*, Forceps applied on both sides.

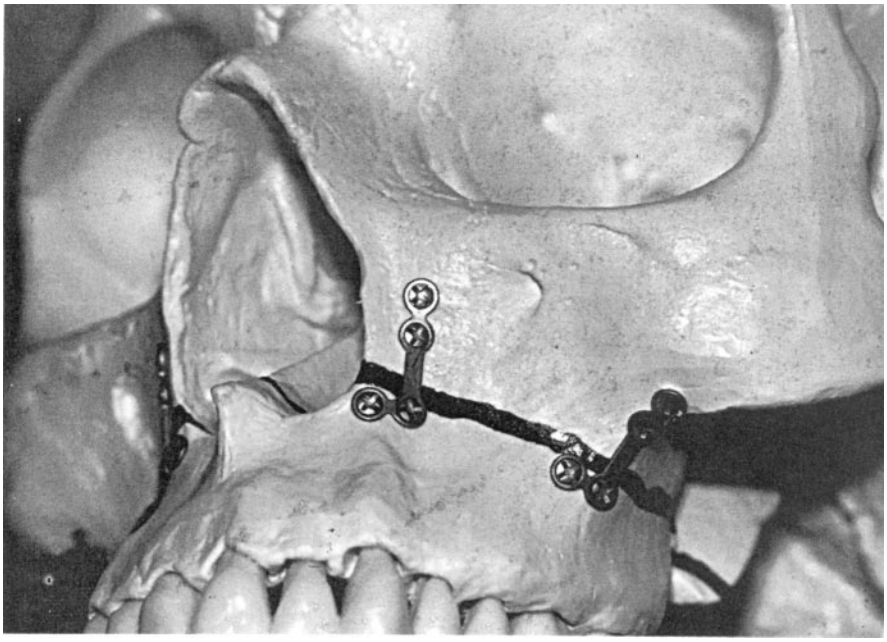


FIGURE 39–48. Miniplate fixation of a Le Fort I fracture. Note the stabilization of the nasomaxillary and zygomaticomaxillary buttresses.

metal plates may impede bony facial growth at suture lines. In adults, the principal advantages are in preventing pain that is occasionally seen with the metal plates and with time eliminating the palpable plate at the lateral and infraorbital rims. The durability and utility of absorbable plating systems have yet to stand the test of time.

Le Fort II fractures require fixation at the infraorbital rim and the zygomaticomaxillary buttresses. If the nasal bones are comminuted, microplates are used on the nasal bones. The infraorbital rims will be fixed with microplates as well. As in most instances, the miniplates have too high a profile and will be obvious through the skin.

The Le Fort III fracture is often a devastating injury usually accompanied by cerebral trauma. It is not uncommon to have a staged type of treatment in which the neurosurgeon controls the intracranial injury and the management of the facial component is limited to the performance of a tracheostomy and the application of arch bars and interdental elastic bands. The treatment of the maxillary fracture is delayed until the patient's neurologic status has stabilized. The shattered maxilla must be fixed between two stable platforms. The cranium provides the superior stabilization point and the mandible the inferior. All displaced fractures of the cranial vault must be restored to their normal anatomic position. Similarly, all fractures of the mandible must be rigidly fixed in correct position to establish the occlusal template for the maxillary dentition to approximate accurately. Rigid fixation of all mandibular fractures is vital to ensure this relationship. Displaced subcondylar fractures must be fixed to provide the stable platform even when only unilateral. Undisplaced subcondylar fractures are usually sufficiently stable to avoid rigid fixation. The goals of reduction and fixation of the Le Fort III fracture must be clearly borne in mind as fixation proceeds. First, the supporting buttresses of the face must be re-established. Second, the functional elements must be restored, such as the correct orbital volume, including an adequately restored orbital floor with orbital contents free of entrapment, a patent nasal airway bilaterally, and maxillary sinuses that will adequately drain. Third, all-important esthetic landmarks must be restored, such as the orbital rims, nasal dorsum, and malar eminences. It is not necessary and, in some instances, is indeed impossible to approximate every small maxillary fragment. The lateral orbital rims and the buttresses are fixed with miniplates and the nasal dorsum and infraorbital rims with microplates (Figure 39–49). Occasionally, small fragments in key positions may require fixation with fine wire or even suture material. The use of titanium wire will prevent interference in subsequent scans.

If rigid fixation is considered stable and if the patient is compliant, intermaxillary fixation may be

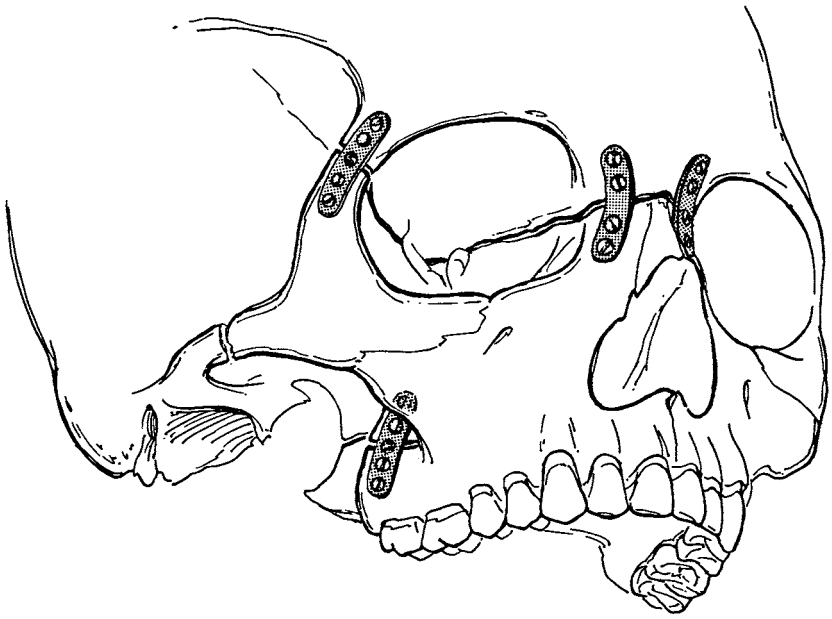


FIGURE 39–49. Three-point rigid fixation of a Le Fort III fracture.

removed at the conclusion of the operation or within the first 1 to 2 weeks after the operation. If the stability is in question (ie, with bone comminution or bone loss), intermaxillary fixation should be left in place for up to 6 to 8 weeks. This will help maintain the occlusion while bone healing occurs.

NASOFRONTAL-ETHMOIDAL FRACTURES AND TRAUMATIC TELECANTHUS

A central blow to the face created by a small-diameter object, such as the head of a hammer, may not produce a Le Fort type maxillary fracture but, if of sufficient force, may cause a severely comminuted nasal fracture that shatters the medial orbital walls and telescopes into the ethmoid block, lodging the fractured bones under the nasal process of the frontal bone (Figure 39–50). Not uncommonly, this nasofrontal-ethmoidal fracture will be accompanied by telecanthus that results from either the shearing away of the medial canthal tendons or a displacement of the lacrimal and maxillary bones into which the tendon or tendons insert. The nasal dorsum is essentially displaced into the fractured anterior ethmoid cells (Figure 39–51). This injury produces a typical clinical picture of a patient with a "pig snout" deformity consisting of a flattened nasal dorsum with an exaggerated excessively superiorly rotated tip (Figure 39–52).

The telecanthus may be unilateral or bilateral. The natural contraction of the orbicularis oculi, whose tendonous extension comprises the bulk of the medial canthal tendon, pulls the medial canthus away from the midline. This produces a blunting of

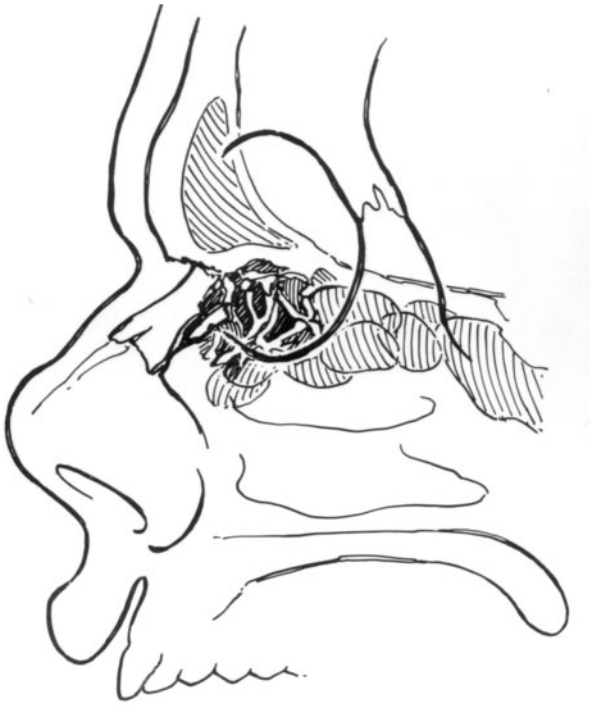


FIGURE 39–50. Nasofrontal complex fracture. Note nasal bones driven under the frontal bones. Reproduced with permission from Donald PJ et al.[51]

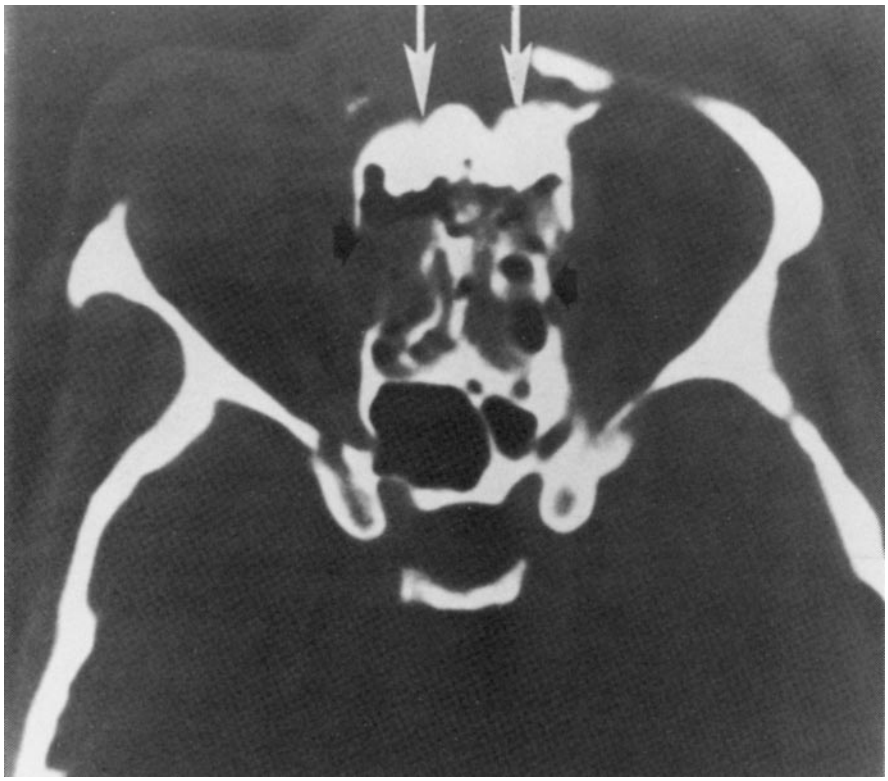


FIGURE 39–51. Computed tomographic scan showing a nasofrontoethmoidal fracture. *Arrows* show direction of displacement. Reproduced with permission from Donald PJ et al.[51]

the medial canthus, with a widening of the medial canthal angle, a shortening of the width of the palpebral fissure, a slight ectropion, and an inferior inclination of the medial canthus (Figure 39–53). The width of each palpebral fissure should normally be the width between both eyes, approximately 25.5 to 37.5 mm in adult women and 26.5 to 37.8 mm in men.[50] With this injury, there is usually the absence of the "bowstring" sign. If tension is created in the medial canthal tendon of the normal eye by pulling laterally on the lateral canthus of the eye, a taut bowstring-like sensation is felt by the palpating finger. If, instead, there is a soft, loose, boggy feeling at the inner canthus, then the presence of a ruptured tendon is highly likely. In addition, there may be injury to the lacrimal system with fracturing of the bony nasolacrimal duct. This will produce epiphora. Epiphora may also result from the loss of the effectiveness of the "lacrimal pump" mechanism. This is formed by the splitting of the medial canthal tendon into two slips of tendon, which insert onto the lacrimal bone both anteriorly and posteriorly to the lacrimal sac (Figure 39–54). With slackening of the tendon by displacement of the medial canthal tendon, the pump is unable to function effectively.

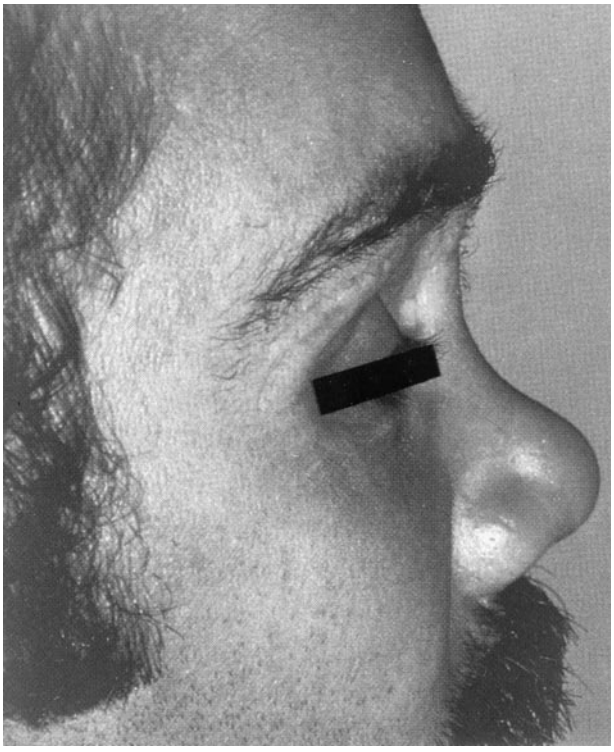


FIGURE 39–52. "Pig snout" deformity characteristic of a nasofrontoethmoidal complex fracture. Reproduced with permission from Donald PJ et al.[51]

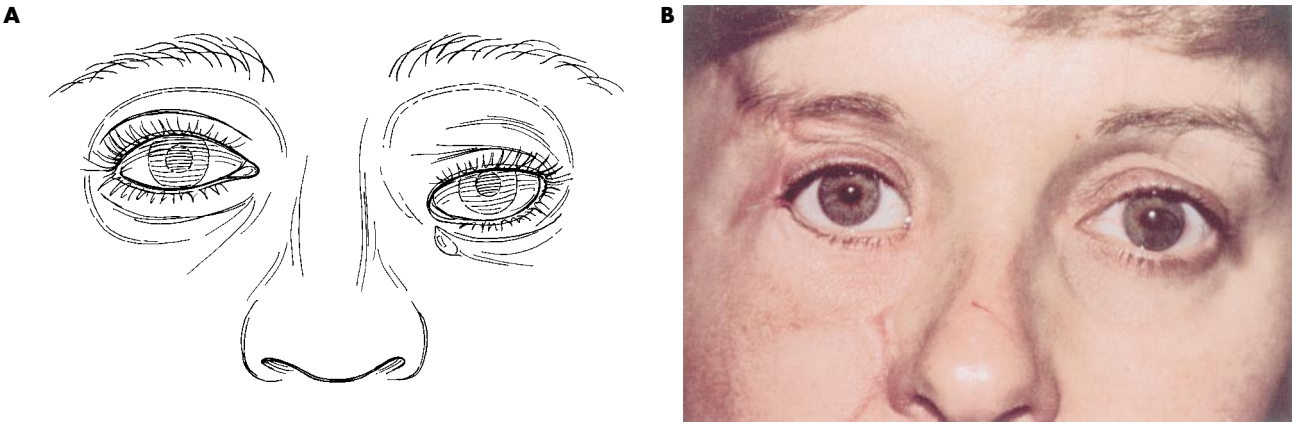
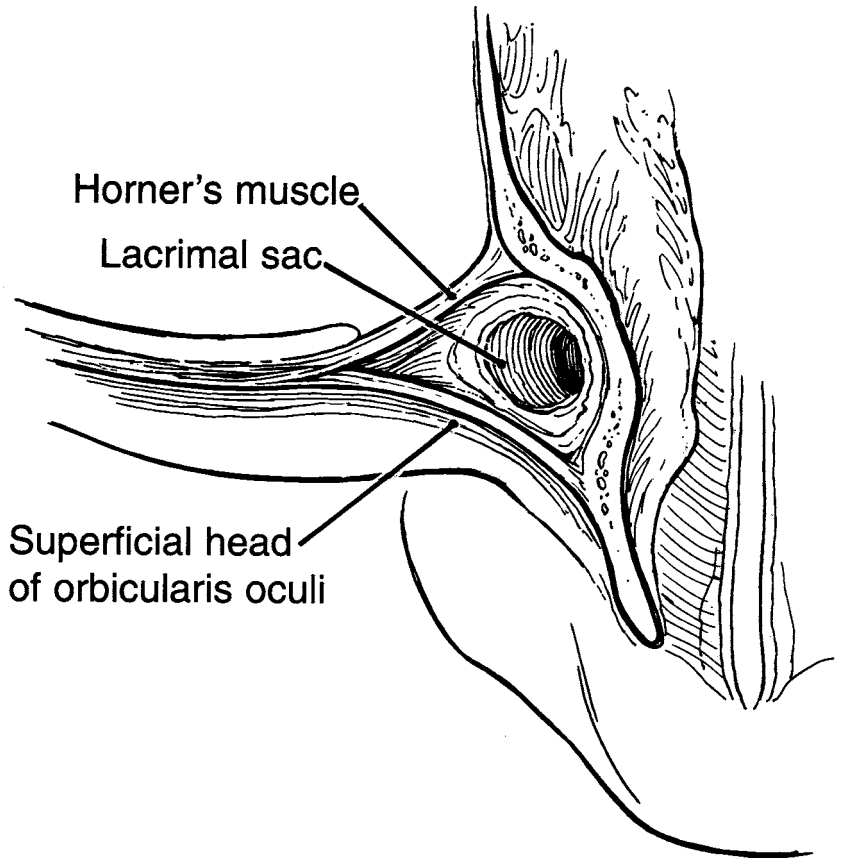


FIGURE 39–53. *A*, Diagram showing telecanthus. *B*, Patient with unilateral telecanthus.

The repair of the nasoethmoidal fracture consists of disimpacting the nose from the collapsed ethmoid sinuses and, if there is telecanthus, repairing one or both canthal tendons. It is often a difficult task and should be undertaken as soon as possible after the accident. One method, using bilateral Lynch incisions, of disimpacting the fragments in the ethmoid is to dissect the periorbita away from the fractured lamina papyracea on both sides, insert two bone hooks into the ethmoid block, and turn the hooks nasally and then apply a down and forward traction (Figure 39–55). This will usually pull the fractured nasal bones into their normal anatomic position, where they can then be fixed to the frontal bone with a microplate. A badly comminuted nasal fracture may have insufficient bone

FIGURE 39–54. Horner's muscle. Reproduced with permission from Donald PJ et al.[51]



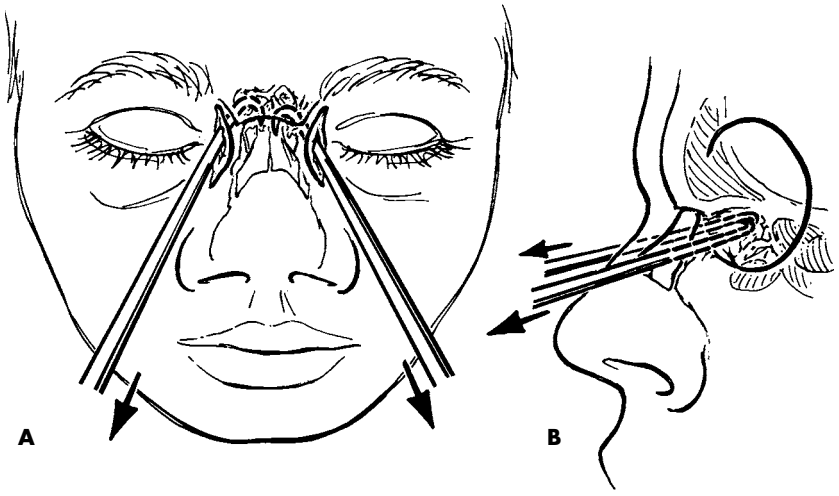


FIGURE 39–55. Reduction of frontoethmoidal fracture. *A*, Stout bone hooks are placed through the lamina papyracea into the ethmoid fracture. *B*, Forward and downward traction is exerted until fracture is reduced. Reproduced with permission from Donald PJ et al.[51]

to hold even the small 0.8 mm screws required for microplate fixation. A solution to this is the use of suspension wires and lead plates.[51] The wires are driven from one side of the nose to the other with the nasal skeleton in the reduced position. The wires penetrate skin, nasal bone, and ethmoid sinuses on one side and then pass through the septum and the same structures of the other side (Figure 39–56). It must be remembered that the plates act simply as anchoring sites for the wires and not as a means of narrowing the nose. Overtightening of the wires can lead to underlying skin necrosis. The wires are removed after 3 weeks.

Obtaining an excellent cosmetic result in the reduction of dislocated medial canthal tendons is one of the most demanding tasks in the surgery of facial trauma. The repair is best done as soon as possible after the accident. Late repairs are notoriously noted for their less than optimal results. Often the tendon comes away with a fragment of attached lacrimal bone. In those instances, two small holes are placed in the bone that will serve as anchor points for the wire used in fixation. If no anchoring bone exists, then some kind of stabilization must be created to prevent the wire from pulling through. The Kazanjian button may be created by twisting the two wires passed through the tendon into the configuration of a rosette.[35] Another method is the use of a Dacron bolster (Figure 39–57), which is usually used to stabilize hemostatic sutures placed in the aorta.[51] In a unilateral canthal repair, the wire can be secured to the bone of the medial wall of the opposite orbit. If enough bone of adequate strength is present on the ipsilateral side, then that would be preferable. A bone graft on the opposite side of the nose can be used or even two holes in a microplate that has been used to stabilize the nasal bones. It is vital to place the tendon high and posterior with some small overcorrection as some degree of relapse is almost inevitable.

The worse injury is the bilateral canthal tendon avulsion. The Converse-Hogan open sky technique is probably the most effective (Figure 39–58).[52] This method wires one canthal tendon to that of the opposite side and supports the skin of the nasal area with lead plates.

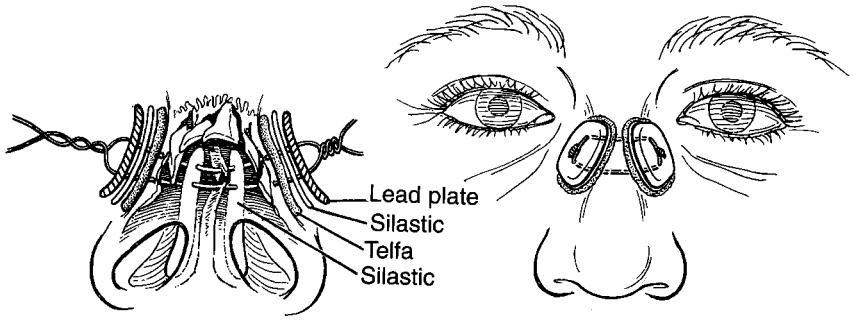


FIGURE 39–56. Nasal bones fixed with lead plates. Reproduced with permission from Donald PJ et al.[51]

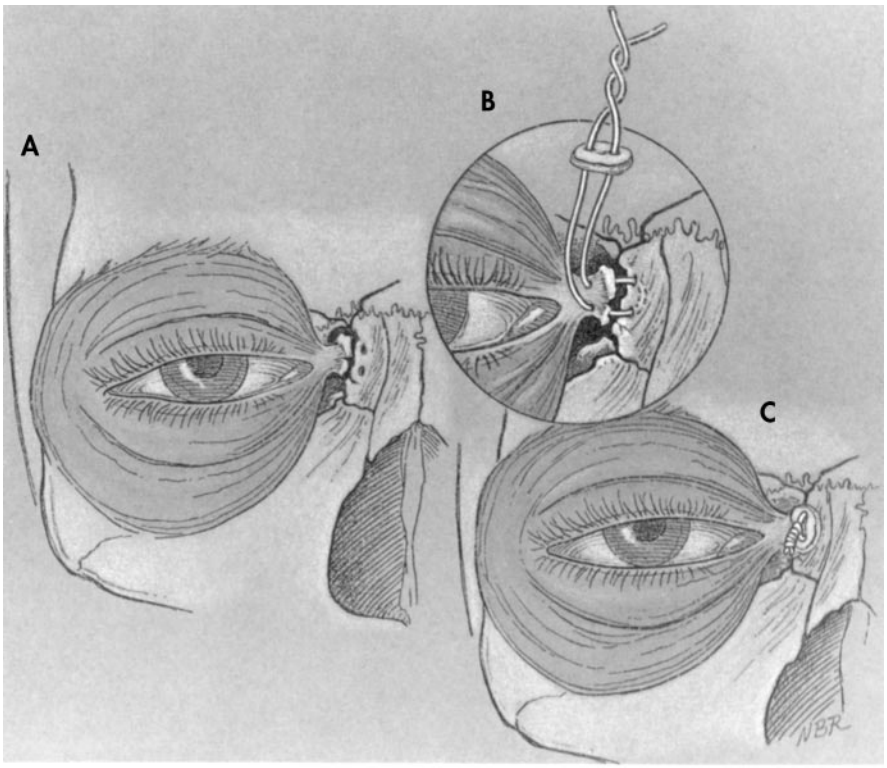


FIGURE 39–57. Standard method of fixation of unilateral medial canthal tendon avulsion. *A*, Fixation to medial orbital wall with wire. Two small holes are drilled in the lacrimal bone. *B*, Wire is passed through the tendon and the bone. Wire is passed through a small Dracon felt bolster to prevent wire tearing through the tendon. *C*, Wire is gently twisted for the ligature. Reproduced with permission from Donald PJ et al.[51]

Bone Grafting

In some patients with midfacial fractures, bone comminution or bone loss prevents proper re-establishment of facial buttresses. Loss of vertical height with midfacial shortening can result. When this condition is recognized, immediate bone grafting is indicated.[48] Grafting is usually performed by harvest of split calvarial bone grafts. These grafts have a good viability and stabilize fixation in cases with significant bone loss. In these patients, intermaxillary fixation usually should remain in place for at least 6 weeks.

In patients with severely comminuted nasal bones in whom an esthetically acceptable nose cannot be created, an onlay graft cantilevered from the forehead using a microplate can restore a more normal nasal dorsum. Small areas of avulsed or missing bone can also be replaced with hydroxyapatite bone cement.[53]

FRONTAL SINUS FRACTURES

Despite the use of the shoulder harness in automobiles, frontal sinus fractures still occur with regularity. The old-fashioned lap belt that restrained the passenger in the seat during the impact of a car crash unfortunately allowed the face to move forward, striking the dashboard with the frontal calvarium and often fracturing the frontal sinus. A considerable amount of force is required to fracture the anterior wall of the frontal sinus because of the arch configuration and its thickness. In Nahum's study, it was apparent that it took more foot-pounds of force to break the frontal sinus anterior wall than any other facial bone, including the symphyseal area of the mandible and the zygomatic body.[54]

The problems of the fracture of the frontal sinus center around the issues concerning the cavity, duct, and mucosa. The cavity on each side has a thick anterior wall but very thin posterior and inferior walls. It has only one point of egress: a funnel-shaped structure at the anteromedial part of the frontal sinus floor that empties into the middle meatus of the nose. According to Lang, 77.3% of individuals have a frontonasal duct and 22.7% have a foramen.[55] A fractured or even damaged duct has a marked propensity to undergo stenosis. The mucosa of the sinus cavity acts in a uniquely pathologic manner when traumatized. It has a tendency not to reconstitute itself with normal-appearing and normal-behaving mucosa as the other paranasal sinuses do when they are injured. In the frontal sinus, damaged mucosa has a propensity for subepithelial fibrosis, especially in the area of

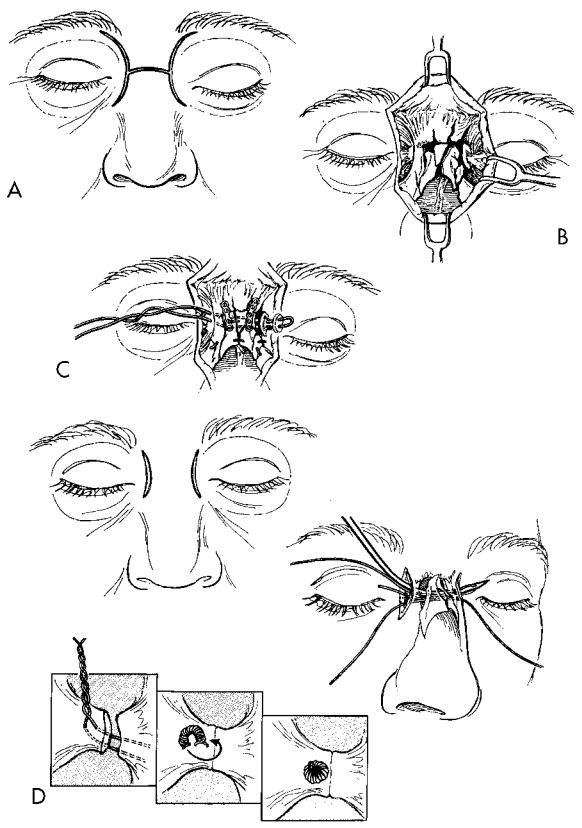


FIGURE 39–58. Bilateral medial canthal tendon avulsion repaired using the Converse-Hogan "open sky" technique. *A*, Incisions are opened. *B*, Fractures are reduced. *C*, Two pairs of wires are inserted—one through the medial canthal tendons (*above*), the second through adjacent nasal bones (*below*). *D*, Wires are tightened, bringing the canthal tendons into position. The wires are twisted together over each avulsed tendon, pulling the tendons together. Tendons are supported and soft tissue is splinted by Telfa pad-Silastic-lead plate assembly (see Figure 39–57).

the frontonasal ducts. Furthermore, it has a tendency to form cysts, some of which will enlarge and erode bone; the resulting mucocele may become infected and form a mucopyocele. Both have the notable characteristic of eroding bone.[56,57]

For the sake of treatment considerations, frontal sinus fractures can be classified as anterior wall, posterior wall, nasofrontal duct, and "through and through." They can be further classified as linear, displaced, compound, and compound with "missing bone." A treatment algorithm has been designed that specifically addresses each site and each type of fracture, bearing in mind that seldom do single fractures occur in isolation. However, with each site and each type, different therapeutic solutions need to be invoked. Therefore, in any given patient, some combination of procedures should be employed to repair the fracture adequately and safely.

Linear fractures of the anterior frontal sinus wall require no operative intervention. Depressed fractures must be opened, the fragments elevated, and all mucosa entrapped between the fragments excised. The fragments are restored to their normal anatomic position and fixed with miniplates. Exposure is through a coronal scalp flap or the so-called "butterfly incision" (Figure 39–59). If the patient has a compound fracture, the exposure can be gained by extending the laceration in a natural crease line in the forehead.[51]

Displaced posterior wall fractures are notoriously difficult to diagnose even with fine-cut CT scanning. In the authors' opinion, all posterior wall frontal sinus fractures should be opened. Even when a small amount of displacement occurs, mucosa can be entrapped and grow into the anterior cranial fossa or form a mucocele. Because the victims of trauma are often very hard to follow clinically, the first subsequent presentation of a patient that has been followed rather than operated on may be in the emergency department with meningitis or a brain

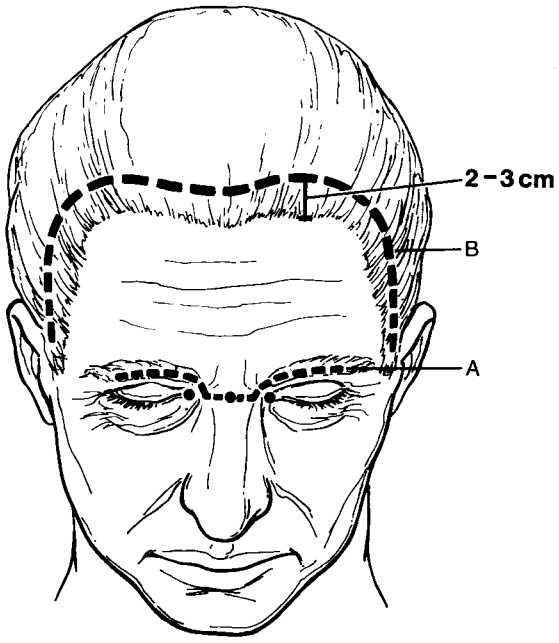


FIGURE 39–59. Incisions used for access for frontal sinus fracture repair. *A*, Butterfly incision. *B*, Coronal scalp incision.

abscess. The treatment of choice is the osteoplastic flap and fat obliteration procedure.

The area of the frontal sinus that is the most difficult to visualize radiographically is the region of the frontonasal duct. The most important radiographic sign is the presence of fluid in the sinus as demonstrated by sinus opacification. If this persists over a 3-week period, the index of suspicion regarding the possibility of a duct fracture is high. If the patient sustained an anterior wall fracture and early intervention for its repair is undergone, then, at the same time, a sinus endoscope can be placed through the site of the fracture and the duct visualized directly. If, however, there is a suspicion of an isolated duct fracture, then a trephine hole in the orbital roof will provide access to wash the sinus out with saline, install some cocaine, and then place a radiopaque dye within the sinus cavity. A subsequent radiograph will show presence of the dye within the nose if the duct is patent and absence of the contrast material if the duct is fractured. Surgical treatment of the fractured duct is somewhat controversial. The author prefers to do the osteoplastic flap and fat obliteration operation to eliminate the duct. Others prefer the Lynch procedure combined with a nasal septal or medially based lateral nasal wall flap, the so-called Sewall-Boyden flap.[58] Removal of the intersinus septum will allow egress of mucus from the fractured to the nonfractured side. More lately, the interest in functional endoscopic sinus surgery has prompted some to do a conservative removal of the duct through the nose. These procedures have yet to stand the test of time.

The most serious of all of the frontal sinus injuries is the "through-and-through fracture."[59] In this dramatic injury, the frontal skin is lacerated, both anterior and posterior walls of the sinus are fractured, often with extensive comminution, the dura is torn, and the brain is contused (Figure 39–60). The frontal sinus is often merely an inferior extension of a larger compound cranial fracture.

Approximately 50% of individuals afflicted by a through-and-through injury die at the scene of the

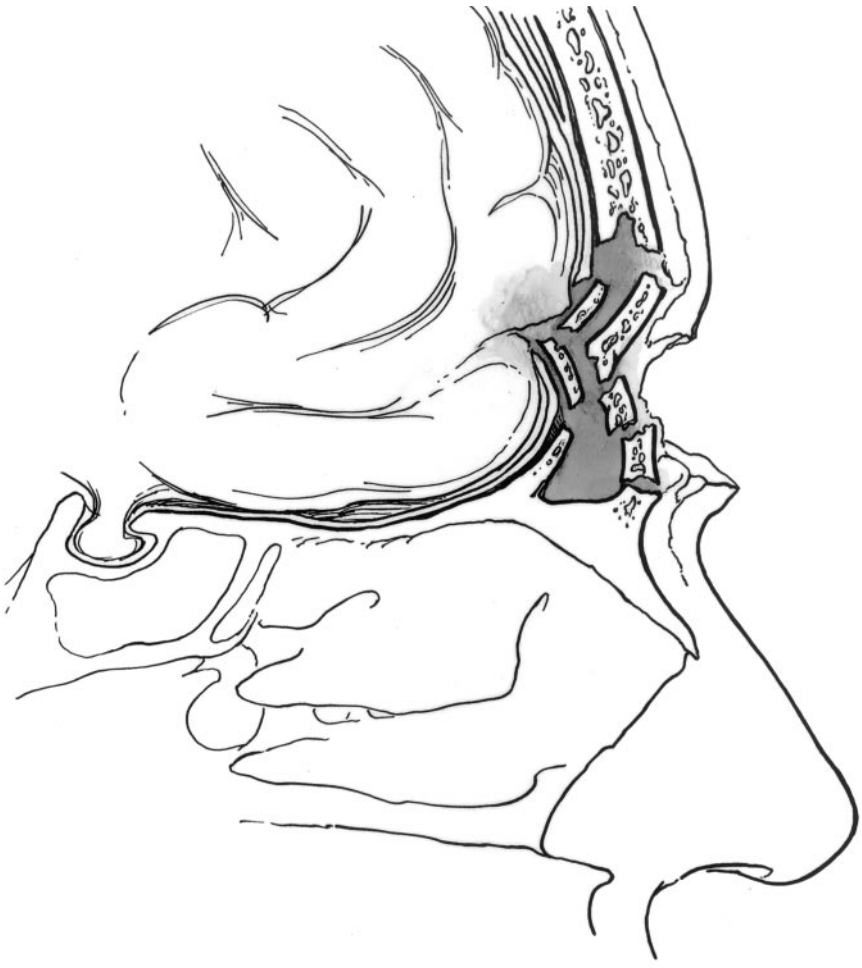
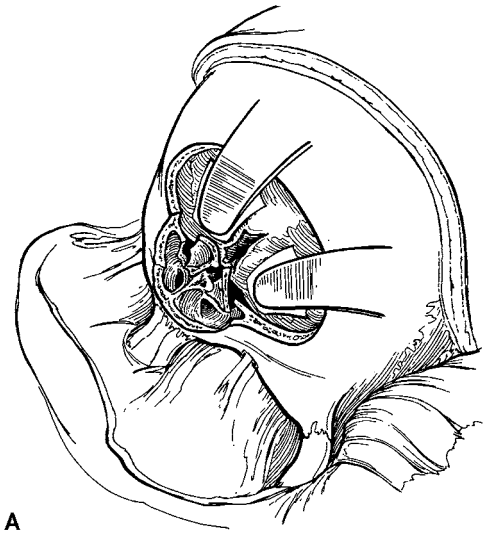
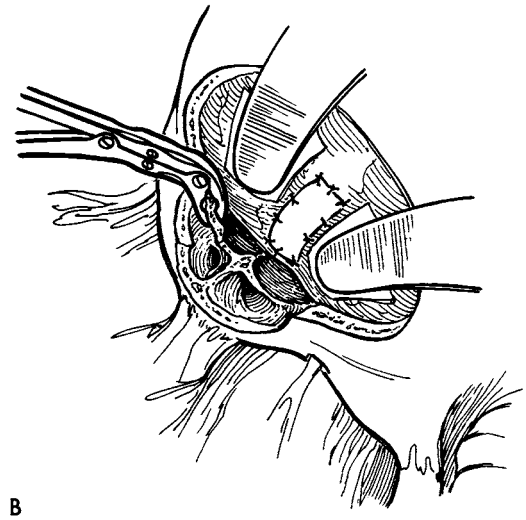
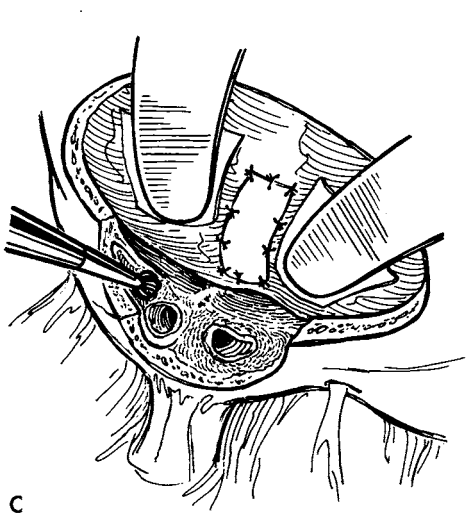
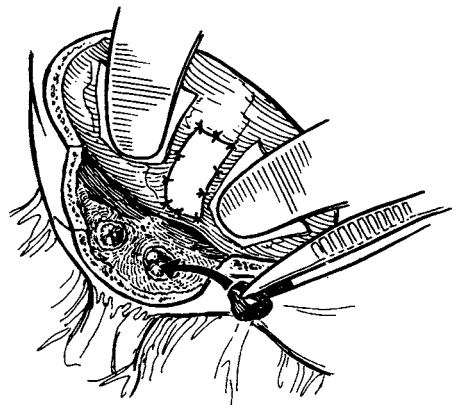


FIGURE 39–60. Through-and-through fracture of the frontal sinus. Reproduced with permission from Donald PJ et al.[51]









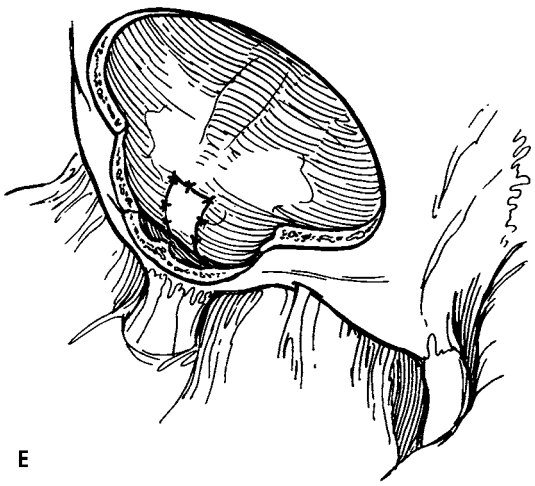


FIGURE 39–61. Repair of through-and-through fracture of the frontal sinus. *A*, Anterior craniotomy done. Anterior wall fragments are cleansed and stored. Fracture of both sinus walls, dural rent, and damaged brain exposed. *B*, Brain débrided, dura patched, and posterior wall remnants removed with double-actioned rongeur. *C*, Removal of the remaining posterior wall of the frontal sinus, producing cranialization. *D*, The mucosa of the sinus is removed from the sinus lumen, the nasofrontal duct mucosa is inverted on itself, and the frontonasal ducts are plugged with temporalis muscle. *E*, Brain migrates forward to fill the defect. Reproduced with permission from Donald PJ et al.[51]

accident. Often the first encounter with the patient by the otolaryngologist is during the craniotomy being done by the neurosurgeon to remedy the patient's intracranial injuries. Usually, a large anterior craniotomy has been done. Once the neurosurgeon has controlled the central hemorrhage, removed the devitalized brain, and repaired the dura, the otolaryngologist needs to manage the severely damaged frontal sinus. This is best done by the cranialization procedure[59] (Figure 39–61). The anterior wall fragments are removed, divested of mucosa, and thoroughly irrigated and débrided to ensure the removal of all of the contaminants acquired at the trauma scene. The bone fragments are stored in povidone-iodine until the end of the procedure. The posterior wall of the frontal sinus is completely removed with a rongeur and a cutting bur. The anterior cranial fossa is now entirely in continuity with the cavity of the frontal sinus. The cutting bur is used to thoroughly remove all vestiges of sinus mucosa including the small projections into the vascular pits in the frontal sinus anterior wall and floor. This is the most important step in the procedure. Any remaining mucosa provides the potential for its regrowth and subsequent infection. The mucosa of the nasofrontal duct is inverted on itself toward the anterior ethmoid sinuses. The duct is plugged with a block of temporalis muscle or bone dust. The anterior wall fragments are restored, and the dural graft will expand with time and fill the new space created in the anterior cranial fossa.

CONCLUSIONS

In summary, treatment of facial fractures requires a multisystem approach. All bony and soft tissue injuries should be diagnosed, and reconstitution of all tissue layers should be performed, if possible. The advancement of technology has enabled rigid fixation to become the standard of care for the fixation of most facial fractures. More precise stability and fixation of fractures have become possible, and intermaxillary fixation is used less frequently. Well-planned incisions minimize scarring. Adequate exposure, precise reduction, and stable fixation remain the hallmarks of treatment of facial fractures.

REFERENCES

1. Dingman RO, Natvig P. Surgery of facial fractures. Philadelphia: WB Saunders; 1964.
2. Sykes JM, Donald PJ. In: Meyerhoff WL, Rice DH, editors. Reconstruction of the head and neck. Philadelphia: WB Saunders; 1992. p. 923.
3. Sykes JM, Murakami C. Principles of local flaps in head and neck reconstruction. Oper Tech Otolaryngol Head Neck Surg 1993;4:2.
4. Angle EH. Classification of malocclusion. Dent Cosmos 1899;41:248.
5. Dingman RO, Natvig P. Surgery of facial fractures. Philadelphia: WB Saunders; 1969.
6. Mathog R. Maxillofacial trauma. Baltimore: Williams & Wilkins; 1984.
7. Chayra GA, Meador LR, Laskin DM. Comparison of panoramic and standard radiographs for the diagnosis of mandibular fractures. J Oral Maxillofac Surg 1986;44:677–9.
8. Hemmings KW. Fracture of the cervical spine complicating bilateral fractures of the mandible. A case report. Br J Oral Maxillofac Surg 1985;23:279–83.
9. Kruger E, Schilli W. Oral and maxillofacial traumatology. Chicago: Quintessence; 1982.
10. Foster CA, Sherman JE. Surgery of facial bone fractures. New York: Churchill Livingstone; 1987.
11. Hoffman-Axthelm W. The treatment of mandibular fractures and dislocations: a historical perspective. In: Kruger E, Schilli W, editors. Oral and maxillofacial traumatology. Chicago: Quintessence; 1982. p. 17.
12. Gilmer TL. A case of fracture of the lower jaw with remarks on treatment. Arch Dent 1887;4:388.
13. Kellman R. Repair of mandibular fractures via compression plating and more traditional techniques: a comparison of results. Laryngoscope 1984;94:1560–7.
14. Szabo G, Kovacs A, Pulay G. Champy plates in mandibular surgery. Int J Oral Surg 1984;13:290–3.
15. Gaynor E, Stephanak W. Prosthetic splinting as an alternative in the treatment of mandibular fractures. Ear Nose Throat J 1983;62:32–43.
16. Chuong R, Donoff R. Intraoral open reduction of mandible fractures. Int J Oral Surg 1985;14:22–8.
17. Donald PJ, Bernstein L. Open reduction for subcondylar fractures of the mandible. In: Bernstein L, editor. Plastic and reconstructive surgery of the head and neck: the 3rd International Symposium. Vol 2. New York: Grune and Stratton; 1981. p. 357–63.
18. Brown E, Obeid G. Simplified method for the internal fixation of fractures of the mandibular condyle. J Oral Maxillofac Surg 1984;22:145–50.
19. Wennogle C, Delo R. A pin in groove technique for reduction of displaced subcondylar fractures of the mandible. J Oral Maxillofac Surg 1985;43:659–65.

20. Fernandez J, Mathog R. Open treatment on condylar fractures with biphase technique. Arch Otolaryngol Head Neck Surg 1987;113:262–3.
21. Ikemura K. Treatment of condylar fractures with other mandible fractures. J Oral Maxillofac Surg 1985;43:810–3.
22. Kellman RM, Marentette LJ. Atlas of craniomaxillofacial fixation. New York: Raven Press; 1995.
23. Freihofer HM, Sailer HF. Experiences with intraoral wiring of mandibular fractures. J Maxillofac Surg 1973;1:248.
24. Larsen OD, Nielsen A. Mandibular fractures: II. A follow-up study of 229 patients. Scand J Plast Reconstr Surg 1976;10:219–26.
25. Michelet FX, Deymes J, Dessus B. Osteosynthesis with miniaturized screwed plates in maxillo-facial surgery. J Maxillofac Surg 1973;1:79–84.
26. Champy M, Lodde JP, Schmitt R, et al. Mandibular osteosynthesis by minature screwed plates via a buccal approach. J Maxillofac Surg 1978;6:14–21.
27. Glineburg RW, Laskin DM, Blaustein DI. The effects of immobilization of the primate temporomandibular joint. A histological study and histochemical study. J Oral Maxillofac Surg 1982;40:3–8.
28. Mekubjian S. Mandibular fracture in a five week old infant. J Oral Maxillofac Surg 1985;43:814–5.
29. McGuirt WF, Salisbury PL. Mandibular fractures: their effect on growth and dentition. Arch Otolaryngol Head Neck Surg 1987;113:257–61.
30. Leake D. Long term follow up of fractures of the mandible condyle in children. Plast Reconstr Surg 1971;47:127.
31. Boyne PJ. Osseous repair and mandibular growth after subcondylar fractures. J Oral Surg 1969;25:300.
32. Close L, Lomba J. Facial reconstruction following blast injury. Head Neck Surg 1983;6:639–52.
33. Shuker S. Management of comminuted mandibular war injuries with multiple circumferential wires. J Oral Maxillofac Surg 1986;44:152–5.
34. Binder P. Evaluation of the eye following periorbital trauma. Head Neck Surg 1978;1:134–47.
35. Converse JM. Surgical treatment of facial injuries. 3rd ed. Baltimore: Williams and Wilkins; 1974.
36. Fryer MP, Brown JB, Davis G. Internal wire pin fixation for fracture dislocation of the zygoma: twenty-year review. Plast Reconstr Surg 1969;44:576–81.
37. Luhr H. Vitallium Luhr systems for reconstructive surgery of the facial skeleton. Otolaryngol Clin North Am 1987;20:573–606.
38. Champy M, Lodde JP, Kahn JL, Kielwasser P. Attempt at systemization in the treatment of isolated fractures of the zygomatic bone: technique and results. J Otolaryngol 1986;15:39–43.
39. Eisele D, Duckert L. Single point stabilization of zygomatic fractures with minicompression plates. Arch Otolaryngol Head Neck Surg 1987;113:267–70.
40. Burres S, Cohn A, Mathog R. Repair of orbital blowout fractures with Marlex mesh and Gelfilm. Laryngoscope 1981;91:1881–6.
41. Weintraub B, Cacin R, Jacobs M. Extrusion of an infected orbital floor prosthesis after 15 years. Plast Reconstr Surg 1981;68:506.
42. Kohn R, Romano PE, Puklin JE. Lacrimal obstruction after migration of orbital floor implant. Am J Ophthalmol 1976;82:934–6.
43. Fujii N, Yamshiro M. Classification of malar complex fractures using computed tomography. J Oral Maxillofac Surg 1983;41:562–7.
44. Godoy J, Mathog R. Malar fractures associated with exophthalmos. Arch Otolaryngol 1985;111:174–7.
45. Le Fort R. Etude experimentale sur les fractures de la machoire supèrieure. Rev Chir (Paris) 1901;23:208–360.
46. Stanley RB. Reconstruction of the midfacial vertical dimension following Le Fort fractures. Arch Otolaryngol 1984;110:571–5.
47. Manson P, Hoopes J, Su C. Structural pillars of the facial skeleton: an approach to the management of Le Fort fractures. Plast Reconstr Surg 1980;66:54–62.
48. Manson P, Crawley W, Yaremchuk M. Midface fractures: advantages of immediate extended open reduction and bone grafting. Plast Reconstr Surg 1985;76:1–12.
49. Close LG. Fractures of the maxilla. Ear Nose Throat J 1983;62:365–70.
50. Gunter H. Konstitutionelle Anomalien des Augenabstrandes und der Interorbitalbrilite. Virchows Arch A Pathol Anat Histopathol 1933;290–373.
51. Donald PJ, Gluckman JL, Rice DH. The sinuses. Vol I. New York: Raven Press; 1995.
52. Converse JM, Hogan VM. Open-sky approach for reduction of naso-orbital fractures. Case report. Plast Reconstr Surg 1970;46:396–8.
53. Friedman CD, Constatino PD, Tagaki S, Chow LC. BoneSource hydroxyapatite cement: a novel biomaterial for craniofacial reconstruction. J Biomed Mater Res 1998;43:428–32.
54. Nahum AM. The biomechanics of maxillofacial trauma. Clin Plast Surg 1975;2:59–64.

55. Lang J. Clinical anatomy of the nose, nasal cavity and paranasal sinuses. New York: Georg Thieme Verlag; 1989.
56. Donald PJ. Tenacity of frontal sinus mucosa. Otolaryngol Head Neck Surg 1979;87:557–66.
57. Donald PJ, Ettin M. The safety of frontal sinus fat obliteration when sinus walls are missing. Laryngoscope 1986;96:190–3.
58. Barron SH, Dedo HH, Henry CR. The mucoperiosteal flap in frontal sinus surgery (the Sewall-Boyden-McNaught operation). Laryngoscope 1973;83: 1266–80.
59. Donald PJ. Frontal sinus and nasofrontoethmoidal complex fractures. Self-instructional package. No. 804000. Alexandria (VA): American Academy of Otolaryngology-Head and Neck Surgery; 1980.

CHAPTER 40

Reconstruction of the Maxilla and Mandible

David Reiter, MD, DMD

The goals of head and neck reconstructive surgeons, namely, the restoration of function and esthetics in an expeditious manner with minimal surgical morbidity,[1] are tested as rigorously by defects of the jaws as by any other problem we face. We are in the midst of a revolution in reconstructive surgery. Advances in vascularized tissue transfer and custom fabrication of reconstructive flaps are occurring rapidly. Although research propels us toward a future in which we will no longer need to remove tumors and provide replacement tissues, we must maintain currency with such methods while we wait for what O'Leary et al aptly termed in 1994 "the molecular reconstructive 'cocktails' from the shelf of the next millennium."[2]

No more than 20 years ago, a chapter on reconstruction of the jaws would have focused on the mandible, for which there were many surgical options (no matter how unsatisfactory they might have been). Maxillary defects were almost exclusively a problem for the maxillofacial prosthodontist, and few surgical options existed. With the perfection and acceptance of vascularized free tissue transfer, mandibular repair and replacement have become routine. After a decade of clinical experience, the fibula has emerged as most facial plastic and reconstructive surgeons' donor site of choice for management of mandibular defects.

Conversely, surgical repair of most maxillary defects was limited to rudimentary adaptation of regional flaps until the 1990s. Whether because of pessimism over the potential to achieve successful surgical reconstruction or because of optimism that ongoing research in free tissue transfer would soon produce new methods far superior to the compromises then available, the literature of maxillary defect management focused largely on prosthetics until the end of the twentieth century. As free tissue transfer became routine, interest in surgical reconstruction of the maxilla accelerated. A MEDLINE search for articles on "maxilla and reconstruction" produced 11 works from 1960 to 1980, 51 from 1980 to 1990, and 96 from 1998 to 2000. There are currently multiple soft and hard tissue options in clinical use. Rapid refinement and ongoing research in methods and materials for maxillary reconstruction now recapitulate the evolution of mandibular reconstruction through the 1990s. It is an exciting time for head and neck and facial plastic surgeons interested in reconstruction of the jaws.

The jaws are the nucleus of many of the most sophisticated structural systems in the body. Although other bones in the body support more weight or more directly protect individual vital structures, only the jaws serve so many functions in so many ways. Masticatory forces approach 10,000 pounds per square inch between upper and lower molars, and load rates exceed 2,000 N/second.[3] Yet a grain of sand measuring only 0.06 mm between the same teeth causes the muscles of mastication to release their load within hundredths of a second. The jaws anchor and support the upper portals of the aerodigestive tract. They surround or otherwise protect important neurovascular conduits, including some of the major extracranial vasculature of the head and neck. They house and nurture developing tooth buds in the neonate and child. They facilitate respiration, speech, chewing, swallowing, communication, facial expression, and thus social interaction.

When part or all of the maxilla or mandible is lost to disease or injury, it would be ideal to provide exact restoration or reconstruction. Were it possible to provide a perfect replacement, function and appearance would not suffer, and the patient would have no post-treatment deficit. Although advances in modern reconstructive methodology have brought us closer to this goal, it remains elusive.

A damaged structure can be restored (returned to its normal state using original or identical materials and methods), reconstructed (brought to its normal state using new or additional materials), or replaced. All three alternatives exist for management of defects or loss of the jaws, although some of the most sophisticated restorative techniques are not yet available for routine clinical use. This chapter will address the management of each defect not only in terms of lost structure but also with consideration of functional deficits. As so many advances in head and neck reconstruction are being made every month, we will discuss some of the most promising methods on the cutting edge despite their current status as investigational. History suggests that many will have become routine between preparation and publication of this work.

Table 40–1 summarizes the spectrum of structural and functional deficits attendant on loss of part

TABLE 40–1. Defects, Deficits, and Options for Management of the Maxilla and Mandible

Defect	*	*Structural Deficit*	*Functional Deficit*	*Reconstructive Options*
			Maxilla	
Oroantral fistula	1–3	Lost integrity of antrum	Loss of air and water seal	Local flaps, buccal fat pad
Alveolar ridge loss	2–3	Tooth-bearing tissue	Dentition lost, possible oroantral fistula	Prosthesis, bone graft, distraction
Palatal shelf	1–3	Teeth plus antrum/ nasal cavity	Chewing, swallowing, nasal regurgitation, speech	Prosthesis, free flap (myocutaneous, osteomyocutaneous)
Hemimaxilla	2–4	All of above plus lip and check support	All of above	All of above
Hemimaxilla plus orbital floor	3–4	All of above plus orbital integrity and support	All of above plus diplopia	All of above
Hemimaxilla plus orbit	4–5	All of above plus contents of orbit	All of above plus sequelae of exenteration	All of above; megaflap
More than hemimaxilla	5	All of above plus midfacial structure	All of above plus nasal/ lip/midfacial support	All of above
			Mandible	
Body posterior to mental foramen	2	Lower facial support, part of dental arch	Chewing, swallowing, appearance	Plate, soft tissue, bone graft or flap (fibula, scapula, iliac; radius if small defect)
Ramus	3	Jaw stability and function	Chewing, swallowing, appearance	Plate, free flap (fibula)
Condyle	3	As above	As above	Plate with prosthesis, free fibula
Anterior arch	4–5	Tongue, lip, and facial support, teeth	Eating, speech, airway integrity, appearance	Free flap (fibula, scapula, iliac)
Hemimandible	3–4	All of above	All of above	Plate, free flap (fibula, scapula, iliac crest)
More than hemimandible	5	All of above	All of above	Free flap—fibula is overwhelmingly tissue of choice

*Relative severity on a 1 to 5 scale (5 = worst).

or all of the jaws, along with relative severity and current recommendations for management.

RESTORATION

Restoration is the building up of something to its original and complete form using original materials and methods.[4] Distraction osteogenesis (DO) is probably the only management method with any clinical utility at present that approaches true restoration of lost bone. It spans or augments bony defects by traction-induced neo-osteogenesis of the remaining original structure. Anecdotally accepted to have been introduced by Cordovilla in 1905 and first published by Putti in 1921,[5] this technique was brought to practical clinical utility with work by Ilizarov and others in the 1990s.[6] The methodology includes interruption of the cortical envelope (if intact) with maintenance of periosteal integrity to the mobilized end plate. Pins or other secure devices are anchored in the mobile segment of bone, and controlled force is applied to the anchors. Healing occurs within the periosteal envelope by neo-osteogenesis between the advancing end plate and the remaining bone. Segmental ends can be advanced toward each other to close gaps, and anchors in depressed or malaligned segments can be used to direct neo-osteogenesis for augmentation or realignment of the defect. There are now commercial appliances and systems available for DO, and the literature is replete with references to its clinical utility in management of defects of the jaws (Figure 40–1).[7–9]

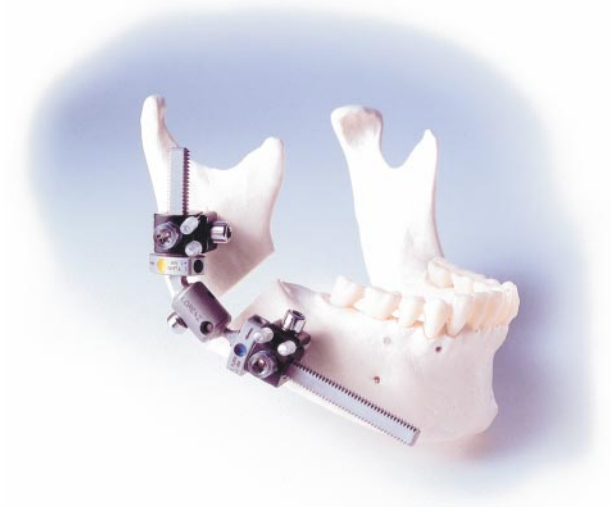


FIGURE 40–1. External distraction osteogenesis device. Courtesy of W. Lorenz Surgical.

Constantino et al demonstrated the ability to fill 2.5 cm defects in dog mandibles using DO in 1993.[10] Regeneration was achieved at a rate of 1 mm per day, and function was normal at 1 year in all test animals. More recent studies verify this work and support its promise.[11,12] As mechanical limitations in appliance design are overcome, DO will become a routine clinical tool. Recent reports of clinical use in humans support the impression that DO will soon be a mainstay in reconstruction of bony defects of the jaws.[13,14]

The use of osteoinductive morphogens, although not yet a routine clinical tool, offers a second alternative for restoration of bony architecture. Bone morphogenetic proteins (BMPs) belong to the family of transforming growth factor-beta substances and play an important role in the growth and development of numerous tissues.[15] Work to date is promising, but clinical application of BMPs to stimulate replacement bone has not reached the stage of practical utility. Work is being done on application of BMPs in absorbable collagen sponge[16] and different forms of hydroxyapatite,[17] but such techniques have not yet reached practical application. It appears clear that BMPs will play a significant role in the true restoration of bony architecture in the head and neck as soon as the proper combination of morphogen, carrier, and methodology is found.

RECONSTRUCTION

To reconstruct is to rebuild.[3] There are many options for reconstruction of the jaws. Modern osseous autografting dates to 1867 and the pioneering work of Ollier.[18] Tibial bone was used prior to World War I, with the iliac crest rising in popularity as a donor soon after World War II.[19] Calvarial bone has been used for about 15 years but was boosted in popularity by the publication of a technique for temporal harvesting by Spear and Wiegering in 1987.[20] Rib has long been used for grafting in the head and neck. The safety and utility of autogenous bone grafting were documented by series such as Kline and Wolfe's 1,000 personal cases and 12,672 surveyed cases.[21] Thus, free autogenous bone grafting was a workhorse in the management of maxillary and mandibular defects until the era of practical free tissue transfer, despite the relatively high failure rate in recipient beds contaminated by oral secretions.[22] Osseous autografts remain useful in the management of small (< 4 cm) defects of the mandible and

in closure of many defects resulting from maxillectomy for which vascularized osteocutaneous free tissue transfer is neither required nor justified.[23] So, despite the wide acceptance of vascularized free tissue transfer and its use by the author since 1993, free nonvascularized bone grafts continue to have a limited but important role in head and neck reconstruction. There is even recent literature on improving the outcome of osseous autografts.[24]

Osseous allografts have been used in jaw reconstruction for many years. Freeze-dried irradiated fibula and other processed cadaver tissues were popular for a brief period in the 1980s, and autografting of explanted tumorous mandibles after chemical manipulation and irradiation to 120 cGy was advocated briefly.[25] However, more modern methods of management have largely relegated the above to historical significance.

The use of bone-like alloplastic substances as scaffolding for neo-osteogenesis is well documented and supported by histologic[26] and clinical[27] reports. One exception to this is the use of hydroxyapatite in mandibular defects secured with a reconstruction plate,[28] a technique not supported by any current literature or experience.

Functional and Cosmetic Reconstruction

Functional reconstruction is defined as creation of sufficient structure to return an organ or region to function, even though neither function nor structure exactly replicates the pretreatment or preinjury state. This describes the vast majority of flap and graft procedures as well as the use of hardware and/or soft tissue to replace bone. Appearance and/or function are often enhanced by substituting tissue of another kind for the resected specimen. Thus, for example, resected tongue tissue is routinely replaced with regional or free flaps despite the current impossibility of restoring volitional mobility to the reconstruction. Perioral soft tissues can be replaced with a variety of soft tissue flaps and grafts, but this will not restore the function of the oral sphincter. Similarly, ablation of a maxillectomy defect with soft tissue and/or bone may enhance oral function and/or appearance, but the region cannot be returned to its prepathologic state.

Soft tissue can also be used to augment a lateral mandibular defect, effacing the cosmetic deformity of mandibular loss and deviation but not restoring the native structure or its actual function. However, the reported success rate of this technique falls off rapidly beyond the first postoperative year, even in lateral mandibular defects posterior to the mental foramen (the most forgiving location in terms of reconstructive requirement and success). Reports of 40 to 50% failure after 1 year are common for techniques of mandibular reconstruction using a combination of soft tissue and plating, even when vascularized free tissue transfer is used for the soft tissue component.[29] Plate breakage is common, as is exposure, leading some authors to recommend plates and soft tissue only for temporary use in mandibular reconstruction.[30] Most authors have found this to be true regardless of the kind of plate chosen. One study suggests that modern plate technology and materials science have not lessened the incidence of plate exposure and have not eliminated the problem of plate breakage,[31] although another found the latest lower-profile mandibular reconstruction plates to have breakage and extrusion rates of about 4% each in a series of 27 cases,[32] which is certainly an acceptable rate of failure, given the problem and alternatives for management.

The wide variation in reported plate breakage rates demonstrates, at least in part, the sensitivity of this complex methodology to technical factors such as the mechanics of hardware placement. The use of a given plate-and-screw system must be taught rigorously and in detail if the method is to achieve clinical success. The use of non–self-tapping screws requires meticulous attention to the process of drilling and tapping holes in bone. If the drill and tap are not inserted and maintained perfectly straight along the desired axis or the threads are not cut perfectly perpendicular to the axis of the hole, the screw will be loose in the hole and stabilization will not be achieved. Even the use of self-tapping screws requires great care in the drilling of the holes and driving of the screws, for the same reasons.

On a temporary basis, the mandibular bone segments to either side of a defect can be stabilized and the defect spanned with a number of plating systems. Although far from a complete functionally restorative technique, proper selection and placement of a plate can provide a patient with satisfactory appearance, stabilization, and comfort for relatively normal social function, with the notable exception of chewing and swallowing. Liquids can be taken fairly

well by patients having such procedures, and the success rate is commonly reported as adequate for patients with poor prognoses or insufficient physiologic reserve to remain under anesthesia on the operating table for an additional few hours for a more formal reconstruction. However, Komisar reported no significant improvement in functional or social rehabilitation of mandibulectomy patients from simple restoration of mandibular continuity with a plate in an analysis of patients for whom vascularized tissue transfers were not provided.[33]

VASCULARIZED FREE TISSUE TRANSFER

Vascularized free tissue transfer has reached routine status in the United States in every academic and most major community medical centers for patients of all ages.[34] A decade ago, one could intelligently question the value of free flap transfers because long-term outcome studies had not been done, few surgeons and medical centers had the technical ability and physical plant required for optimum microsurgical management, and techniques had not been refined sufficiently to offer consistently high success rates. However, each of those questions has been answered well. Numerous studies support the statement that vascularized free flap reconstruction is both highly successful and cost effective for appropriately chosen head and neck surgery patients.[35,36] See Chapter 41 for the spectrum of available free flaps.

MANAGEMENT OF MAXILLARY DEFECTS

Goals of Management

The goals set forth by Hoopes and Edgerton in 1966 for surgical reconstruction of maxillectomy defects remain as valid today as when they were published, namely, obtain a healed wound, restore palatal competence and function, support the orbit or fill the orbital cavity in exenteration, obliterate the maxillectomy defect, and restore facial contour.[37]

Relevant Maxillary Anatomy

The maxilla has been described as a geometric structure with six walls[38] (Figure 40–2). The roof of the maxilla is the floor of the orbit above it. The medial wall delineates the lateral limit of the nasal cavity and contains the nasolacrimal duct. The maxillary floor consists of the hard palate and the alveolar ridge. The walls of the maxilla delineate the maxillary sinus, being contained within the central portion of the maxilla. In addition, many of the muscles of facial expression and mastication are anchored on the maxilla. The orbicularis oris and oculi, zygomaticus, buccinator, and various levators and depressors, together with the overlying skin and/or intraoral mucosa, constitute the units of the lower eyelid, cheek, upper lip, and oral commissure. The medial and lateral pterygoid muscles originate at the pterygoid plates along the posterior wall of the maxilla. Finally, the two horizontal and three vertical buttresses of the maxilla are responsible for midfacial projection and vertical facial height.

Muzaffar et al called the maxilla the "keystone of the midface."[39] They elegantly described it as uniting "… the key mid-facial elements—the orbits, the zygomaticomaxillary complex, the nasal unit, and the stomatognathic complex" into a functional and esthetic unit. Most maxillary defects are thus complex in nature, involving more than one surface or cavity and affecting more than one function of the midface. Quoting Muzaffar et al again, "In reconstructing maxillary defects secondary to trauma, ablative tumor surgery, or congenital deformities, the surgeon must obliterate the defect, provide adequate struc-

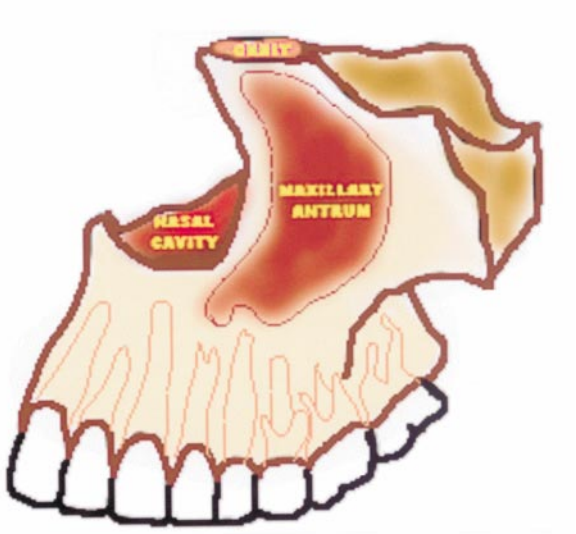


FIGURE 40–2. Maxillary anatomy.

tural support to each of the midfacial units, and address the essential functions of the midface."[39]

The maxilla is often described as being a trilevel structure, with unique functions attributed to each level. The floor of the maxilla houses the upper dental arch and separates the oral cavity from the antrum. The middle level contains the antrum and the path for its aeration. It serves as a vertical buttress for dissipation of masticatory forces and lends support and definition to the lips and the soft tissue vault of the nose. The upper level defines and supports the orbit and cheek and carries vasculature and sensory innervation to the midface.

The maxillae separate several cavities from one another. The alveolar ridges and hard palate separate the oral cavity from the antra, and the central portion of the hard palate separates the oral cavity from the nasal cavity. The medial maxillary walls separate the antra from the nasal cavities, and the orbital processes separate the orbital contents from the antra. The posterior walls of the maxillae separate the antra from the pterygopalatine fossae.

Spectrum of Maxillary Loss

The tissue deficit from maxillary surgery or injury may be as small as a single dental alveolus or as large as the entire party wall between the nasal and oral cavities. Even a tiny defect such as an oroantral fistula through the alveolus of an extracted tooth can cause major dysfunction and require repair. Further, surgical repair of even a tiny defect such as this has a significant failure rate, and much effort has been directed toward such reconstruction.[40] Management of an aggressive maxillary sinus tumor can create a defect that spans the distance from the oral cavity to the anterior part of the skull base. Repair may require re-creation of the roof of the mouth, the floor of the orbit, and support for the cheeks and lips. Aggressive ablation of such a tumor may also require resection of the orbital confines and contents, further complicating reconstruction.

Maxillofacial Prosthodontics

Prostheses have been the mainstays of management of maxillary defects for hundreds of years. Most maxillary defects resulting from oncologic surgery were managed with removable prosthetic obturation until the last decade of the twentieth century.[41] As long as there is a defect amenable to obturation, a prosthesis remains a viable option for many patients suffering loss of part or all of the maxilla, despite the compromises inherent in such management. Dental reconstruction and rehabilitation can be integral with palatal defect management by including prosthetic dentition in the prosthetic unit. Of course, retention of a prosthesis requires sufficient hard and/or soft tissue to anchor, support, and stabilize the device.

Removable palatal prostheses are more acceptable to patients comfortable with the concept of removable dental prostheses than to those with a full complement of natural teeth, and the edentulous proportion of most of the world's population is dropping. A recent British study found so predictable an improvement in the oral health of the general population that a lack of natural teeth will affect 25% of the British population between 65 and 74 years of age in 2008, down from 52% 20 years before.[42] In the United States, the rate is expected to decline from an observed 37% of all Americans in 1997 to 15% in 2020.[43] This suggests a decline in the number of future patients likely to accept and be satisfied with removable maxillary prosthetics. But it is important to note that education, economic status, and social function (the same socioeconomic factors that correlate strongly with facial trauma and with cancer of the head and neck) are independent predictors of low functional dental status.[44] As the population of the United States is growing at a rate close to that of reduction in tooth loss,[43] the absolute number of edentulous patients in a given age group is expected to remain constant. These are the patients who are both at greatest risk for head and neck cancer and least likely to expend the resources for state-of-the-art reconstruction. As a result, the role of removable maxillofacial prosthodontics in maxillary reconstruction will not likely dwindle below its current level of activity.

On the other hand, a progressive increase in oral health awareness and improvement in the oral health of civilized societies suggests an increasing demand for dental rehabilitation as an integral component of maxillofacial reconstruction. Advances in implantology have made osseointegrated implantation of posts, anchors, and other means of securing dental arch restorations and external maxillofacial prostheses readily available to those with the economic resources to obtain them. The spectrum of available management methods

now ranges from a removable obturator with rudimentary dentition and facial support to vascularized free osteocutaneous flap reconstruction with osseointegrated dental implants and fixed full-arch dental restorations. As a result, it is now more important than ever to identify the needs and desires of individual patients regarding maxillofacial reconstruction and rehabilitation.

One of the biggest problems facing the patient in need of a maxillary prosthesis is the availability of a maxillofacial prosthodontist. As Gotay and Yates pointed out, "… the majority work in private practice or a dental school, are primarily involved in clinical work or teaching, and see few cancer patients. Inadequate reimbursement and lack of collaboration with surgeons were cited as barriers to involvement in cancer patient care."[45] Cost-benefit analyses have been done comparing microsurgical reconstruction with traditional regional flap methods, but no such analysis appears in the literature comparing maxillofacial prosthodontic management with surgical reconstruction. As attention is directed toward the cost effectiveness of clinical methodology, removable prosthetic management of maxillary defects may emerge as a societal choice. However, it is far from ideal and necessitates many functional and esthetic compromises much better addressed by modern methods of tissue transfer.

Although head and neck surgeons traditionally express concern about the quality of prosthetic devices available to their patients, studies have shown that "Correlations between anatomic conditions and denture quality and patient satisfaction are weak."[46] Further, authors repeatedly suggest that psychological factors predominate in the acceptance of and adaptation to removable dentures. As a result, the lack of insurance coverage for maxillofacial prosthodontics is the biggest single factor controlling the availability of these important clinical services to head and neck cancer patients.

There are many patients for whom a removable prosthetic management plan is preferable to surgical reconstruction of maxillary defects. This group includes patients who have had multiple recurrences, patients with medical compromise affecting the likely success of surgical reconstruction, and patients whose personal philosophies favor seeking the quickest, simplest treatment that will allow them reasonable function despite some major compromises from ideality.

Surgical Considerations for Optimal Management with an Obturator

Because removable prosthodontic management of maxillectomy defects remains a valuable tool for the head and neck surgeon, it is appropriate to discuss the special surgical considerations necessary to optimize results. Davison et al included concise guidelines for maxillectomy to facilitate the best possible fit, retention, and function of an obturator in a recent work on reconstruction of maxillectomy defects.[47] These include the following:

- Leaving the largest amount of hard palate possible without compromising margins
- Preserving periodontal support for sound teeth adjacent to the defect (eg, making adjacent bony cuts through the center of the next alveolus, not between)
- Lining the soft tissue cavity with split-thickness skin
- Positioning the graft so that the resulting scar band at its junction with residual buccal mucosa aids retention of the obturator flange
- Removing turbinate if it might compromise seating of an adequate obturator
- Preserving nondiseased medial palatal mucosa for use as coverage of the medial bony margin
- Resecting all of the soft palate when more than half must be taken anteroposteriorly for tumor management

Surgical Management of Maxillary Defects

Oroantral Fistulae and Other Defects of the Posterior Part of the Maxilla Although maxillary reconstruction is most often associated with management of head and neck cancer, the most common causes of oroantral communications are dental.[40] Communication between the oral cavity and the maxillary antrum causes chronic sinusitis when it persists over time, compounding the problem of achieving closure of defects resulting from dental disease and its treatment.

Unfortunately, the historical mainstay of management for most maxillary defects, the palatal obturator, is of little help in managing oroantral fistulae in the edentulous patient. The most important factor in retention of an upper dental prosthesis is the integrity of the seal of the base of the appliance against the oral tissues. An oroantral communica-

tion allows air to get beneath the flange of the prosthesis, preventing the development of an adequate seal to retain the device.

This problem deserves specific mention both because of its commonality and the generally acknowledged difficulty in achieving closure.[48] Oroantral fistulae have perplexed the surgeon for centuries. Closure techniques reported in the latter half of the twentieth century include gold foil (a method in intermittent use for over a hundred years),[49] methylmethacrylate,[50] skin flaps, mucosal flaps,[40] and transposition of the buccal fat pad.[51] Although most otolaryngologist-head and neck surgeons have tried to facilitate antral drainage and aeration in the belief that this enhances the success of closure of palatal defects that communicate with the antrum, there is literature suggesting that this is not necessary.[52]

As summarized by Davison et al,[47] flaps of buccal, alveolar, and/or palatal mucosa or mucoperisoteum remain the most commonly used management methods for relatively small openings through the posterior part of the palate into the maxillary sinus. There is growing evidence, however, that the buccal fat pad produces reliable closure with fewer drawbacks than other methods in current use.[51]

Historical Approaches to Maxillary Reconstruction Early attempts at surgical reconstruction of maxillary defects used local or regional soft tissue flaps with limited success.[53,54] Perhaps the first staged maxillary reconstruction to use both bone and soft tissue was reported by Campbell in 1948.[55] He used temporalis muscle and palatal mucosa in the first stage, followed by placement of an iliac crest graft. This reconstruction was reportedly able to retain and support a traditional maxillary denture. Subsequent work by numerous surgeons saw methodologic progress through pedicled regional flaps[56] to pedicled regional flaps with free nonvascularized autogenous bone grafts[39] to free vascularized tissue transfers of all kinds. Prior to the advent of routine free tissue transfers, the mainstays of maxillary reconstruction were pedicled myocutaneous flaps with free nonvascularized cranial, iliac, or rib bone autografts.

Various combinations and permutations of vascularized and nonvascularized soft and hard tissue were then tried, including such innovative juxtapositions as free iliac bone with vascularized jejunal transfer.[57] Advances in microsurgery facilitated major advances in maxillary reconstruction, with reports of success using groin,[58] latissimus,[59] lateral arm,[60] scapula,[61] and radial forearm[62] flaps.

Current Management of Maxillary Defects

Functional wound closure, to prevent central nervous system infection and eliminate dead space, has a higher priority in reconstruction of midfacial defects than does cosmesis.[63] Further, despite major advances in midfacial reconstruction, a prosthesis remains an important part of rehabilitation for many maxillectomy defects in conjunction with vascularized free tissue transfer.[64] Thus, optimal current management of defects resulting from maxillectomy requires an armamentarium encompassing sophisticated surgical skills, prosthodontics, and a rehabilitative team for residual deficits in communicating and taking nutrition.

Optimal management of maxillary defects requires an understanding of the composite nature of the tissue loss. Reconstruction of communicating maxillary defects most often requires replacement of three separate tissue layers, those being the soft tissue linings of the two cavities whose party wall has been violated and the bone of the wall itself. If one considers the functions of these layers rather than their histology, one recognizes that the bone is primarily a barrier to penetration of the soft tissue on either side, with dissipation of forces a secondary but by no means insignificant contribution. Thus, if soft tissue of sufficient strength and puncture resistance could be provided, it might not be necessary to replace the bony component of every communicating maxillary defect. Alternatively, because the consequence of penetration of the party wall between the oral cavity and the antrum is transfer of air and contaminated substances into the antrum, simple obliteration of the antral cavity might suffice in lieu of restoration of bony integrity. Then there would be no open cavity into which penetration could allow spread of alien substances. Both approaches are in current use.

Local and Regional Flaps for Maxillary Reconstruction

Contemporary Use of Local Flaps Even composite defects involving skin and muscle and oroantral communication may be amenable to reconstruction

with local flaps alone. Creative use of cheek, buccal fat, and palatal mucosal flaps may produce excellent results in carefully selected patients.[65] Even the laterally based forehead flap, which was once quite popular for midfacial and intraoral reconstruction,[66,67] still has occasional use.

Temporalis Muscle Flaps. Flaps based on the temporalis muscle are successful and popular for management of intermediate-sized defects of the posterior part of the maxilla, especially when the orbital floor is included in the resection. Temporalis flaps can be muscular,[68] fascial,[69] myofascial,[70] myocutaneous,[71] or osteomyocutaneous.[72] Flaps based on the temporalis muscle are often mobile enough to use on contralateral defects,[73] especially if extended to include attached galea.

The placement of free nonvascularized bone within a temporalis flap has been described for maxillectomy defect reconstruction.[74] However, recent work has suggested that the vascular source for neoepithelialization of the flap's surface originates in the tissue surrounding the defect rather than the flap, possibly explaining less than universal success from use of a free nonvascularized bone graft within a vascularized flap in this application.[75]

Buccinator Musculomucosal Flap. Pribaz et al reported 18 reconstructions of midfacial defects of all kinds using a posteriorly based musculomucosal flap of buccinator that they believed was nourished by the facial vessels.[76] Although not widely known or used, this flap seems to offer a valuable alternative to regional flaps, especially when other local tissue is not available. Licameli and Dolan determined with cadaver dissections that the blood supply is from the buccal (buccinator) artery (a branch of the internal maxillary) and published a more recent series of reconstructions of the soft palate and retromolar trigone with this flap to support further its use.[77]

Chest Flaps The pectoralis major myocutaneous flap has been used for reconstruction of large maxillectomy defects.[78] Although it will fill the defect resulting from a total maxillectomy, with the bulk of muscle obliterating the antral cavity and the orbit if necessary, it does not readily reach high enough without resection of the medial third of the clavicle, and dental rehabilitation is difficult or impossible.

The deltopectoral flap has also been used for obliteration/reconstruction of maxillectomy defects, as typified by Konno et al's series of 46 patients reported 20 years ago.[79] However, this is primarily of use when free tissue transfer is not an option.

Microvascular Free Flaps for Maxillary Reconstruction Taylor et al's original report of successful transfer of vascularized fibula presaged the modern era of jaw reconstruction.[80] Subsequent investigators refined the principles and practice of vascularized fibular transfer, with Hidalgo et al's seminal work over the last decade of the twentieth century presenting and later confirming the reliability and versatility of free osteocutaneous tissue transfer for head and neck reconstruction.[81]

The scapula, fibula, and radius are the most frequently described osteocutaneous flaps for reconstruction of the maxillary region. Each donor site has its own advantages and disadvantages. To quote Schusterman et al,

> … no one flap can be used for all defects; certain flaps are better suited for specific situations. The radial forearm flap is the best choice for small defects around the orbit because it provides a small segment of bone and a thin, pliable paddle of soft tissue. The fibula, iliac crest, and scapula flaps are useful for larger defects involving the maxilla, where bulk and a strong bone buttress are required for prosthetic support and stabilization.[64]

The scapula can be used after an osteotomy to reconstruct the orbital floor and the maxillary buttress simultaneously. Separate skin islands can be used for reconstruction of the palate and the cheek, although parascapular skin and subcutaneous tissue are often quite bulky and not an ideal donor for reconstruction of the cheek and intraoral lining.

Scapular and Parascapular Flaps. The fasciocutaneous parascapular free flap has been in clinical use for about two decades (Figure 40–3). It was defined within the distribution of the circumflex scapular artery by Saijo in 1978.[82] Teot et al described the use of lateral scapular bone 3 years later with the first three reported cases,[83] and the addition of a skin paddle was presented by Nasif's group a year after that.[84] Subsequently, Coleman and Sultan demonstrated that the angular artery, a branch of the thoracodorsal artery, supplies the most inferior aspect of the lateral scapula.[85] The addition of this vascular

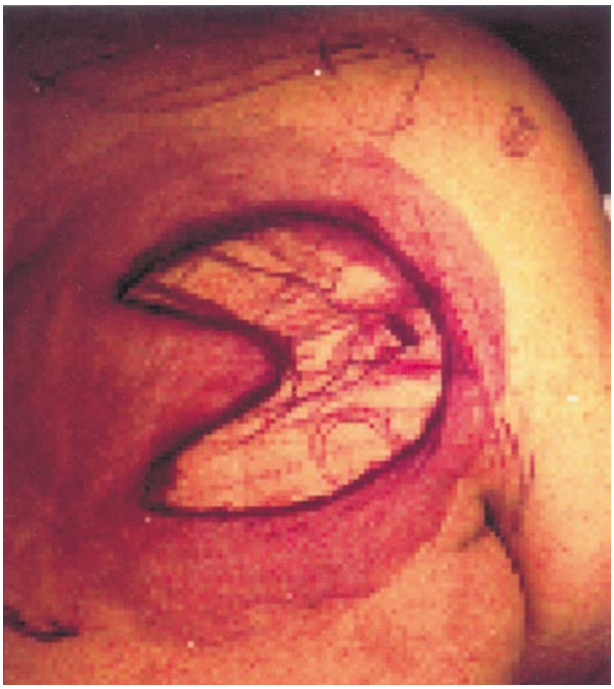


FIGURE 40–3. Parascapular flap donor site.

pedicle allows for two separate bony defects to be reconstructed from the same flap. The lateral border and the inferior tip of the scapula have been used together on their individual pedicles to fill a larger defect than the unaugmented lateral scapular border will span.

Coleman's group strongly advocated the use of the scapular osteocutaneous free flap (SOFF) over fibula and other donor sites, although their series included only four maxillary and midfacial reconstructions among 29 cases.[86] They cite the availability of significantly more skin surface area than can be obtained from fibular or iliac crest osteocutaneous harvests. However, they fail to point out that the latissimus dorsi muscle flap offers far more available skin on the same vascular trunk. Further, scapular bone is often but not always of adequate bulk and dimension to accommodate placement of osseointegrated dental implants.[87] It is useful for orbital floor reconstruction. Of the five patients from Coleman et al's study who remained available for long-term questioning, all complained of difficulty reaching above their heads and lifting objects, and two complained of moderate pain at the donor site.

Modern modifications of the SOFF have significantly increased its utility. However, the morbidity reported by Coleman et al, the unpredictability of bony bulk, and the availability of more bone and larger skin paddles from fibular harvest than that reported by others from scapula prevent the SOFF from being the flap of first choice. The interesting report of prefabricated scapular flaps from Holle's group adds new potential utility to the SOFF.[88] They placed osseointegrated dental implants into the dissected scapular segment and covered it with split-thickness skin grafts around which an expanded polytetrafluoroethylene sheet was placed. The flap was then left in situ for 3 months before harvesting and inserting it. Further improving the outcome was their ability to define preoperatively the anatomy and bulk of the scapula using a digital graphic workstation processing of high-resolution three-dimensional computed tomographic images. With refinement of techniques like that used by Holle's group, the SOFF is becoming a treatment of choice for maxillary reconstruction.

Management of Large Defects following Maxillectomy Large wounds, especially if they involve loss of more than one cutaneous and/or mucosal surface, usually require a rectus abdominis or a latissimus dorsi myocutaneous flap. These flaps can provide skull base coverage as well. Although the radial forearm flap satisfies midfacial soft tissue requirements, the fibular osteocutaneous free flap, now well established for mandibular reconstruction, is considered by many to be the flap of choice for maxillary defects of the hemipalate or greater since these defects require bony replacement to allow for osseointegrated implant use.

Rectus Abdominis Flaps. A large flap of rectus abdominis muscle can be harvested readily on a long pedicle of good caliber. Although usually based on the deep (inferior) epigastric artery, it can even be raised on its superior vasculature if the inferior pedicle is unusable or inadvertently damaged.[89] The rectus abdominis has several advantages over other soft tissue flaps. It can be harvested easily during the resection without repositioning the patient. The pedicle may be as long as 20 cm and is reliably 15 cm or more. A rectus abdominis myocutaneous free flap can be tailored into multiple skin paddles and/or myofascial components to reconstruct cheek, oral lining, and orbital lining. The muscle in this flap is thought by some authors to satisfactorily isolate a free nonvascularized bone graft placed within it to protect it from oral contamination.[74]

By using the rectus abdominis free flap in combination with nonvascularized bone grafts, a three-dimensional defect is reconstructed quite easily because bone, skin, and soft tissue may be inset in their proper positions without compromising the microvascular aspect of the reconstruction. However, most recent literature favors the use of a single donor source over multiple sites and discontinuous tissues.[23]

Latissimus Dorsi Flaps. Vascularized free latissimus dorsi flaps have been used for closure of extensive maxillectomy defects for almost two decades.[59] With so much collective experience, the facial plastic and reconstructive surgeon can rely on flaps supplied by the subscapular artery for reconstruction of moderate to very large maxillectomy defects. The spectrum of flaps supplied by the subscapular system is large,[90] including multicomponent "megaflaps" using various combinations of latissimus dorsi, scapula, parascapular skin, and serratus anterior muscle to reconstruct complex defects involving skin, bone, and mucosa (Figure 40–4). Ample literature exists supporting the utility and success rates of so-called "megaflaps" using skin, muscle, and/or bone supplied by branches of the subscapular-circumflex scapular-thoracodorsal system.[91]

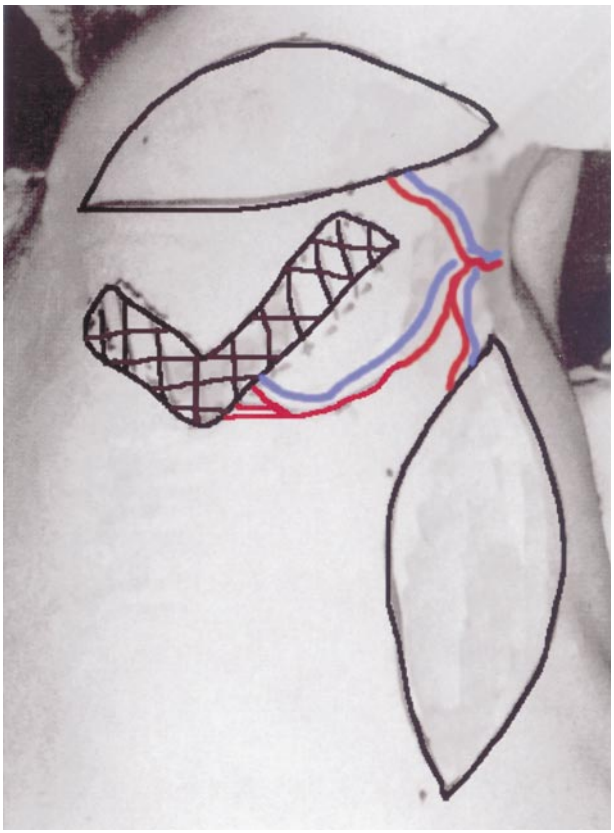


FIGURE 40–4. Megaflap donor site.

Radial Forearm Flaps for Maxillectomy Reconstruction

Microvascular free flaps based on the radial artery have been in clinical use for many years[92] and have achieved wide application in the management of maxillectomy defects.[93] Flaps can be tailored to include various combinations of skin, fat, fascia, tendon, and bone. The limiting factor in harvesting of osseous flaps in the radial artery distribution is the fragility of the remaining radius.

Complications at the donor site occur in about 15% of radial forearm flap harvests. In a large and well-studied series, 100 patients were followed for complications after radial forearm flap transfers.[94] Eighty-six patients were available for follow-up at 3 months and 74 at 1 year. Partial loss of the donor-site skin graft occurred in 14 patients (16%), with exposure of tendons in 11 patients (13%). Delay in healing of the split-thickness skin graft at the donor site occurred in 19 patients (22%). There were six fractures of the remaining radius in 35 patients having osteocutaneous flap harvests (17%). Superficial radial nerve sensation was reduced in 24 patients (32%), 10 reported cold intolerance, and 21 (28%) were unhappy with the appearance of the donor site. Function of the donor arm was restricted in 8 patients (16%) in the fasciocutaneous group, 7 patients (36%) in the composite group without fracture, and in all patients who sustained a postoperative donor site fracture. Detailed measurements of forearm circumference, grip strength, pinch strength, and wrist movements showed greater reduction in these parameters in patients reporting restricted function compared with those reporting normal function, confirming their subjective impressions. Fracture of the radius produced functional impairment in all cases.

Free vascularized radial osteocutaneous flaps have been used to reconstruct maxillectomy defects involving the infraorbital rim and midfacial skin,[95] as well as less extensive tissue loss. Preference for the radial osteocutaneous flap is expressed by sev-

eral authors for defects involving the orbital floor and inferior rim. Vascularized scapular blade is the alternative osteocutaneous flap of choice by some authors.[96]

MANDIBULAR RECONSTRUCTION

Mandibular Anatomy

The mandible is unique among the bones of the body. As part of the stomatognathic system, the mandible is suspended in space by muscles. It is free to move within an envelope of motion defined by its ligamentous attachments to other bones and by interdental contacts in patients with teeth. The ligaments of the temporomandibular joint combine with the stylomandibular ligament to place firm limits on mandibular excursion. Muscular attachments also limit mandibular movement, but these limitations are less well defined because of the distensibility and length tension phenomenon of muscle.

The muscles of mastication control mandibular movement, inserting on a significant portion of the overall surface of the mandible and exerting strong force on it during chewing. Suprahyoid muscles connect the mandible to the hyoid. The genioglossus and mylohyoid form the floor of the mouth, further defining mandibular orientation and movement.

The anatomic regions of the mandible include the condyle and condylar neck, sigmoid notch, coronoid process, ramus, angle, body, alveolar process, parasymphysis, and symphysis. Within the alveolar process are the teeth and supporting structures, including the periodontal ligament and the neurovascular bundles. The inferior alveolar artery, vein, and nerve course through the inferior alveolar canal within the body of the mandible to supply the teeth. The periodontium derives blood supply both from periapical vessels and from perforating vessels in the alveolar bone. The gingival blood supply is extraperiosteal and originates in the vessels of the face and neck.

When natural teeth are lost, the alveolar process resorbs, leaving the reduced height of basal bone of the maxilla and mandible. The relative bulk and anatomic contours of the jaws must be appreciated if appropriate reconstructive methods and materials are to be chosen.

Treatment Planning

It is necessary to consider the bone defect first when planning reconstruction of composite defects of the mandible and lip. Without adequate restoration of skeletal support in the region of the anterior mandibular arch, soft tissue reconstruction cannot achieve firm apposition of the upper and lower lips. Restoration of bony continuity is recommended for all patients with segmental mandibular defects since the rate of hardware failure and/or exposure is reported to be at least 25% by multiple authors.[97]

The mandible may be invaded by tumors of the oral cavity. In these cases, reconstruction must include the restoration of both the soft tissue and bony anatomy. Depending on the extent of tumor resection, the residual mandible may not have sufficient stock to withstand the forces of mastication. In considering reconstruction, important anatomic variables include the location and length of the defect and any associated soft tissue, skin, or neurovascular involvement.[98]

Although it is not clear that all mandibular defects require reconstruction, there is obvious benefit in many instances. For instance, resection of the anterior mandibular arch produces the "Andy Gump" deformity (Figure 40–5), which is a complete loss of anterior oromuscular support and oral competence. Because this is such a debilitating functional and esthetic problem, it is important to reconstruct this defect at the time of resection.[99] The use of an immediate microvascular bone-containing free flap is the optimal method to restore form and function to the lower third of the face and also allows the patient to wear a functional dental prosthesis. Other absolute indications for this type of reconstruction include through-and-through mandibular defects (which require reconstruction not only of bone but also of mucosa and external skin), severe mandibular osteoradionecrosis, and failure of previous attempts at mandibular reconstruction.

A variety of osseocutaneous free flaps have been described to reconstruct the mandible. The fibula is the most advantageous donor site, although the iliac crest may also be used.[100] The free fibula flap provides the greatest length of bone and a minimal donor-site defect. Multiple osteotomies can be made in the bone to properly contour it (Figure 40–6), and dental rehabilitation can be achieved with primary or secondary dental implants. The iliac crest free flap provides

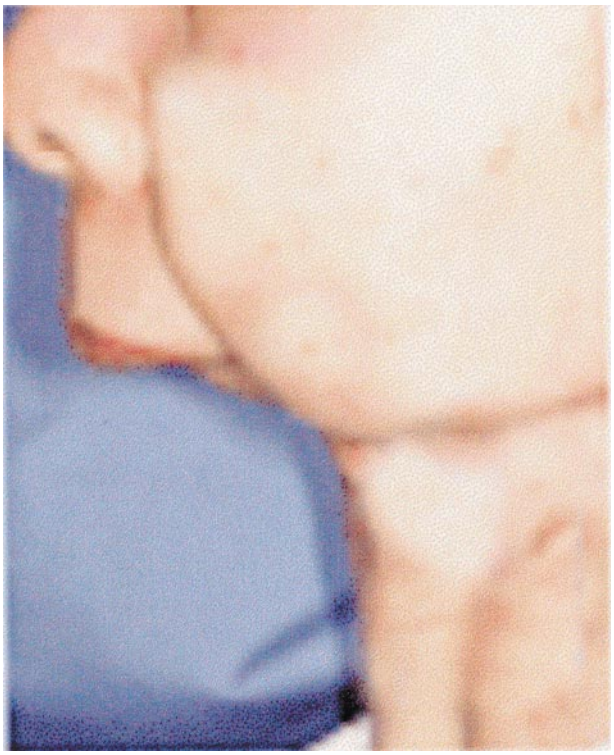


FIGURE 40–5. Unreconstructed mandibulectomy defect ("Andy Gump" deformity).

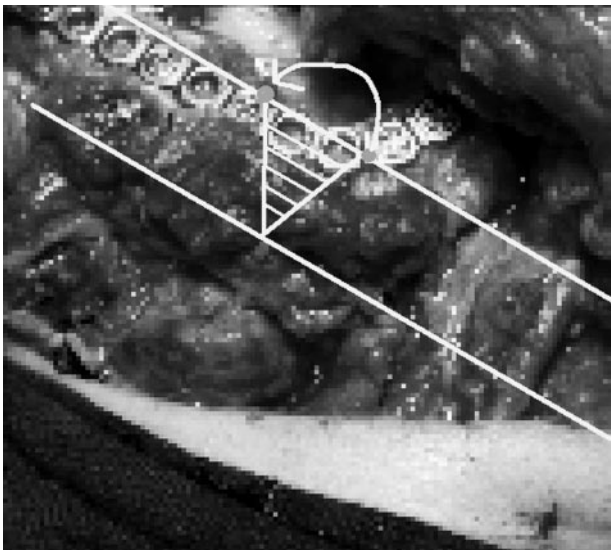


FIGURE 40–6. Fibular free flap with wedge ostectomy to create an angle.

thicker bone than the fibula in most cases, along with an abundance of soft tissue.[101] This combined osseomusculocutaneous flap has found favor with some authors for through-and-through mandibular defects because the muscle reconstructs the mucosal defect and the skin paddle fills the cutaneous defect.

In a comparative study of the suitability of implantable bone for osteointegration of dental implants,[102] the iliac crest was found to offer the best bone stock with the least variability in comparison with the scapula and fibula. Tempering this as a criterion for flap selection is the observation that few patients achieve optimal oral rehabilitation with implant-borne prostheses because of the cost of such management and the relative lack of insurance coverage for it.

Specific Considerations

There is little question that vascularized fibula is the donor site of choice for reconstruction of most mandibular defects.[103] The results of fibular free flap reconstruction of mandibular defects greater than 4 cm regardless of their location along the mandibular arch have been good. Defects anterior to the mental foramen require aggressive bony reconstruction to prevent the collapse of the mandibular arch (Figure 40–7). The fibular free flap provides about 25 cm of bone similar in cross-sectional shape and area to native mandibular body. The graft can be osteotomized without compromising segmental viability because the blood supply is both periosteal and endosteal. The vascular pedicle is long, of good cal-

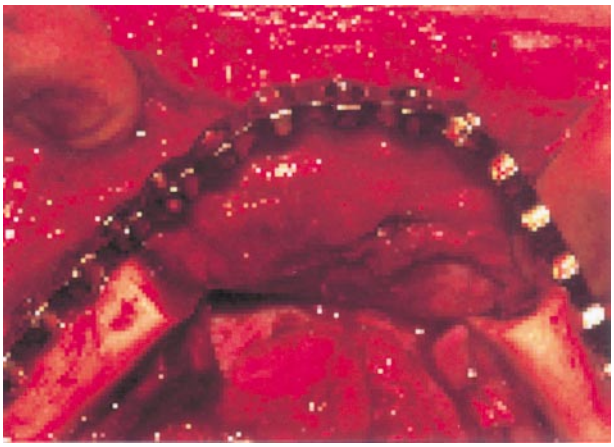


FIGURE 40–7. Fibular free flap reconstruction of an anterior mandibular arch.

iber, and reliable. After initial reports of 30 to 40% failure of the skin paddle of fibular osteocutaneous flaps, published modifications in technique have increased the reliability of the skin paddle to that of the bone flap itself.[104] See Chapter 41 for a more detailed description of the spectrum of available flaps.

Although lateral defects of the mandibular body are better tolerated than are defects of other regions, the mandibular body is also easily reconstructed. The typical defect from unreconstructed lateral mandibular loss includes visible external collapse and deviation of the jaw to the contralateral side (Figure 40–8). This causes a functional impairment that can be improved but not corrected with dental prostheses. Improvements in denture design for mandibulectomy patients continue to be made,[105] but there is good literature support for primary reconstruction of pure lateral mandibular defects.[106] Primary reconstruction of lateral mandibular defects is routinely performed and strongly advocated because the cosmetic and functional residua of lateral resection are significant, even if they are relatively well tolerated. Reconstruction of a lateral mandibular defect with vascularized fibula produces a cosmetic and functional result approximating normalcy (Figure 40–9).

FIGURE 40–8. Unreconstructed lateral mandibular defect.

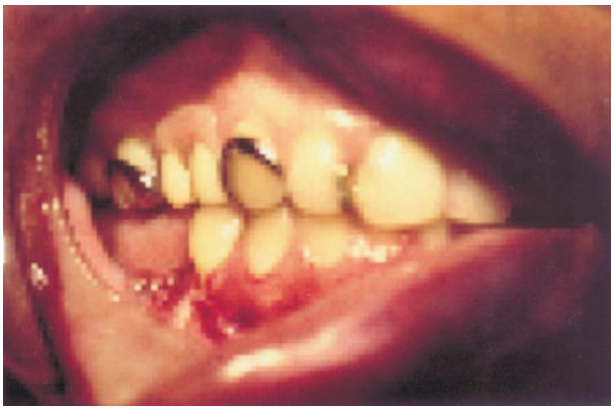


FIGURE 40–9. Normal occlusion after fibular replacement of body, angle, and ramus.

The cutaneous paddle of a fibular osteocutaneous flap can be rendered sensate by anastomosis of the lateral sural cutaneous nerve within the graft to the proximal end of a recipient sensory nerve such as the inferior alveolar, lingual, or great auricular.[107] Published results have been inconsistent, with some authors believing strongly that innervation of flaps and grafts must be restored surgically[108] and others finding return of sensory function in the absence of surgical reinnervation.[109] Further, some authors have suggested that sensation is not even necessary for function of the upper aerodigestive tract and that "Current beliefs about the utility of sensate free tissue transfer and the importance of sensation in swallowing in general may need refinement."[110]

Defects of the Anterior Mandibular Arch

Defects of the anterior mandibular arch cause severe problems that cannot be reversed without formal reconstruction. These problems combine to produce the classic "Andy Gump" deformity, named after the 1917 comic strip character whose appearance suggests loss of the anterior mandibular arch. The loss of anchorage of suprahyoid muscles to the genial tubercles causes collapse of the pharyngeal and hypopharyngeal airway. Detachment of the genioglossus from the mandible destabilizes the tongue and interferes with protrusive movements. Loss of labial support affects competence of the oral sphincter.

When the anterior mandibular arch is removed, there is usually an associated soft tissue

deficit too large to close primarily. A mandibular reconstruction plate can be used to secure the remaining mandibular bodies and restore continuity to the arch. The stabilizing function of the genial musculature can be approximated by suturing the hyoid and tongue base to the plate, but this provides only a crude and limited imitation. Soft tissue of varying character and from a variety of donor sites has been wrapped around such plates to fill the space left by the resected specimen, but this is also a crude management method. And, as cited earlier in this chapter, the rate of loosening and breakage of plates used in this fashion is very high, necessitating secondary reconstructive surgery in at least half of patients so treated over the next 3 years. The effects of shear and moment on security and stability of mandibular reconstruction plates have been demonstrated.[111] As the plate and screws used for reconstruction of the anterior mandibular arch are subject to multiplanar shear, torsion, bending, and loading, the high failure rate of plates not secured to a continuous bony arch is not surprising.

Insetting a vascularized bone graft and securing it firmly between the cut ends of the native mandibular remnants provides stability to the reconstruction. Regardless of the source of the bone graft, the success rate of the reconstruction is greatly enhanced over that of a plate with surrounding soft tissue. Fibula is preferred, as is the use of a skin paddle from the lateral lower leg whenever soft tissue coverage is also required. The inferior border of the fibular graft is aligned with that of the remaining mandible to achieve the best bone apposition as well as the best cosmetic result (Figure 40–10), despite the increased vertical dimension of removable or fixed dental prostheses that may be placed.[112]

In cases requiring replacement of skin of the lower face as well as oral lining, a fibular osteocutaneous flap is used for reconstruction of the mandible and facial skin with a second flap for replacement of oral lining (Figure 40–11). A radial forearm flap is ideal for this purpose. Some authors prefer a scapular-parascapular flap for this problem, resurfacing the face with a skin paddle on latissimus and closing the intraoral defect with the fascia-covered surface of the muscle. One could also accomplish this with two parascapular skin paddles, using the scapula itself for mandibular reconstruction and harvesting the entire flap on the circumflex scapular artery.

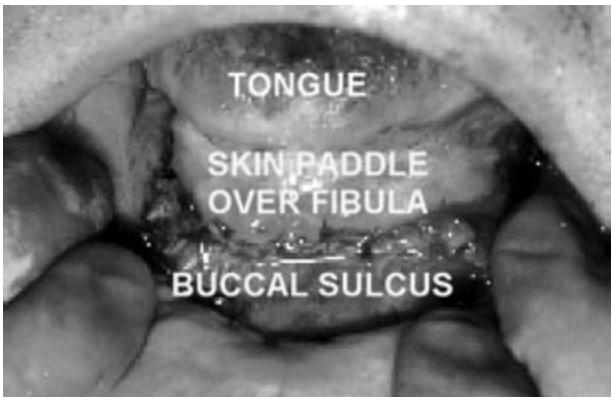


FIGURE 40–10. Insetting of a fibular osteocutaneous flap for anterior reconstruction.

Reconstruction of the Entire Mandible

As the average adult fibula yields at least 25 cm of transferable linear bone, it can be used to fabricate an entire mandible. Osteotomies and wedge ostectomies are routinely used to contour the graft. They do not compromise the viability and healing potential of the graft, even though they have been demonstrated to diminish distal blood flow in the graft.[113] However, care must be taken not to interrupt the periosteum because it carries the blood supply from segment to segment.

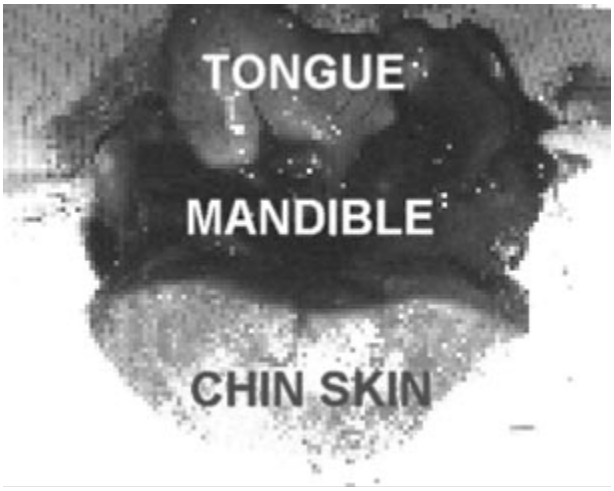


FIGURE 40–11. Through-and-through resection of the tongue, mandible, and chin.

Reconstruction of the Temporomandibular Joint

The spectrum of management methods for loss of the mandibular condyle have ranged from nonintervention, to extension of the smoothed end of a bone graft or flap into the glenoid fossa, to implantation of alloplastic joint prostheses. Explored along the way were methods as varied as securing the resected native condyle to a graft and reinserting it into the fossa, filling the fossa with temporalis muscle and/or fascia, and transplanting the head of the fifth metatarsal. Current work favors the placement of a smoothed fibular graft into the joint space, securing it with a 2-0 braided permanent suture through the articular disk (if it remains) or through an anchoring hole in the bony rim of the fossa (Figure 40–13).[114]

Maintenance of the articular disk is desirable, whenever possible, for improved stabilization of the neocondyle in the glenoid fossa. If the patient has adequate articulating upper and lower posterior teeth to define the vertical dimension of the occlusion, the exact method of managing temporomandibular joint loss is probably not so critical. However, many head and neck cancer patients are fully edentulous at the time of treatment and do not undergo oral rehabilitation. These patients will likely have sufficient difficulty with temporomandibular joint function to prevent them from chewing solid food in the absence of a functional temporomandibular joint reconstruction. Great care must be paid to shaping and placing a neocondyle during reconstruction of mandibular defects in these patients.

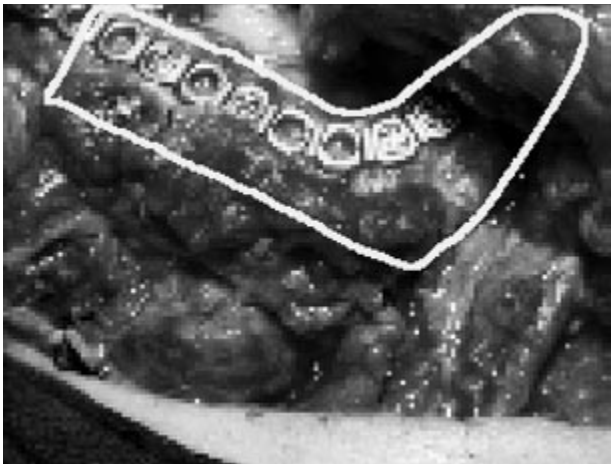


FIGURE 40–12. Shaped fibular end on temporalis muscle as a neocondyle.

Sensory Nerve Considerations in Mandibular Reconstruction

A word must be said about restoration of sensation during mandibular reconstruction. The inferior alveolar nerve will be sacrificed in all mandibular resections posterior to the mental foramen, rendering the ipsilateral lower lip numb. Grafting has been proposed at the time of resection and reconstruction.[115] This is a reasonable component of management, and the facial plastic and reconstructive surgeon should consider this during mandibular reconstruction. Grafting from the contralateral mental nerve has also been reported.[116]

Extensive and Unusual Defects

Extensive tumors of the lateral pharyngeal wall and retromolar trigone can necessitate resection of the mandible and maxilla (Figure 40–13). If the maxillary defect is confined to the lateral palate, alveolar ridge, and pterygoid hamulus, reconstruction of both the mandible and the maxilla can be accomplished with a single fibular osteocutaneous flap (Figure 40–14). Extensive defects of the maxilla and midface that require more skin than can be harvested with a fibular graft may necessitate a megaflap with scapular bone, parascapular skin, and latissimus dorsi muscle. Alternatives include the use of multiple free flaps, although this requires adequate donor vasculature to supply each flap.

Another innovation in reconstruction of the jaws and oral lining is the "bridging flap."[117] Because the peroneal artery and veni comitantes traverse the entire length of a fibular flap, the distal ends of the vessels can nourish a second flap. Thus, a radial forearm flap can be used to provide pliable oral lining over a fibular reconstruction of the mandible. This is especially useful when the oral defect is through and through the facial skin, requiring external as well as intraoral resurfacing (Figure 40–15).

FUTURE OF JAW RECONSTRUCTION

As can be seen, there are many reconstructive options for defects of the jaws. Distraction osteogenesis and physiologic osseous morphogens offer glimpses into

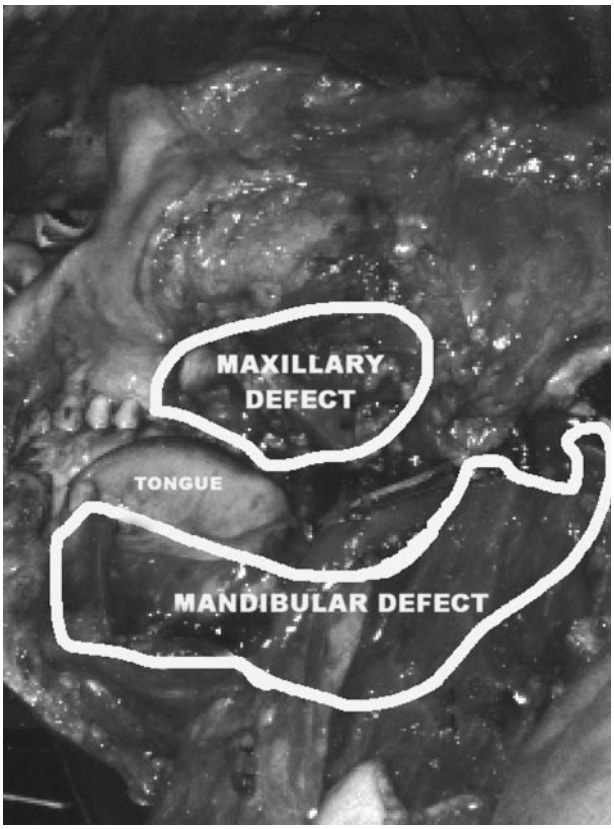


FIGURE 40–13. Composite defect including the maxilla and mandible.

the future of jaw reconstruction. Genetic engineering and flap prefabrication herald the day when tumors will be replaced with well-fitted autografts identical in composition and function to tissues sacrificed for tumor control. As we endure the wait for

FIGURE 40–14. Reconstruction of the defect in Figure 40–13 with a fibular osteocutaneous flap.

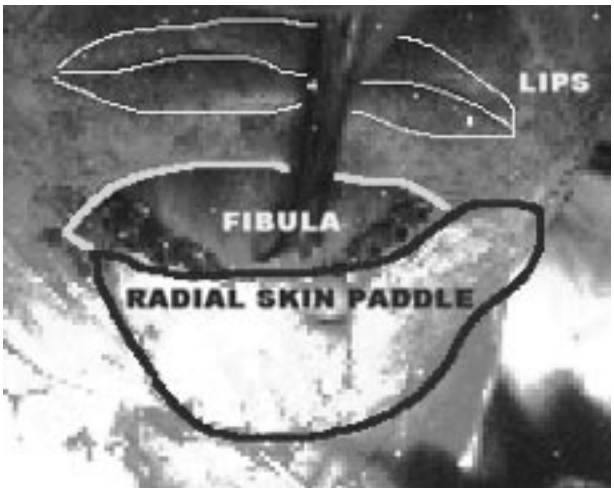


FIGURE 40–15. Reconstruction of the through-and-through defect from resection in Figure 40–11.

"molecular reconstructive 'cocktails' from the shelf of the next millennium," we have sufficient alternatives for reconstruction of the jaws to produce results that were unimaginable to surgeons only 25 years ago.

REFERENCES

1. Kroll SS. An overview of head and neck reconstruction. In: Knoll SS, editor. Reconstructive plastic surgery for cancer. St. Louis: Mosby; 1996.
2. O'Leary MJ, Martin PJ, Hayden RE. The neurocutaneous free fibula flap in mandibular reconstruction. Otolaryngol Clin North Am 1994;27:1081–96.
3. Hobkirk JA, Brouziotou-Davas E. The influence of occlusal scheme on masticatory forces using implant stabilized bridges. J Oral Rehabil 1996;23:386–91.
4. Webster's new universal unabridged dictionary. 2nd ed. New York: Simon and Schuster; 1979.
5. Putti V. The operative lengthening of the femur. JAMA 1921;127:934.
6. Ilizarov GA. Transosseous osteosynthesis: theoretical and clinical aspects of the regeneration and growth of tissue. Berlin: Springer-Verlag; 1992.
7. Rubio-Bueno P, Padron A, Villa E, Diaz-Gonzalez FJ. Distraction osteogenesis of the ascending ramus for mandibular hypoplasia using extraoral and intraoral devices: a report of 8 cases. J Oral Maxillofac Surg 2000;58:593–9.
8. Tehranchi A, Behnia H. Treatment of mandibular asymmetry by distraction osteogenesis and orthodontics: a report of four cases. Angle Orthod 2000;70:165–74.

9. Judge B, Hamlar D, Rimell FL. Mandibular distraction osteogenesis in a neonate. Arch Otolaryngol Head Neck Surg 1999;125:1029–32.
10. Costantino PD, Friedman CD, Shindo ML, et al. Experimental mandibular regrowth by distraction osteogenesis. Arch Otolaryngol Head Neck Surg 1993;119:511–6.
11. Cho BC, Lee JH, Baik BS, et al. Distraction osteogenesis of free interpositional membranous bone: experimental design. J Craniofac Surg 1999;10:123–7.
12. Oda T, Sawaki Y, Fukuta K, Ueda M. Segmental mandibular reconstruction by distraction osteogenesis under skin flaps. Int J Oral Maxillofac Surg 1998;27:9–13.
13. Rachmiel A, Aizenbud D, Eleftheriou S, et al. Extraoral vs. intraoral distraction osteogenesis in the treatment of hemifacial microsomia. Ann Plast Surg 2000;45:386–94.
14. Wangerin K. [Untitled abstract]. Mund Kiefer Gesichtschir 2000;4 Suppl 1:S226–36.
15. Helm GA, Alden TD, Sheehan JP, Kallmes D. Bone morphogenetic proteins and bone morphogenetic protein gene therapy in neurological surgery: a review. Neurosurgery 2000;46:1213–22.
16. Barboza EP, Duarte ME, Geolas L, et al. Ridge augmentation following implantation of recombinant human bone morphogenetic protein-2 in the dog. J Periodontol 2000;71:488–96.
17. Yoshida K, Bessho K, Fujimura K, et al. Enhancement by recombinant human bone morphogenetic protein-2 of bone formation by means of porous hydroxyapatite in mandibular bone defects. J Dent Res 1999;78:1505–10.
18. Ollier L. Traité experimental et clinique de la regeneration des os et de la production artificelle du tissu osseux. Paris: P. Masson et fils; 1867.
19. Robertson IM, Barron JN. A method of treatment of chronic osteitis. J Bone Joint Surg 1946;28:19–24.
20. Spear SL,Wiegering CE. Temporal fossa bone grafts: a new technique in craniofacial surgery. Plast Reconstr Surg 1987;79:531–4.
21. Kline RM, Wolfe SA. Complications associated with the harvesting of cranial bone grafts. Plast Reconstr Surg 1995;95:5–13.
22. Lawson W, Biller HF. Mandibular reconstruction: bone graft techniques. Otolaryngol Head Neck Surg 1982;90:589–94.
23. Foster RD, Anthony JP, Singer MI, et al. Reconstruction of complex midfacial defects. Plast Reconstr Surg 1997;99:1555–65.
24. Lalikos JF, Mooney MP, Janosky J, et al. Bacterial exposure required to induce rabbit calvarial bone graft infection by superficial contamination. J Craniofac Surg 1994;5:247–52.
25. Jisander S, Aspenberg P, Salemark L, Wennerberg J. Mandibular reconstruction by secondary reimplantation of resected segments: a preliminary report. Int J Oral Maxillofac Surg 1995;24:288–92.
26. Holmes RE, Wardrop RW, Wolford LM. Hydroxylapatite as a bone graft substitute in orthognathic surgery: histologic and histometric findings. J Oral Maxillofac Surg 1988;46:661–71.
27. Cottrell DA, Wolford LM. Long-term evaluation of the use of coralline hydroxyapatite in orthognathic surgery. J Oral Maxillofac Surg 1998;56:935–41.
28. Klotch DW, Prein J. Mandibular reconstruction using AO plates. Am J Surg 1987;154:384–8.
29. Blackwell KE, Buchbinder D, Urken ML. Lateral mandibular reconstruction using soft-tissue free flaps and plates. Arch Otolaryngol Head Neck Surg 1996;122:672–8.
30. Ueyama Y, Naitoh R, Yamagata A, Matsumura T. Analysis of reconstruction of mandibular defects using single stainless steel A-O reconstruction plates. J Oral Maxillofac Surg 1996;54:858–62.
31. Klotch DW, Gal TJ, Gal R. Assessment of plate use for mandibular reconstruction: has changing technology made a difference? Otolaryngol Head Neck Surg 1999;121:388–92.
32. Blackwell KE, Lacombe V. The bridging lateral mandibular reconstruction plate revisited. Arch Otolaryngol Head Neck Surg 1999;125:988–93.
33. Komisar A. The functional result of mandibular reconstruction. Laryngoscope 100:364–74.
34. Serletti JM, Higgins JP, Moran S, Orlando GS. Factors affecting outcome in free-tissue transfer in the elderly. Plast Reconstr Surg 2000;106:66–70.
35. Talesnik A, Markowitz B, Calcaterra T, et al. Cost and outcome of osteocutaneous free-tissue transfer versus pedicled soft-tissue reconstruction for composite mandibular defects. Plast Reconstr Surg 1996;97: 1167–78.
36. Gallivan KH, Reiter D. Acute postoperative alcohol withdrawal and free flap manidibular reconstruction. Arch Facial Plast Surg 2001;3:264–6.
37. Hoopes JE, Edgerton MT. Surgical rehabilitation after radical maxillectomy and orbital exenteration. Md State Med J 1966;15:91–7.
38. Cordeiro PG, Santamaria E. A classification system and algorithm for reconstruction of maxillectomy

and midfacial defects. Plast Reconstr Surg 2000; 105:2331–46.
39. Muzaffar A, Adams WP, Hartog JM, et al. Maxillary reconstruction: functional and aesthetic considerations. Plast Reconstr Surg 1999;104:2172–84.
40. Guven O. A clinical study on oroantral fistulae. J Craniomaxillofac Surg 1998;26:267–71.
41. Curtis TA, Beumer J. Restoration of acquired hard palate defects: etiology, disability, and rehabilitation. In: Beumer J, Curtis TA, Marunick MT, editors. Maxillofacial rehabilitation: prosthodontic and surgical considerations. St. Louis: Ishiyaku EuroAmerica; 1996.
42. Hemmings K, Griffiths B, Hobkirk J, Scully C. Clinical review: ABC of oral health. BMJ 2000;321:438–41.
43. Carlsson GE. The geography of prosthodontics. Int J Prosthodont 1997;10:407–10.
44. Appollonio I, Carabellese C, Frattola A, Trabucchi M. Dental status, quality of life, and mortality in an older community population: a multivariate approach. J Am Geriatr Soc 1997;45:1315–23.
45. Gotay CC, Yates JW. The long-term impact of a training program in maxillofacial prosthodontics. J Cancer Educ 1988;3:181–6.
46. Carlsson GE. Clinical morbidity and sequelae of treatment with complete dentures. J Prosthet Dent 1998;79:17–23.
47. Davison SP, Sherris DA, Meland NB. An algorithm for maxillectomy defect reconstruction. Laryngoscope 1998;108:215–9.
48. Salins PC, Kishore SK. Anteriorly based palatal flap for closure of large oroantral fistula. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1996;82:253–6.
49. Goldman EH, Stratigos GT, Arthur AL. Treatment of oroantral fistula by gold foil closure: report of a case. J Oral Surg 1969;27:875–7.
50. Awang MN, Greene MW, King RC, et al. Closure of oroantral fistula. Int J Oral Maxillofac Surg 1988; 17:110–5.
51. el-Hakim IE, el-Fakharany AM. The use of the pedicled buccal fat pad (BFP) and palatal rotating flaps in closure of oroantral communication and palatal defects. J Laryngol Otol 1999;113:834–8.
52. Car M, Juretic M. Treatment of oroantral communications after tooth extraction. Is drainage into the nose necessary or not? Acta Otolaryngol (Stockh) 1998;118:844–6.
53. Edgerton MT, Zovickian A. Reconstruction of major defects of the palate. Plast Reconstr Surg 1956;17: 105–12.
54. Elliott RA. Use of nasolabial skin flap to cover intraoral defects. Plast Reconstr Surg 1976;58:201–4.
55. Campbell HH. Reconstruction of the left maxilla. Plast Reconstr Surg 1948;3:66–70.
56. Konno A, Togawa K, Iizuka K. Primary reconstruction after total or extended total maxillectomy for maxillary cancer. Plast Reconstr Surg 1981;67:440–8.
57. Inoue T, Harashina T, Asanami S, Fujino T. Reconstruction of the hard palate using free iliac bone covered with jejunal flap. Br J Plast Surg 1988;41:143–7.
58. Panje WR, Krause CJ, Bardach J, Baker SR. Reconstruction of intraoral defects with the free groin flap. Arch Otolaryngol 1977;103:78–85.
59. Baker SR. Closure of large orbito-maxillary defects with free latissimus dorsi myocutaneous flaps. Head Neck Surg 1984;6:828–35.
60. Matlou HS, Larson DL, Kuhn JC, et al. Lateral arm free flap in oral cavity reconstruction: a functional evaluation. Head Neck Surg 1989;11:205–10.
61. Swartz WM, Bains JL, Newton ED, et al. The osteocutaneous scapular flap for mandibular and maxillary reconstruction. Plast Reconstr Surg 1986;77: 530–8.
62. Vaughan ED. The radial forearm free flap in orofacial reconstruction: personal experience in 120 consecutive cases. J Craniomaxillofac Surg 1990;18:2–13.
63. Gurtner GC, Evans GR. Advances in head and neck reconstruction. Plast Reconstr Surg 2000;106:672–82.
64. Schusterman MA, Reece GP, Miller MJ. Osseous free flaps for orbit and midface reconstruction. Am J Surg 1993;166:341–5.
65. Pandolfi PJ, Yavuzer R, Jackson IT. Three-layer closure of an oroantral-cutaneous defect. Int J Oral Maxillofac Surg 2000;29:24–6.
66. Chambers RG, Cohn ES. Palatal reconstruction utilizing retrieved forehead flap. J Surg Oncol 1975;7: 191–7.
67. Kavarana NM. Use of a folded forehead flap for reconstruction after a large excision of the full thickness of the cheek. Plast Reconstr Surg 1975;56: 629–32.
68. Dexter WS, Jacob RF. Prosthetic rehabilitation after maxillectomy and temporalis flap reconstruction: a clinical report. J Prosthet Dent 2000;83:283–6.
69. Rose EH, Norris MS. The versatile temporoparietal fascial flap. Adaptability to a variety of composite defects. Plast Reconstr Surg 1990;85:224–2.
70. Goel A, Gahankari D. Extended subgaleal fascia-pericranial temporalis flap for skull base reconstruction. Acta Neurochir (Wein) 1995;135:203–5.

71. Koch A, Hollmann K, Undt G, et al. Reconstruction of the orbital roof after tumor resection with a pedicled muscle calvarial flap—the calvaria temporalis flap. Fortschr Kiefer Gesichtschir 1994;39:50–2.
72. Andres CJ, Newton AD, Schriener JE, Shore JW. Orbital prosthesis following temporal muscle or forehead flap reconstruction: use of optics and illusions. J Prosthet Dent 1992;67:390–3.
73. Fukuta K, Jackson IT, Collares MV, et al. The volume limitation of the galeal temporalis flap in facial augmentation. Br J Plast Surg 1991;44:281–4.
74. Cordeiro PG, Santamaria E, Kraus DH, et al. Reconstruction of total maxillectomy defects with preservation of the orbital contents. Plast Reconstr Surg 1998;102:1874–84.
75. Cheung LK. Microvascular network of the healing surface over the temporalis flap in maxillary reconstruction. Int J Oral Maxillofac Surg 1999;28: 469–74.
76. Pribaz J, Stephens W, Crespo L, Gifford G. A new intraoral flap: facial artery musculomucosal (FAMM) flap. Plast Reconstr Surg 1992;90:421–9.
77. Licameli GR, Dolan R. Buccinator musculomucosal flap: applications in intraoral reconstruction. Arch Otolaryngol Head Neck Surg 1998;124:69–72.
78. Thomas WO. Facial arteriovenous malformation managed with ablative surgery and dual rotational flap reconstruction. South Med J 1994;87:1178–82.
79. Konno A, Togawa K, Iizuka K. Primary reconstruction after total or extended total maxillectomy for maxillary cancer. Plast Reconstr Surg 1981;67:440–8.
80. Taylor GI, Miller GD, Ham FJ. The free vascularized bone graft. A clinical extension of microvascular techniques. Plast Reconstr Surg 1975;55:533–44.
81. Hidalgo DA, Disa JJ, Cordeiro PG, Hu QY. A review of 716 consecutive free flaps for oncologic surgical defects: refinement in donor-site selection and technique. Plast Reconstr Surg 1998;102:722–32; discussion 733–4.
82. Saijo M. The vascular territories of the dorsal trunk: a reappraisal for potential flap donor sites. Br J Plast Surg 1978;31:200–4.
83. Teot L, Bosse JP, Moufarrege R, et al. The scapular crest pedicled bone graft. Int J Microsurg 1981;38:257–62.
84. Nasif TM, Vidal L, Bovet JL, et al. The parascapular flap: a new cutaneous microsurgical free flap. Plast Reconstr Surg 1982;69:591–600.
85. Coleman J, Sultan M. The bipedicled osteocutaneous scapular flap: a new subscapular system free flap. Plast Reconstr Surg 1991;87:682–92.
86. Coleman SC, Burkey BB, Day TA, et al. Increasing use of the scapula osteocutaneous free flap. Laryngoscope 2000;110:1419–24.
87. Frodel JL, Funk GF, Capper DT. Osseointegrated implants: a comparative study of bone thickness in four vascularized bone flaps. Plast Reconstr Surg 1993;92:449–55.
88. Holle J, Vinzenz K, Würinger E, et al. The prefabricated combined scapula flap for bony and soft-tissue reconstruction in maxillofacial defects—a new method. Plast Reconstr Surg 1996;98:542–52.
89. Canales FL, Furnas H, Glafkides M, et al. Microsurgical transfer of the rectus abdominis muscle using the superior epigastric vessels. Ann Plast Surg 1990;24:534–7.
90. Uglesic V, Virag M, Varga S, et al. Reconstruction following radical maxillectomy with flaps supplied by the subscapular artery. J Craniomaxillofac Surg 2000;28:153–60.
91. Yamamoto Y, Nohira K, Minakawa H, et al. The combined flap based on a single vascular source: a clinical experience with 32 cases. Plast Reconstr Surg 1996;97:1385–9.
92. Soutar DS, Scheker LR, Tanner NS, McGregor IA. The radial forearm flap: a versatile method for intraoral reconstruction. Br J Plast Surg 1983;36:1–8.
93. Santamaria E, Granados M, Barrera-Franco JL. Radial forearm free tissue transfer for head and neck reconstruction: versatility and reliability of a single donor site. Microsurgery 2000;20:195–201.
94. Richardson D, Fisher SE, Vaughan ED, Brown JS. Radial forearm flap donor-site complications and morbidity: a prospective study. Plast Reconstr Surg 1997;99:109–15.
95. McLoughlin PM, Gilhooly M, Phillips JG. Reconstruction of the infraorbital margin with a composite microvascular free flap. Br J Oral Maxillofac Surg 1993;31:227–9.
96. Swartz WM, Banis JC, Newton ED, et al. The osteocutaneous scapular flap for mandibular and maxillary reconstruction. Plast Reconstr Surg 1986;77:530–45.
97. Nicholson RE, Schuller DE, Forrest LA, et al. Factors involved in long- and short-term mandibular plate exposure. Arch Otolaryngol Head Neck Surg 1997;123:217–22.
98. Futran ND. Mandibular defect. In: Gates GA, editor. Current therapy in otolaryngology-head and neck surgery. 6th ed. St Louis: Mosby; 1994. p. 327–30.
99. Urken ML, Cheney ML, Sullivan MJ, Biller HF. Atlas of regional and free flaps for head and neck reconstruction. New York: Raven Press; 1995.

100. Hidalgo DA. Fibula free flap: a new method of mandibular reconstruction. Plast Reconstr Surg 1989;84:71–9.
101. Urken ML, Weinberg H, Vickery C, et al. The internal oblique-iliac crest free flap in composite defects of the oral cavity involving bone, skin, and mucosa. Laryngoscope 1991;101:257–70.
102. Moscoso JF, Keller J, Genden E, et al. Vascularized bone flaps in oromandibular reconstruction. A comparative anatomic study of bone stock from various donor sites to assess suitability for enosseous dental implants. Arch Otolaryngol Head Neck Surg 1994;120:36–43.
103. Futran ND, Alsarraf R. Microvascular free-flap reconstruction in the head and neck. JAMA 2000;284:1761–3.
104. Schusterman MA, Reece GP, Miller MJ, Harris S. The osteocutaneous free fibula flap: is the skin paddle reliable? Plast Reconstr Surg 1992;90:787–93.
105. Cheng AC, Wee AG, Morrison D, Maxymiw WG. Hinged mandibular removable complete denture for post-mandibulectomy patients. J Prosthet Dent 1999;82:103–6.
106. Anthony JP, Foster RD, Kaplan MJ, et al. Fibular free flap reconstruction of the "true" lateral mandibular defect. Ann Plast Surg 1997;38:137–46.
107. Rhee JS, Weisz DJ, Hirigoyen MB, et al. Intraoperative mapping of sensate flaps. Electrophysiologic techniques and neurosomal boundaries. Arch Otolaryngol Head Neck Surg 1997;123:823–9.
108. Hayden RE. Free flap transfer for restoration of sensation and lubrication to the reconstructed oral cavity and pharynx. Otolaryngol Clin North Am 1994;27:1185–99.
109. Shindo ML, Sinha UK, Rice DH. Sensory recovery in noninnervated free flaps for head and neck reconstruction. Laryngoscope 1995;105:1290–3.
110. Bastian RW, Riggs LC. Role of sensation in swallowing function. Laryngoscope 1999;109:1974–7.
111. Yi Z, Jian-Guo Z, Guang-Yan Y, et al. Reconstruction plates to bridge mandibular defects: a clinical and experimental investigation in biomechanical aspects. Int J Oral Maxillofac Surg 1999;28: 445–50.
112. Inigo F, Rojo P, Ysunza A, Jimenez Y. Three different techniques for mandibular reconstruction after hemimandibulectomy. J Craniofac Surg 1997;8: 58–64.
113. Gur E, Chiodo A, Pang CY, et al. The vascularized pig fibula bone flap model: effects of multiple segmental osteotomies on growth and viability. Plast Reconstr Surg 1999;103:1436–42.
114. Wax MK, Winslow CP, Hansen J, et al. A retrospective analysis of temporomandibular joint reconstruction with free fibula microvascular flap. Laryngoscope 2000;110:977–81.
115. Wessberg GA, Wolford LM, Epker BN. Simultaneous inferior alveolar nerve graft and osseous reconstruction of the mandible. J Oral Maxillofac Surg 1982; 40:384–90.
116. Kaban LB, Upton J. Cross mental nerve graft for restoration of lip sensation after inferior alveolar nerve damage: report of a case. J Oral Maxillofac Surg 1986;44:649–51.
117. Sanger JR, Matloub HS, Yousif NJ. Sequential connection of flaps: a logical approach to customized mandibular reconstruction. Am J Surg 1990;160: 402–4.

CHAPTER 41

Flaps and Grafts in the Head and Neck

Jane Y. Yang, MD, Marc R. Rosen, MD, William M. Keane, MD

The primary role of oncologic surgery is the curative resection of tumor. Equally important, however, is the subsequent reconstruction and rehabilitation, which should lead to an optimal functional and cosmetic result. The type of flap or graft employed in reconstruction is dictated by the location and size of the defect, and every attempt should be made to replace like tissue with like tissue. Simple mucosal defects may be allowed to heal secondarily or may require minor grafting procedures. Larger resections incorporating skin, mucosa, muscle, or bone demand more complex reconstruction, employing free flaps or compound myocutaneous flaps. The ideal reconstructive modality should be performed as a single-stage procedure while providing effective rehabilitation with minimum donor-site morbidity. A wide variety of reconstructive choices are available today. It is important to be familiar with these options while understanding their advantages and limitations.

GRAFTS

Smaller oral cavity and oropharyngeal defects may be allowed to heal by secondary intention. This is appropriate when subsequent wound contracture produces a minimal functional deficit. Defects of the hard palate, where underlying palatal bone has been preserved, are best allowed to heal by secondary intention. Similarly, defects of the tonsillar area, soft palate, buccal mucosa, and tongue heal effectively by secondary intention with minimal contracture. Conversely, only the smallest defects of the anterior floor of the mouth may be allowed to heal by secondary intention as scar contracture typically produces some degree of ankyloglossia.

Graft Physiology

Epithelial grafts are free tissue grafts that depend entirely on the recipient bed for their viability and so require a well-vascularized host bed. Revascularization of skin and mucosal grafts follows three phases: *serum imbibition* (plasmotic circulation), *revascularization*, and *organization*.[1]

Serum imbibition begins when the graft is placed on the recipient site and lasts about 48 hours. A thin layer of fibrin clot forms between the host bed and the skin graft, temporarily anchoring it in place. Nutrients from the plasma transudate, which come from the dilated capillaries of the host bed, diffuse into the graft.[2]

Revascularization of the graft begins with the ingrowth of the vascular buds from the host bed. The process of revascularization, although somewhat controversial, is believed to occur by two processes, *neovascularization* (ingrowth of vessels from the host into the graft) and *inosculation* (direct anastomoses of graft and host vessels).[1]

The *organization* phase begins with the infiltration of leukocytes soon after the graft is placed on the host. As the graft revascularizes, the leukocytes disappear. Fibroblasts begin to infiltrate the fibrin network by day 4, and as the fibroblasts proliferate, the fibrin resorbs. By day 9, the process is complete.

Types of Grafts

The location and size of the defect typically dictate the type of graft to be selected for reconstruction.

Split-Thickness Skin Graft Split-thickness skin grafts (STSGs) consist of the epidermis and a portion of the dermis. The thickness of the graft varies. Thin grafts are harvested at a thickness of 0.005 to 0.012 in, moderate grafts at 0.012 to 0.018 in, and thick grafts at 0.018 to 0.028 in. Thinner grafts are more readily nourished and have a high survival rate but are more liable to contract. Because adnexal structures (hair follicles, sebaceous and sweat glands) are located in the deeper portion of the der-

mis, thicker grafts may grow hair and reflect donor-site morphology (Figure 41–1).[2]

Split-thickness grafts are excellent choices for coverage of areas that require surveillance for tumor recurrence. Split-thickness skin grafts work well in the oral cavity and oropharynx when bulk is not required and when coverage is needed to prevent excessive wound contraction.

They are often meshed, allowing for expansion to cover more surface area. This is particularly appropriate for large superficial defects of the scalp. Powered dermatomes enable these grafts to be harvested with ease and precision. The anterior thigh, with its large surface area, is the favored donor site. Critical to the successful take of the skin graft is immobilization. This can be accomplished in some cases by simply suturing the central portion of the graft, but, in other cases, specifically contoured bolsters may be needed to ensure fixation.

Full-Thickness Skin Grafts Full-thickness skin grafts (FTSGs) contain the epidermis and the entire dermal layer. These grafts are most commonly used to cover defects of the nasal tip, lateral ear, or eyelids after the excision of cutaneous skin malignancies. The color and texture of the donor and recipient site should match as closely as possible. Preauricular, postauricular, and supraclavicular skin is ideal for most facial defects, whereas upper eyelid skin should be used to replace upper and lower lid defects. Because of their thickness, FTSGs do not contract. This quality, however, also leads to slower revascularization and a lower take rate. Unlike the STSG donor site, the FTSG donor site cannot re-epithelialize on its own and thus must be closed primarily or be covered with a STSG.

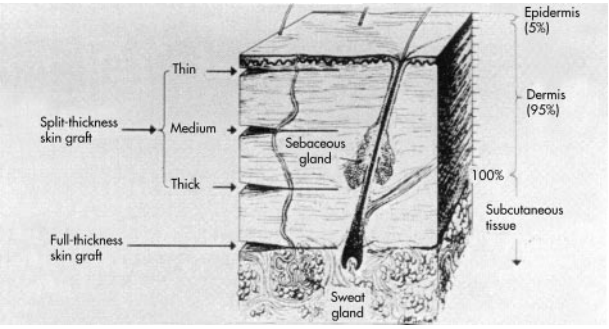


FIGURE 41–1. Diagrammatic representation of skin and subcutaneous tissue. The thickness of various types of skin grafts is illustrated. Reproduced with permission from Petruzzelli GJ and Johnson JT.[2]

Dermal Grafts Dermal grafts are mentioned primarily for historical interest. These grafts are harvested after the epidermal layer has been raised. They have been used in the past for oral and pharyngeal reconstruction and have been placed in the subcutaneous layer to improve facial contouring. These grafts have been used to protect the carotid artery when salivary leaks and skin breakdown might lead to carotid exposure and blowout. This technique has been largely abandoned because of complications of epithelialization of the graft on exposure to air or saliva and the growth of retained epithelium beneath the skin, resulting in cholesteatoma formation.

Mucosal Grafts Buccal mucosa is the usual donor site for mucosal grafts. The graft obtained should not exceed 2 × 3 cm to prevent excessive scarring with subsequent trismus. Buccal grafts are often used for airway reconstruction for which desquamating skin is not desirable.

FLAPS

Classification

A flap is a unit of tissue that may be advanced or rotated into a surgical defect while maintaining its original blood supply. Understanding its vascular supply is crucial to flap design and success. Flaps are defined as "random," "axial," or "free" depending on their pattern of blood supply. Random flaps derive nutrition from local, small, unnamed vessels of the dermal-subdermal plexus and are therefore severely restricted with respect to size. Axial pattern flaps are based on recognizable, usually named arteriovenous pedicles. An axial flap may include a "random" extension at the distal portion of the flap, increasing the amount of skin available. Free flaps are pedicled, axial pattern grafts detached from a distant site (Figure 41–2).

Flaps are classified as local, regional, and distant depending on their location. Tissues that are immediately adjacent to the defect are called local flaps. Local flaps usually have a random vascular pattern. Regional flaps possess an axial vascular pattern and are typically advanced and transferred from a more distal site while maintaining its vascular pedical attachment. Free or distant flaps are pedicled

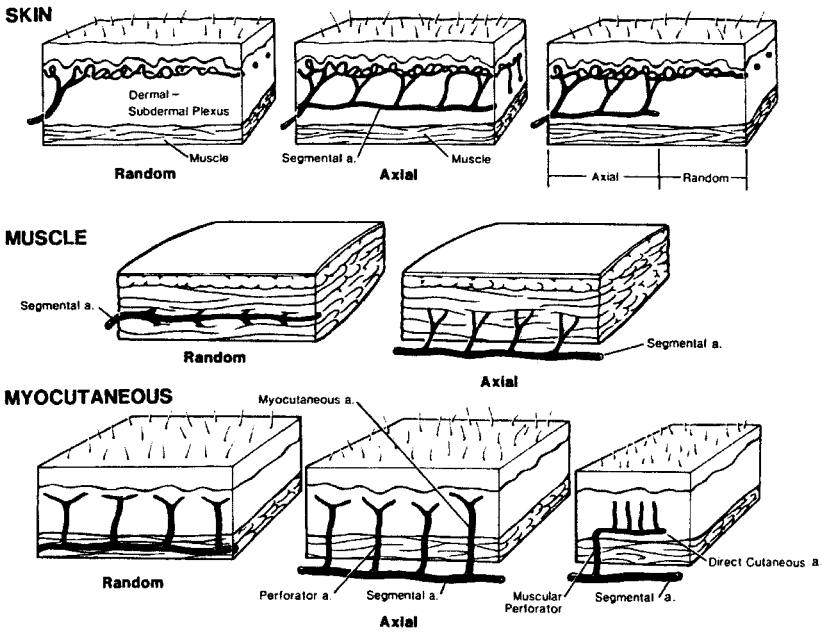


FIGURE 41–2. Diagrams illustrating the blood supply pattern to skin, muscle, and myocutaneous flaps. Reproduced with permission from Paparella MM, Shumrick DA, Gluckman JL, Otolaryngology. Vol IV. 3rd ed. Philadelphia: WB Saunders; 1991.

composite grafts that require separation of the vascular pedicle from the donor site with subsequent reanastomosis within the recipient bed.

Pedicled Regional Flaps

Prior to the development of pedicled regional flaps, head and neck reconstruction options consisted of primary closure, local flaps, or delayed procedures requiring prolonged hospitalization. These options were less than effective for large defects. Primary closure of the oral cavity often reduced the functional volume or tethered the tongue, causing ankyloglossia, dysarthria, or dysphagia. Such techniques also led to stenosis of the oropharynx and hypopharynx, with similar morbidity. The limitations of blood supply with random pattern flaps frequently created a relatively short flap, and attempts to reconstruct more distant sites were often associated with necrosis of the distal part of the flap. In addition, local flap rotation often resulted in stricture and tethering. Delayed flaps required multiple surgical procedures over a prolonged period of time.

Bakamjian introduced the two-stage deltopectoral flap for pharyngoesophageal reconstruction in 1965.[3] Other regional flaps were subsequently introduced, but it was only when Ariyan developed the pectoralis major myocutaneous flap in 1979 that the use of regional pedicled flaps became widespread.[4] Regional flaps dominated the reconstructive scene in the 1980s, with the pectoralis myocutaneous flap taking the lead while becoming the "workhorse" for head and neck reconstruction.

Many of these pedicled flaps can be modified to include a bone or cartilage graft. This modification has been largely replaced with free osseocutaneous flaps, which offer superior reconstructive options with a higher success rate.

A wide variety of pedicled regional flaps are available (Table 41–1) and are described in more detail.

Pectoralis Major Myocutaneous Flap Since the introduction of the pectoralis major myocutaneous flap in 1979, it has become the most commonly used pedicled regional flap for head and neck reconstruction while providing coverage for most defects of the head and neck from the scalp and skull base to the cervical region and hypopharynx.

The pectoralis major muscle is a large fan-shaped structure located over the anterior upper chest. This muscle is attached laterally to the medial surface of the humeral head, medially to the sternum, and inferiorly to the lower ribs. It adducts and medially rotates the humerus. The pectoral branch of the thoracoacromial artery is its primary blood supply, with a minor contribution by a small branch of the lateral thoracic artery. Its vascular pedicle is located in the clavipectoral fascia as it runs on the undersurface of the pectoralis major muscle along

TABLE 41–1. Flaps in Head and Neck Reconstruction

Flap	*Type*	*Vascular Supply*	*Sensory*
Pectoralis major	Pedicled	Thoracoacromial a.	Insensate
Trapezius	Pedicled	Transverse cervical, dorsal scapular, occipital a., perforating branches of the paraspinous vasculature	Insensate
Sternocleidomastoid	Pedicled	Occipital, suprascapular, external carotid/superior thyroid a.	Insensate
Deltopectoral	Pedicled	Internal mammary a.	Insensate
Temporalis	Pedicled	Deep temporal a.	Insensate
Platysma	Pedicled	Facial a.	Insensate
Latissimus dorsi	Pedicled or free	Thoracodorsal a.	Long thoracic n.
Gastric pull-up	Pedicled	Right gastric and gastroepiploic a.	Insensate
Radial forearm	Free	Radial a.	Lateral antibrachial cutaneous n.
Lateral arm	Free	Posterior branch of the radial collateral a.	Inferior lateral brachial cutaneous n.
Lateral thigh	Free	Perforators of the profunda femoris a.	Lateral femoral cutaneous n.
Rectus abdominis	Free	Deep inferior epigastric a.	Insensate
Fibula	Free	Peroneal a.	Lateral sural cutaneous n., sural communicating n.
Iliac crest	Free	Deep circumflex iliac a.	Insensate
Scapula/parascapula	Free	Circumflex scapular a.	Insensate
Jejunum	Free	Superior mesenteric a.	Insensate
Gastro-omental	Free	Right gastroepiploic a.	Insensate

a = artery; n = nerve.

an axis from the shoulder to the xiphoid (Figure 41–3).[5] The pectoralis major muscle receives a minor contribution from the lateral thoracic artery; however, this vessel is typically divided, allowing for a more effective arc of rotation.

This flap has many advantages for head and neck reconstruction. It is reliable, easy to raise, and in close proximity to the head and neck recipient site. It can be raised quickly with the patient in the supine position, and the donor-site defect can usually be closed, primarily when a relatively large skin paddle is harvested.

This flap may be fashioned as a myocutaneous or myofascial flap for the reconstruction of oral cavity or pharyngeal defects.[6] The myofascial flap eliminates several disadvantages of the myocutaneous flap, including excessive bulk, chest wall deformity, and transposition of hair to the aerodigestive tract. The incidence of partial flap loss ranges from 4 to 14% and total flap necrosis from 1 to 7%.[7]

The pectoralis major myocutaneous flap continues to play an important role in head and neck reconstruction, reflecting its versatility, reliability, and ease of harvest.

Trapezius Myocutaneous Flap The trapezius is a large, broad, flat, triangular muscle of the back and shoulder. It originates on the external occipital protuberance near the nuchal line and spinous processes of C7 to T12 and inserts onto the lateral

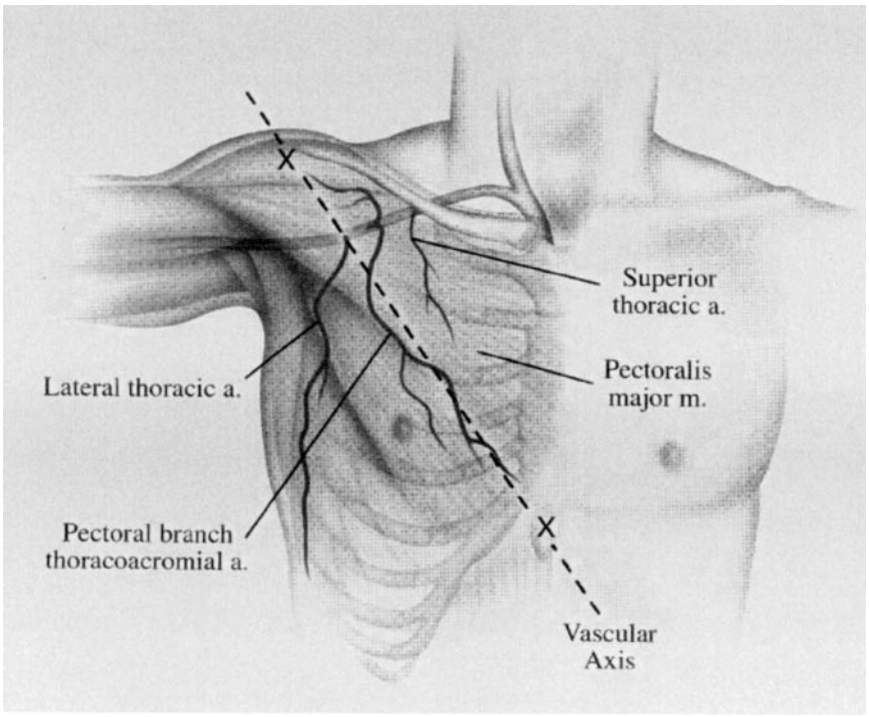


FIGURE 41–3. Illustration of the major blood supply to the pectoralis major myocutaneous flap. Reproduced with permission from Jacono AA, Moscatello AL. Pedicled myocutaneous flaps in head and neck surgery. Oper Tech Otolaryngol Head Neck Surg 2000;11(2).

third of the clavicle, the acromion, and the spine of the scapula. The trapezius acts to steady the scapula. Four arteries, the transverse cervical, dorsal scapular, occipital, and perforating branches of the paraspinous vasculature, supply this broad muscle. Although the transverse cervical is its primary blood supply, collateral circulation allows for the design of three musculocutaneous flaps through different pedicles, the superior flap, island (lateral) flap, and lower (posterior) flap.

The superior flap, first described by Conley in 1972, receives its primary vascular supply from the perforating branches of the posterior intercostal arteries with minor contribution by the occipital branches.[8] This flap requires a simple dissection without the need for identification of the feeding vessels and is useful for closure of ipsilateral regional skin or pharyngeal defects. It is a superiorly based flap and, as such, is not prone to wound separation secondary to gravitational pull (Figure 41–4).

The trapezius muscle is located outside the radiated field and its blood supply is not typically threatened during radical neck dissections.

A major disadvantage is its limited length, preventing the flap from extending across the midline. The most distal portion of the flap, lateral to the acromion, may not survive if the transverse cervical artery has been ligated during a radical neck dissection.[9] In these cases, a delayed procedure may be required to maintain the viability of the distal portion of the flap. The donor site usually requires a STSG.

The lateral island trapezius flap, first presented by Panje[10] and Demergasso[11] in 1978, is based on the transverse cervical vessels (Figure 41–5). This is the least reliable of the three flaps as mobilization is dependent on the course and length of the transverse cervical artery and vein. Both vessels must be patent for this flap to survive, so this flap is not recommended after radical neck dissection in which sacrifice of these vessels often takes place. This flap works well for soft tissue defects of the ipsilateral lateral and anterior neck, oral cavity, and pharynx. The flap itself is technically easy to develop, and the donor-site defect may be closed primarily.

The use of the lower trapezius myocutaneous island flap (LTMIF) was first introduced by Baek et al in 1980.[12] This flap provides the greatest length and arc of rotation and therefore is the most versatile. It is primarily supplied by the transverse cervical vessels, but the dorsal scapular artery feeds the distal portion of the flap. Both arteries usually arise from the thyrocervical trunk, with a great deal of variability in their origin and course. The LTMIF is often based solely on the transverse cervical vessels as the preservation of the dorsal scapular vascular

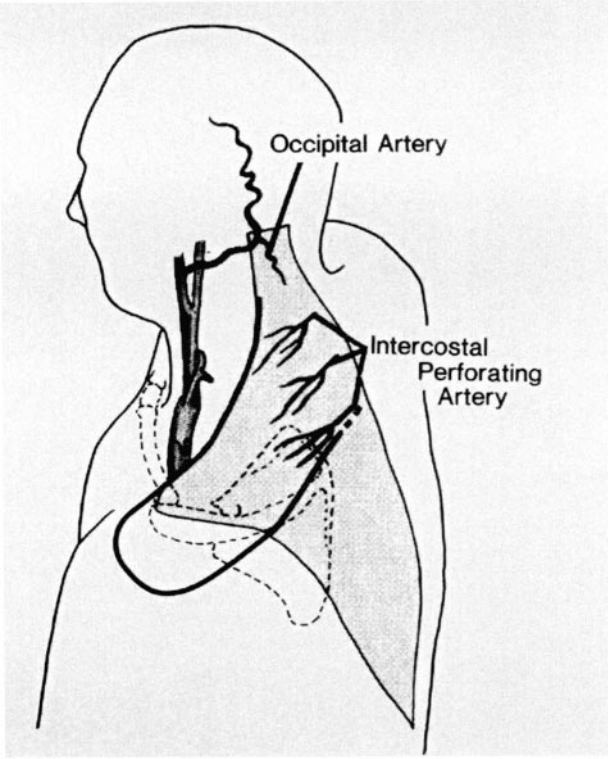


FIGURE 41–4. The superior trapezius myocutaneous flap received its primary blood supply from perforating branches of paraspinous arteries, with minor supplement from the occipital artery. Reproduced with permission from Panje WR, Schuller DE, Shagets FW. Musculocutaneous flap reconstruction of the head and neck. Philadelphia: Lippincott Williams and Wilkins; 1989.

bundle limits the flap's arc of rotation. This may place the distal portion of the flap at risk for partial necrosis, especially if the dorsal scapular artery is a dominant feeding vessel. The skin paddle is designed over the portion of the trapezius muscle bound by the vertebrae and the medial border of the scapula[13] (Figure 41–6).

The LTMIF provides effective coverage for lateral skull base and midface defects. It is a thin, pliable flap with an excellent arc of rotation. The function of the upper portion of the trapezius muscle is often preserved, and the donor defect may be closed primarily.

A significant disadvantage is the need to place the patient in a lateral decubitus position to harvest the flap. A history of prior radiation therapy alone did not affect the rate of flap necrosis, but a higher rate of flap necrosis in patients with a prior history of neck dissection has been noted.[14] Previous radical neck dissection is therefore considered a contraindication to its use.

Sternocleidomastoid Myocutaneous Flap Jianu first reported the sternocleidomastoid myocutaneous flap for facial reanimation in 1908.[15] This muscle originates with two heads and ascends obliquely to insert onto the lateral surface of the mastoid process and into the lateral half of the superior nuchal line. The sternocleidomastoid muscle receives its blood supply from the occipital, superior thyroid, external carotid, and suprascapular arteries.[16] The flap is harvested as a superiorly or inferiorly based flap and may incorporate a portion of the clavicle when used for mandibular reconstruction.

Since its initial description, it has not been overwhelmingly embraced. It is the least reliable of all pedicled flaps.[17] Its random pattern blood supply with the associated tenuous perfusion to its cutaneous portion results in a higher rate of failure.

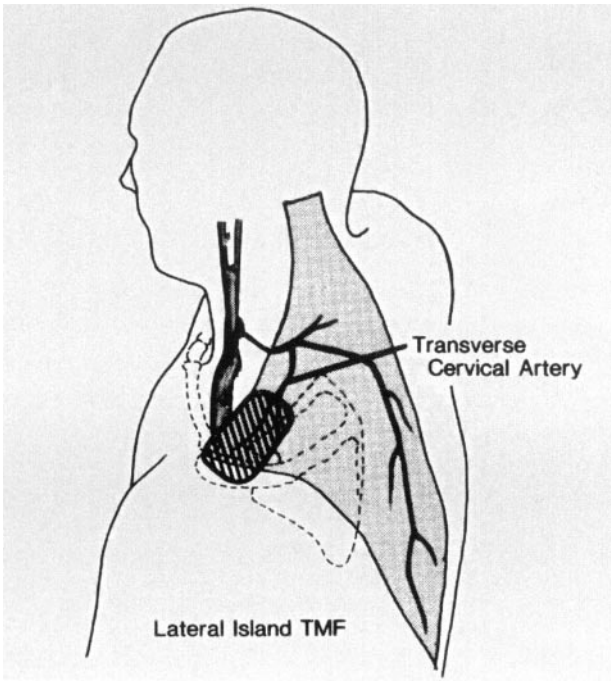


FIGURE 41–5. Transverse cervical artery through its superficial branch provides the primary blood supply to the island trapezius musculocutaneous flap. Reproduced with permission from Panje WR, Schuller DE, Shagets FW. Musculocutaneous flap reconstruction of the head and neck. Philadelphia: Lippincott Williams and Wilkins; 1989.

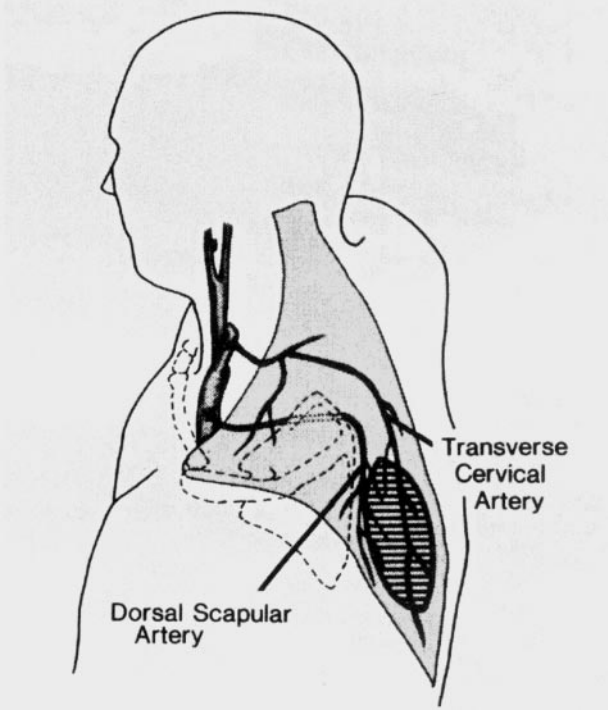


FIGURE 41–6. The lower trapezius myocutaneous island flap receives its primary blood supply from the descending branch of the transverse cervical artery with the dorsal scapular artery contributing to the distal portion of the flap. Adapted from Panje WR, Schuller DE, Shagets FW. Musculocutaneous flap reconstruction of the head and neck. Philadelphia: Lippincott Williams and Wilkins; 1989.

Sebastian et al reported an incidence of total flap loss of 7.5% and superficial skin loss of 22.7% in his large series.[18] This flap is not recommended in previously irradiated necks, and oncologic principles should not be compromised while preserving the sternocleidomastoid muscle for which nodal disease dictates a more radical procedure.[18] Therefore, the sternocleidomastoid myocutaneous flap has limited use but may be applicable for small oral and laryngopharyngeal defects.

Deltopectoral Flap Bakamjian popularized the use of the deltopectoral flap in 1965, and it remained the flap of choice for head and neck reconstruction until pedicled myocutaneous flaps and free flaps were introduced.[3] This thin flap is easy to harvest and leaves no functional deficit.

It is a fasciocutaneous flap based on the perforating branches of the internal mammary arteries. The vessels of the flap run in a plane superficial to the pectoralis major and deltopectoral muscle, requiring the flap to be elevated in a subfascial plane. Limiting elevation to within 2 cm of the sternal border prevents disruption of the feeding vessels. The deltoid branch of the thoracoacromial artery and the anterior circumflex humeral artery supplies the distal portion of the flap overlying the deltoid muscle.

The risk of flap necrosis increases as the flap is extended over the deltoid region, often requiring a delayed procedure. In some series, the rate of flap necrosis reached 25%.[19] To minimize this risk, Kingdom and Singer modified this dissection to include the first four perforators while limiting the medial dissection to the nipple line.[19] In addition, a retention suture is placed at the pivot point, counteracting the gravitational pull on the flap.[20] With these modifications, no major flap loss was reported, and only minor flap necrosis was noted in less than 20% of their cases.[19] Previous radiation or radical neck surgery does not impact the survival of the flap.[3,19] The donor-site defect requires coverage with a skin graft.

Pedicled myocutaneous flaps and free flaps are currently the method of choice over the deltopectoral flap for the reconstruction of oral cavity and pharyngeal defects; the deltopectoral flap continues to be an excellent method for cervical skin and soft tissue defects and remains a viable option for mucosal defects (Figure 41–7).

Temporalis Muscle Flap The temporalis muscle flap was first described by Verneuil in 1872 for interposition arthroplasty in temporomandibular joint ankylosis.[21] Since then, this flap has been used for augmentation of regional tissue defects, for the reconstruction of palatal defects, and to provide reanimation of unilateral facial paralysis.[22]

This fan-shaped muscle arises from the inferior temporal line, fills the entire temporal fossa, and inserts onto the coronoid process of the mandible. The deep temporal artery provides its vascular supply. The anterior third of the muscle is thicker, whereas the middle and posterior thirds are longer and thinner. These qualities make the central portion of the flap useful for facial reanimation. The muscle and periosteum are harvested down to the zygomatic arch and can be detached from the arch and coronoid process to increase the arc of rotation. The inferior rotation of the flap over the zygomatic arch will distort the facial contour. This problem is

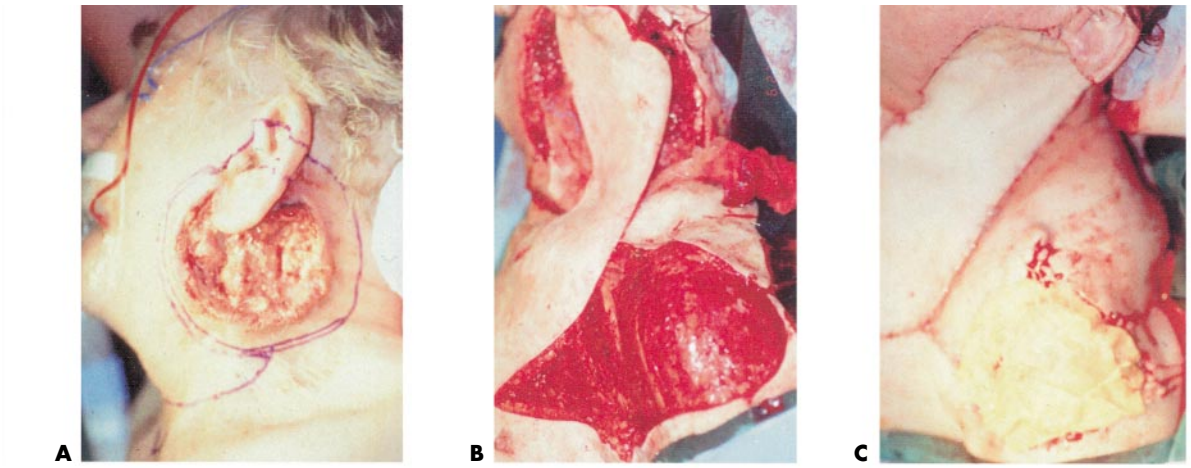


FIGURE 41–7. *A*, Squamous cell carcinoma of the parotid with skin exposure. *B*, Deltopectoral flap raised and rotated into the defect. *C*, The deltopectoral flap is inset into the cervical defect. Xeroform bolster placed over the split-thickness skin graft.

solved by detaching the zygomatic arch prior to rotation of the flap and then reattaching the arch with reconstruction plates.[23]

A major complication of the temporalis muscle flap is injury to the frontal branch of the facial nerve. Clauser et al noted an incidence of 19.2% transient paresis and 2.7% permanent paralysis.[22] Fibrosis of the retromolar trigone region has been experienced in oropharyngeal reconstruction and is best managed with postoperative jaw opening exercises.[22]

This flap is reliable but has a limited role beyond facial reanimation. Its primary use is to cover small to moderate-sized defects of the orbit, lateral aspect of the face, and lateral part of the skull base.

Temporoparietal Fascial Flap The temporoparietal fascia can be used as a pedicled or free flap. This thin, pliable flap may resurface a significant area without bulk.

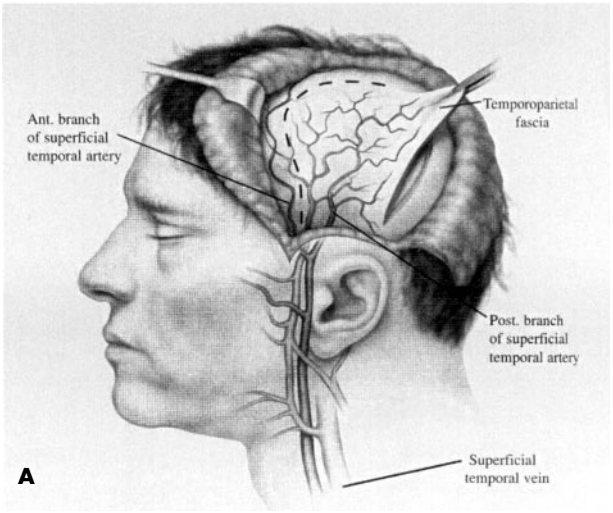
This fascia is the superficial covering that overlies the temporalis muscle and continues on as the superficial aponeurotic system inferiorly and the galea aponeurosis superiorly. Its vascular supply arises from the superficial temporal artery and vein and may be isolated 3 cm superior to the root of the helix prior to its division into an anterior and a posterior branch. The anterior branch, which extends into the frontal region, is often ligated 3 to 4 cm from its takeoff. Further peripheral dissection risks injury to the frontal branch of the facial nerve. The posterior branch, which extends into the parietal region, acts as the distal vascular pedicle (Figure 41–8). Dissection of the vascular pedicle is best monitored with intraoperative Doppler ultrasonography. The auriculotemporal nerve runs with the vascular pedicle and is usually sacrificed, leading to temporal numbness.

The donor site is well hidden within the hairline, creating minimal morbidity. Alopecia is occasionally produced.

This well-vascularized fascial flap is useful for resurfacing the orbit, midface, and lateral oral defects and is available for auricular reconstruction. The flap may include the overlying hair-bearing skin to reconstruct a brow and mustache.[24] The inclusion of a split calvarial bone graft with overlying galea aponeurosis creates a composite flap for the reconstruction of the orbit or zygoma.

Platysma Myocutaneous Flap The platysma myocutaneous flap has several advantages for use in oral cavity reconstruction.[25] It is located within the surgical field, easy to elevate, and pliable and has minimal donor-site morbidity.

This flap is created from the broad sheet of muscle that arises from the fascia covering the upper parts of the deltoid and pectoralis major muscles. It ascends across the clavicle and up the neck, reaching the lower border of the mandible and lateral half of the lower lip. The platysmal myocutaneous flap



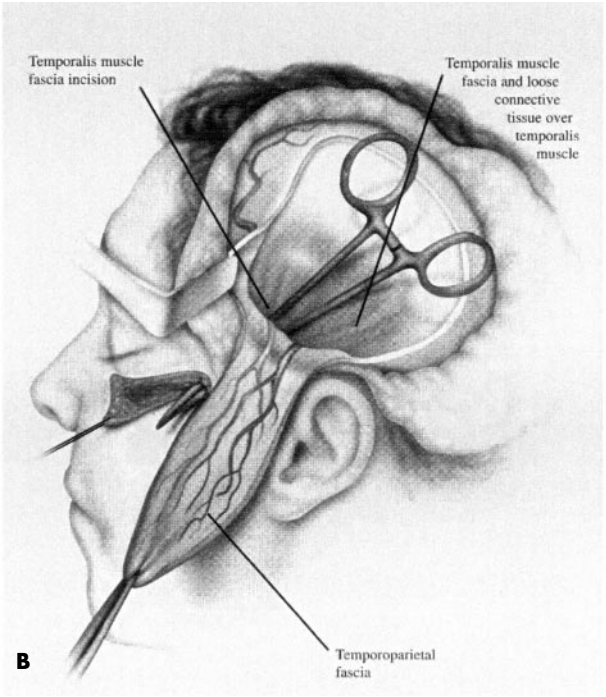


FIGURE 41–8. *A*, The posterior branch of the superficial temporal artery is usually the axial artery to the temporoparietal fascial flap. The superficial temporal vein usually accompanies it but can vary in relative position to the artery or even be completely absent. The incision sometimes requires a posterior T-extension of the vertical opening to maximize exposure. *B*, The temporoparietal flap has excellent range, as demonstrated here. The hemostat shows the safe tunnel made under the deep temporalis fascia over the muscle from the donor site to the subciliary incision. This maneuver allows protection of the facial nerve. Reproduced with permission from Friedman M. Operative techniques in otolaryngology-head and neck surgery—flaps for head and neck reconstruction. Part 1. Philadelphia: WB Saunders; 2000.

receives its blood supply from the submental branches of the facial artery. The platysmal myocutaneous flap can be used even if the ipsilateral facial artery is ligated as long as the submental branch distal to the submandibular gland is patent.[25,26]

This flap is not recommended in necks that have been irradiated or have undergone previous neck dissections.[26,27] A relative contraindication is the presence of ipsilateral facial paralysis because of atrophy of the muscle.[28] A high incidence of complication in patients who received neoadjuvant chemotherapy was noted by Verschuur et al.[25]

Indications for platysma myocutaneous flaps are in the reconstruction of a relatively small floor of mouth, gingival, and oropharyngeal defects not requiring bulk.

Latissimus Dorsi Myocutaneous Flap Quillen et al first described the latissimus dorsi myocutaneous pedicled flap for head and neck reconstruction in 1978.[29] This thin flap, with its large surface area, is extremely versatile for several reasons. Its size provides coverage of massive defects of the orbitomaxillary and lateral skull base region in a single-stage procedure without excessive bulk. The length of the pedicle enables it to reach nearly any area of the head and neck, and the pedicle may also be transferred as a free flap. Dynamic, sensate reconstruction after total glossectomy has been achieved by adding a hypoglossal-thoracodorsal nerve graft.[30]

The latissimus dorsi is a large, flat, triangular muscle that sweeps over the lumbar region and lower thorax. It arises from the posterior iliac crest, thoracolumbar fascia, external oblique fascia, and lower six vertebrae. It inserts into the intertubercular groove of the humerus. This muscle acts to adduct, extend, and medially rotate the humerus. Its blood supply is from the thoracodorsal artery. After the subscapular artery has branched to give off the circumflex scapular artery, it continues as the thoracodorsal artery. There it is joined by the vein and thoracodorsal nerve, which provides the muscle's

motor innervation. This neurovascular bundle enters the muscle at a single neurovascular hilum on its internal (costal) surface 6 to 12 cm from the subscapular artery and 1 to 4 cm medial to the lateral border of the muscle.[29] The artery gives off one to three large branches to the serratus anterior muscle before dividing at the hilum to supply the latissimus. The long thoracic nerve provides cutaneous sensory innervation. The vascular supply of the latissimus dorsi muscle is not affected by prior radiation. Reconstruction can be performed employing either a pedicled or free flap technique with equally high success rates[30] (Figure 41–9).

The flap does have some disadvantages. A lateral decubitus position is needed to harvest this flap, creating a need for intraoperative repositioning. In addition, when pedicled, the flap is passed through the axilla between the pectoralis major and minor muscles.

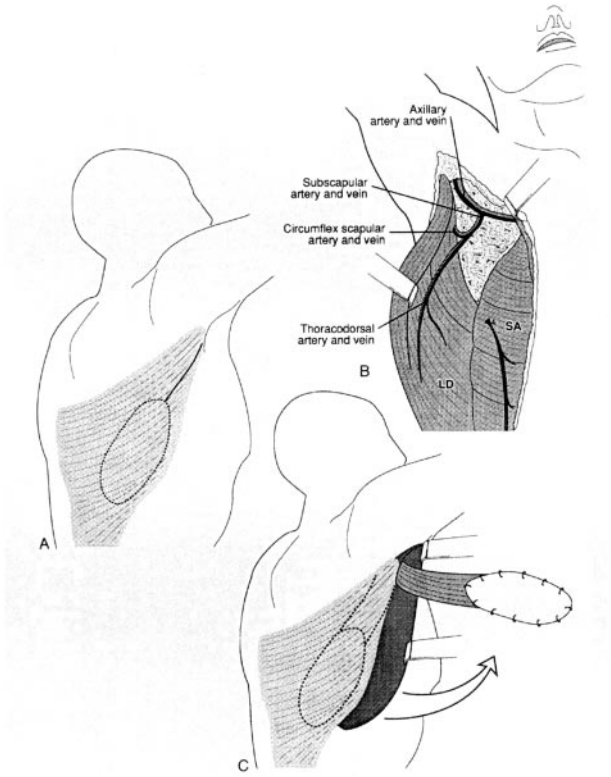


FIGURE 41–9. The latissimus dorsi myocutaneous flap: *A*, flap design; *B*, flap dissection; *C*, rotation of flap. Reproduced with permission from Shindo ML, Sullivan MJ. Muscular and myocutaneous pedicled flaps. Otolaryngol Clin North Am 1994;27:166.

Postoperatively, this puts the vascular pedicle at risk of compression in the region of the anterior axilla by the upper arm. The reconstructive surgeon needs to be familiar with the anatomy of the axilla, the axillary vessels, and their variations. Finally, owing to the shoulder dysfunction that may arise, this flap should be avoided if the trapezius muscle, pectoralis major muscle, or spinal accessory nerve has previously been compromised.[23,31]

In summary, this is a reliable and versatile flap that works well as a pedicled or free flap. It may provide coverage to distant areas, such as the scalp, lateral part of the skull base, and orbitomaxillary region, where a flap with a large surface area is required. It is also an excellent option for dynamic reconstruction after total glossectomy.

Gastric Pull-up The gastric pull-up plays an important role in reconstruction of total pharyngoesophageal defects when the tumor extends into the cervical esophagus. The gastric pull-up allows for single-stage reconstruction with good oral rehabilitation. The gastric tube is a well-vascularized flap that may be used in a previously irradiated or operated field. The technique requires a general surgical team to mobilize the stomach and to create the gastric tube. Posterior mediastinal dissection is performed through both the diaphragm and a cervical approach (Figures 41–10 and 41–11). With adequate mobilization, the distal portion of the flap may reach the soft palate. The flap is contraindicated in patients with severe pulmonary and hepatobiliary disease. Mortality rates as high as 18% and an incidence of severe complications reaching 50% have been reported.[32,33] The morbidity and mortality can be decreased to acceptable levels in carefully selected patients.

Free Flaps

Although microvascular free tissue transfer was introduced in the 1970s, it was not until the late 1980s that free flaps became popular. Pedicled regional flaps dominated the 1980s, reflective of their ease of use with shortened operative time while eliminating the skills required in microvascular surgery.

The use of microvascular free tissue transfer blossomed in the late 1980s and has become one of the primary reconstructive modalities in head and neck surgery. Flap survival rates ranging from 91 to

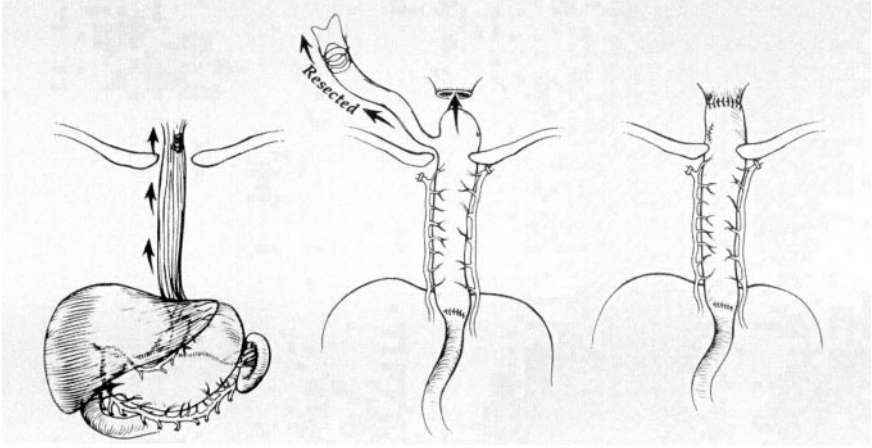


FIGURE 41–10. Technique of gastric pull-up. Reproduced with permission from Schusterman MA, Shestak K, deVries EJ, et al. Reconstruction of the cervical esophagus: free jejunal transfer versus gastric pull-up. Plast Reconstr Surg 1990;85:17.

99% are routinely achieved, reflecting a more predictable axial blood supply.[34,35] Free tissue transfer also allows for more donor-site options relating to size, thickness, and the need to include bone or innervation potential. Single-stage reconstruction was now possible in defects that had previously required a delayed procedure. Prior radiation did not affect flap survival.[36,37]

The success of free tissue transfer goes far beyond designing of flap dimensions. It requires

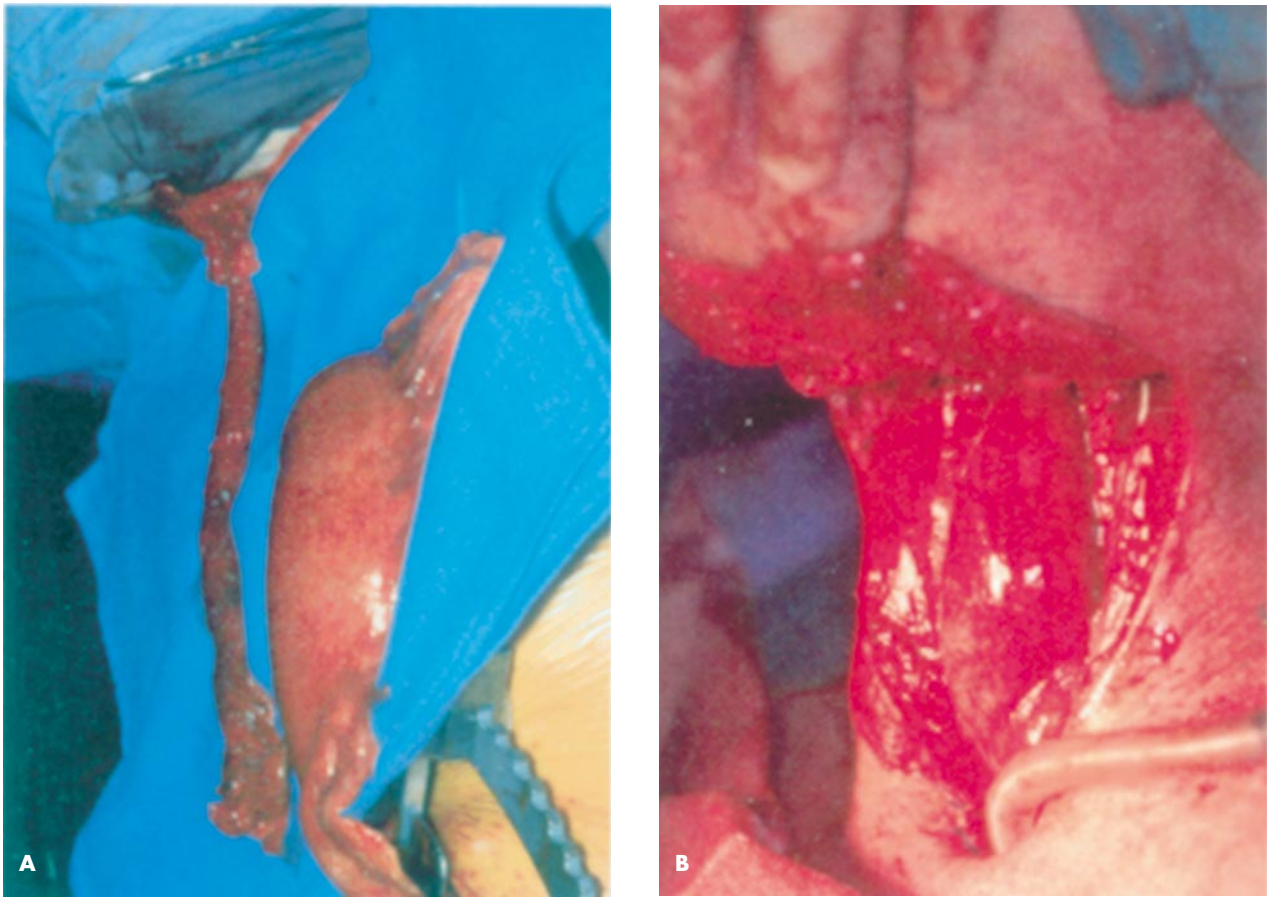


FIGURE 41–11. A, Proximal esophagus resected en bloc with hypopharynx. Note the mobilized gastric tube. *B*, Gastric tube in place. Note the pharyngogastric anastomosis.

careful planning and begins with an understanding of both the donor-site vascular pedicles and recipient-site vessel options. Careful harvest of the nutrient vascular pedicle is required, taking into account the length of pedicle needed to prevent tension on the vascular anastomosis while avoiding excessive length, which might result in twisting and kinking. The highly vascular nature of the head and neck region normally provides a plethora of recipient vessels to choose from. Previous or associated surgical resection required in a neck dissection may severely limit recipient vessel options. Prior radiation to the neck may lead to atherosclerosis of the recipient artery.[38] Determination of the length of the vascular pedicle and choice of recipient vessels are ultimately made at the time of resection and reconstruction.

The most critical aspect of microvascular reanastomosis is the accurate apposition of the cut edges of the vessels and maintenance of intimal continuity. The venous anastomosis is typically more demanding technically, given the thin and delicate walls of veins, and may lead to venous thrombosis, the most common cause of flap failure. The critical period for venous thrombosis is the first 3 to 5 days.

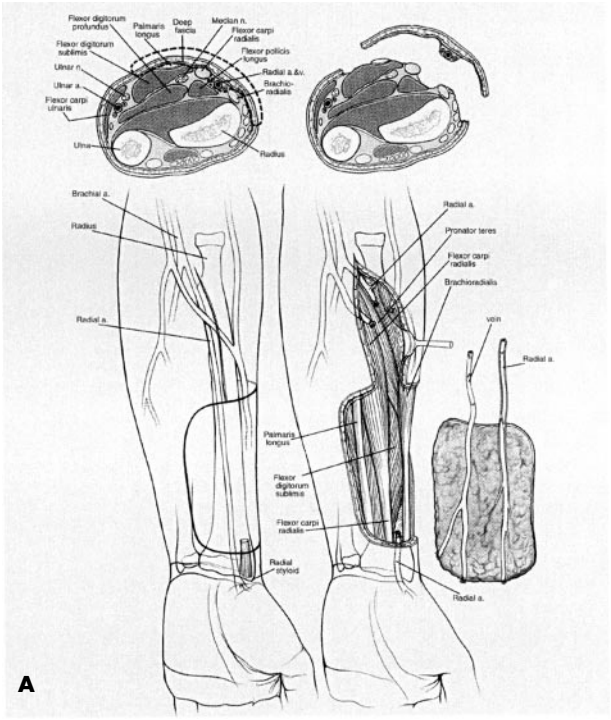
A variety of choices are generally available in the neck for recipient vessel selection, although it does partially depend on the region to be reconstructed. The most common choices for recipient arteries are the ipsilateral lower branches of the external carotid artery or branches of the thyrocervical trunk.[38] When ipsilateral recipient vessels are not available, contralateral neck vessels may be used as an alternative option. Interposition vein grafts may be necessary if the pedicle length is inadequate.[38] The primary recipient veins are the external and internal jugular veins and the transverse cervical vein. When these options are not available on the ipsilateral neck, the contralateral neck veins would be the next alternative.[38]

Some have raised the question of cost effectiveness in the use of microvascular free tissue transfer. Several studies have shown that the cost of free flaps is ultimately comparable to pedicled regional flaps. This slightly higher cost is more than offset by the superior functional results and lower complication rates.[39,40]

Radial Forearm Free Flap The radial forearm free flap is the most frequently chosen free flap for head and neck reconstruction. This pliable, robust flap has the ideal thinness for pharyngeal, palatal, and floor-of-mouth defects. With the addition of sensory reinnervation, it has become quite useful for tongue and hemilaryngopharyngeal reconstruction.

This flap is harvested from the volar surface of the forearm. The vascular supply is the radial artery and the cephalic vein or the venae comitantes (Figure 41–12).

Although a composite bone graft can be obtained by adding a portion of the radius, no more



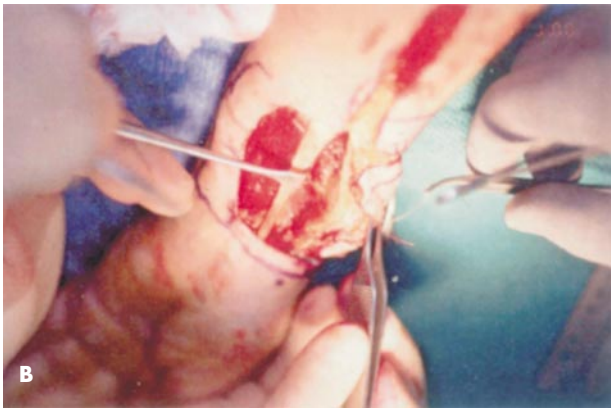


FIGURE 41–12. A, Pictorial representation of radial fasciocutaneous flap. Reproduced with permission from Silver CE, Rubin JS, editors. Atlas of head and neck surgery. Philadelphia: Churchill Livingstone; 1999. *B*, Intraoperative harvest of radial forearm free flap.

than 40% of the radius can be used without significantly increasing the risk of a radius fracture. Such fractures may adversely affect wrist mobility and hand strength. Because of the small size of this bone graft and the significant morbidity related to the fracture, other composite grafts, such as the fibular and the scapular flaps, are usually better choices when bone is required.

The ability to provide sensation to the flap has also made this flap particularly advantageous in oral cavity and pharyngeal reconstruction. The lateral antibrachial cutaneous nerve, which runs with the cephalic vein, is the nerve most commonly used for reinnervation. The medial antibrachial cutaneous nerve is also available.

The biggest concern with harvesting this flap is adequate perfusion of the hand by the ulnar artery via the palmar arch. It is crucial that the Allen test is properly performed preoperatively to assess the perfusion of the hand by the ulnar artery when the radial artery is occluded. Doppler ultrasonographic evaluation is also helpful in evaluating patency of the ulnar artery. Despite careful preoperative testing, if inadequate perfusion of the hand is discovered, a reverse vein graft can be placed to reconstitute flow through the radial artery.

Complications primarily arise at the donor site. In addition to the risk of radius fracture, failure to raise the flap in a suprafascial plane may lead to poor skin graft take and tendon exposure with subsequent wrist and hand dysfunction.[41]

Lateral Arm Fasciocutaneous Flap The lateral arm flap is a moderately thin and hairless flap. Its main advantage over the radial forearm flap is its minimal donor-site morbidity. The donor site can usually be closed primarily, no immobilization is required, and vascular compromise is not a concern.

This fasciocutaneous flap is centered on the intermuscular septum at the distal half of the lateral part of the upper arm. The flap is supplied by the posterior branch of the radial collateral artery, the terminal branch of the porfunda brachii artery. The vascular pedicle runs with the radial nerve along the anterolateral portion of the brachioradialis muscle and measures 10 to 12 cm. The posterior branch then enters the intermuscular septum. The venous drainage is via the cephalic vein and the venae comitantes. The inferior lateral brachial cutaneous nerve is harvested with the flap to provide sensation to the flap (Figure 41–13).

The primary limitation for this flap is the small caliber of the vascular pedicle with its associated decreased survival rate.[23]

Lateral Thigh Fasciocutaneous Flap The lateral thigh flap is a lower extremity version of the lateral arm flap while providing a much larger surface area. Hayden and Dreschler have described harvesting a graft measuring 27 × 14 cm; however, the exact limits of the graft size have not been established.[42] This flap has been used successfully in reconstructing total laryngopharyngectomy and total glossectomy defects and may effectively resurface large soft tissue defects in the cervical region. The donor site is usually closed primarily leaving an esthetically acceptable linear scar with minimal donor-site morbidity.

The skin paddle is centered over the intermuscular septum, which separates the vastus lateralis and the iliotibial tract anteriorly from the biceps

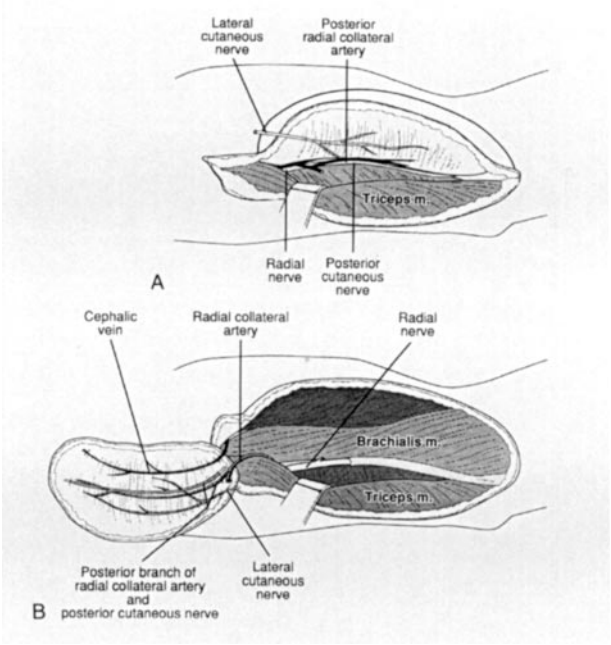


FIGURE 41–13. The lateral arm flap. *A*, Initial elevation of the posterior flap and identification of the vascular pedicle. *B*, The entire flap elevated and pedicled on the radial collateral artery. Reproduced with permission from Shindo ML, Sullivan MJ. Soft-tissue microvascular free flaps. Otolaryngol Clin North Am 1994; 27:178.

femoris posteriorly (Figure 41–14). It can be roughly located by drawing a line from the greater trochanter to the lateral epicondyle. The third perforator of the profunda femoris artery, which is usually the dominant vascular supply to the flap, can be found emerging from the septum at the midpoint of this line. Although the third perforator is usually the dominant blood supply, any one of the second to fourth perforators may prove to be the dominant vessel. A sensate flap can be obtained by harvesting the lateral femoral cutaneous nerve.

Preoperative evaluation is extremely important. If peripheral vascular disease exists, another donor site should be chosen. Because the thickness of the flap depends on the body habitus of the patient, the donor site must be carefully evaluated to match the needs of the recipient defect. This flap is rarely used in obese patients.

The complication rate compares favorably to other more commonly used free and pedicled flaps.[42] Donor-site morbidity includes seroma formation and compartment syndrome. The risk of compartment syndrome has been significantly minimized by wide undermining to avoid excessive tension and use of an active drain.

Latissimus Dorsi Free Flap The latissimus dorsi flap can be harvested as a pedicled or free flap. Although the pedicled flap can reach up to the orbitomaxillary and lateral skull base lesions, the free flap allows coverage of large scalp and cranial defects anywhere on the calvarium.

The anatomy of this flap was discussed earlier in the pedicled regional flaps section.

A vascular pedicle as long as 12 cm can be obtained by dissecting up to its takeoff from the axillary vessels.

As with the pedicled version, a lateral decubitus position is required in its harvesting. In addition, donor-site morbidity with shoulder dysfunction may develop.

Rectus Abdominis Free Flap The popularity of the rectus abdominis free flap relates to its ease to harvest, long vascular pedicle, and excellent reliability, with a success rate as high as 97.5%.[43] It provides a large surface area and bulk, making it particularly useful in extensive skull base defects. It is also effective in covering the orbitomaxillary regions and reconstructing total glossectomy defects. The bulk of the flap depends on body habitus. The flap can be harvested using the entire rectus as a small single island flap or as multiple skin islands. Attempts at a sensate flap have generally not been successful.

The flap is thicker inferiorly and then thins out superiorly above the costal margin.

The rectus abdominis muscle arises from the pubic crest and symphysis pubis and inserts into the fifth to seventh costal cartilages. The muscle is enveloped anteriorly and posteriorly by the rectus

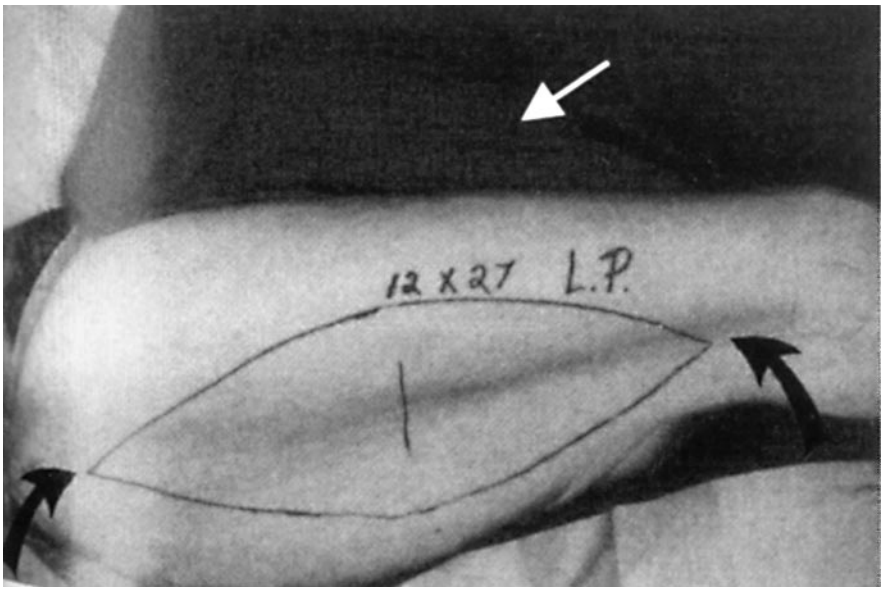


FIGURE 41–14. Preoperative planning of lateral thigh free flap harvest. The axis of the flap is based on the posterolateral intermuscular septum (*arrows*). Reproduced with permission from Hayden RE, Desch DG. Lateral thigh free flap for head and neck reconstruction. Laryngoscope 1999;109:1490.

sheath. It has two dominant vascular pedicles, the deep inferior epigastric and the superior epigastric vessels. The two vascular systems anastomose in the region above the umbilicus through a system of small caliber vessels called "choke vessels."[44] The deep inferior epigastric artery, with its large vessel diameter, is the preferred vascular pedicle for harvest. It gives rise to the major musculocutaneous perforators and thus supplies a larger area of skin (Figure 41–15, A). The largest concentration of cutaneous perforators is located in the paraumbilical region, with the cutaneous vessels radiating superiorly and laterally toward the costophrenic angle. The flap is designed to capture these perforators (Figure 41–15, B). The pedicle length can be extended up to 15 cm. The deep inferior epigastric artery and vein can be found entering the lateral border of the rectus muscle halfway between the pubis and umbilicus, about 3 to 4 cm below the arcuate line. Topographically, the arcuate line is located approximately at the level of the anterior superior iliac spine. The arcuate line is an important anatomic marker, not only to identify the vascular pedicle but also to identify the transition zone where the posterior sheath is composed of all three aponeuroses of the oblique muscles superiorly and only the transversalis fascia inferiorly. This creates a potential weak area below the arcuate line that puts it at risk for a ventral hernia if the anterior sheath is not carefully reapproximated in this region.

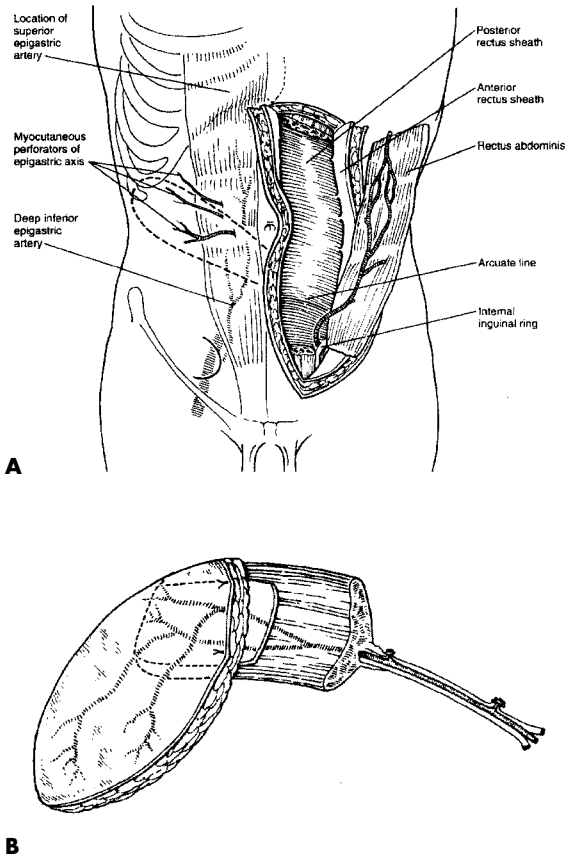


FIGURE 41–15. A, Anatomy of the rectus abdominis muscle illustrating the tendinous inscriptions and blood supply. *B,* Oblique design of the skin paddle on the rectus muscle free flap. Reproduced with permission from Myers EN, editor. Operative otolaryngology-head and neck surgery. Vol I. Philadelphia: WB Saunders; 1997.

Donor-site morbidity is a major limitation, and this includes acute pain, ventral hernia, gait disturbance, and paresthesias with radiculopathy.

Contraindications for this flap are previous abdominal surgery, intrapelvic surgery, inguinal herniorrhaphy, and vascular surgery on the iliac system.[23]

Osteocutaneous Composite Free Flaps

Fibular Free Flap The fibula provides the longest bone graft, up to 25 cm, and thus allows reconstruction of the mandible from condyle to condyle. Because it is a non–weight-bearing bone, it can be harvested in its entirety, leaving only a 6 to 7 cm segment proximally and distally. This flap can be raised as a bone-only graft or as a composite graft by adding soft tissue and skin (Figure 41–16).

Several advantages of this flap make it particularly desirable for mandibular reconstruction. The cross-sectional area of the fibula is similar to that of the mandible. Osseointegrated dental implants can be placed primarily to provide masticator function. Finally, the cutaneous portion of the flap can be made sensate by the anastomosis of the lateral sural cutaneous nerve of the calf to the lingual nerve. The sural communicating nerve can also be anastomosed to the inferior alveolar nerve to provide sensation to the lower lip.[45]

The peroneal artery and vein provide the blood supply to the fibula via two routes. The endosteal blood supply is provided by the nutrient vessels that enter the fibula at the junction of the proximal and middle thirds of the bone. Several segmental perforators provide the periosteal blood supply. Thus,

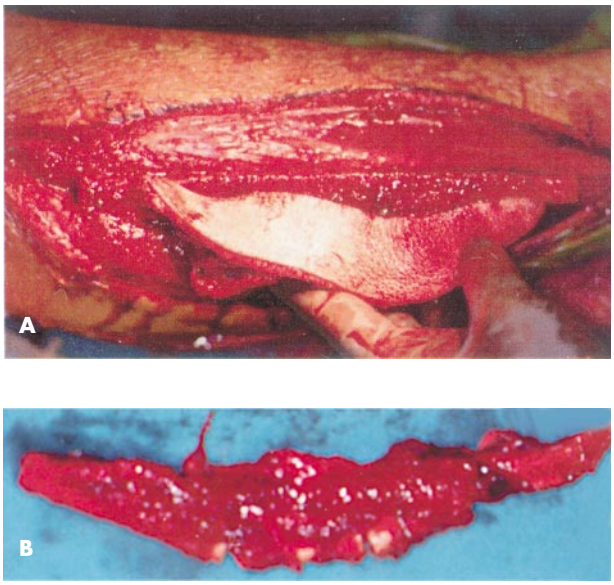


FIGURE 41–16. A, Composite fibula flap with skin, muscle, and bone. *B*, Bone-only fibula flap with osteotomies.

multiple osteotomies can be performed to contour the bone graft without threatening its viability. The septocutaneous perforators and the musculocutaneous perforators through the flexor hallucis longus and soleus muscles provide the vascular supply to the overlying lateral skin. Because of the importance of the musculocutaneous perforators in maintaining skin viability, it is important to harvest a cuff of soleus muscle and the flexor hallucis muscle when obtaining a composite graft (Figure 41–17).

Preoperative assessment of a lower extremity vascular pattern is accomplished with traditional or magnetic resonance angiography. These studies provide crucial information regarding peripheral vascular disease and an anomalous blood supply to the foot, which may be compromised with sacrifice of the peroneal artery.[46,47] A history of lower extremity fracture is a contraindication for this flap.

The most serious donor-site morbidity is the potential vascular compromise to the foot secondary to inadequate collateral flow after the peroneal artery is harvested.[46] This complication can be avoided through appropriate preoperative angiography. Vascular compromise may also occur as a result of an excessively tight closure. Distal pulses should be monitored in the postoperative period. Other donor-site complications include cold intolerance, edema, and weakness in dorsiflexion of the great toe owing to either injury to a branch of the peroneal nerve or scarring of the flexor hallucis longus. The common peroneal nerve should be identified early in the dissection to avoid traction injury, which may lead to an equinovarus deformity and anesthesia along the anterolateral surface of the lower leg and dorsum of the foot. A 6 to 7 cm segment of fibula should be preserved both proximally and distally to preserve the common peroneal nerve proximally and maintain stability of the ankle joint distally. In spite of these precautions, patients may experience pain and weakness on ambulation for several months.[46]

Iliac Crest Free Flap The iliac crest composite graft reflects the natural curve and length of the mandible (up to 16 cm), making it a good option for mandibular reconstruction. This graft provides adequate width and depth for placement of osseointegrated dental implants but does not offer the option of sensate rehabilitation.

The deep circumflex iliac artery supplies the iliac bone and arises from the lateral aspect of the external iliac artery approximately 1 to 2 cm above the inguinal ligament. It gives off the ascending branch to supply the internal oblique muscle. The vascular pedicle runs along the internal aspect of the iliac bone. Thus, osteotomies can be performed on the external cortex to contour the neomandible. The deep circumflex iliac artery gives off several perforators that traverse through the external oblique, internal oblique, and transversalis fascia to provide the blood supply to the overlying skin. The cutaneous perforators exit the external oblique in a zone 9 cm posterior to the anterior superior iliac spine and about 2.5 cm medial to the iliac crest. The skin paddle should be centered on a line drawn from the anterior superior iliac spine to the inferior angle of the ipsilateral scapula. The cutaneous portion makes this flap bulky and limits rotation of the soft tissue. The thin internal oblique can also be harvested to provide intraoral reconstruction by wrapping it around the iliac bone graft and allowing it to mucosalize secondarily. The skin paddle may then be used to reconstruct the external cutaneous defect or to act as an external monitor for flap viability (Figure 41–18).

Donor-site morbidity is the primary limitation. Prolonged donor-site pain and gait disturbances are common.[48]

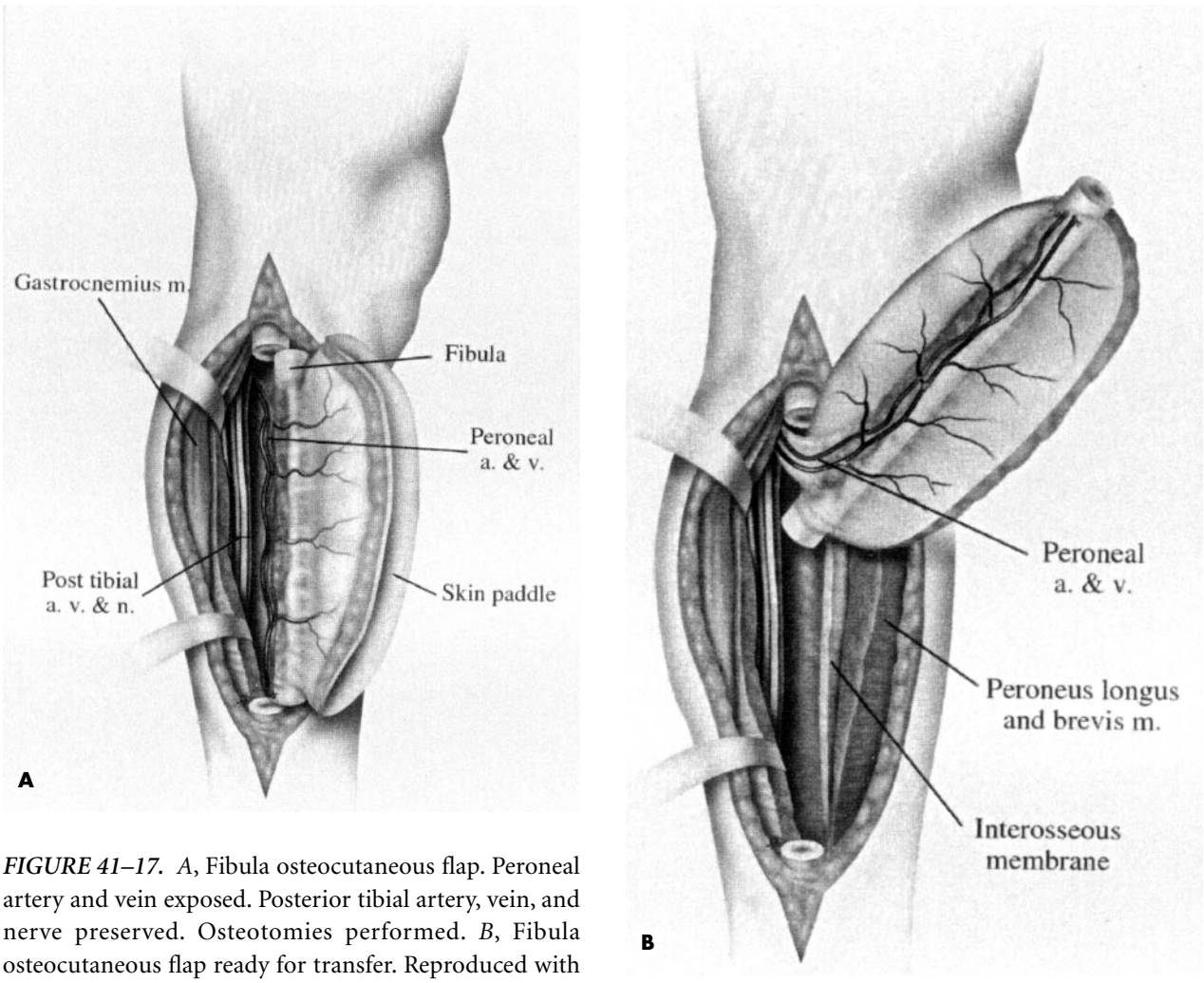


FIGURE 41–17. A, Fibula osteocutaneous flap. Peroneal artery and vein exposed. Posterior tibial artery, vein, and nerve preserved. Osteotomies performed. *B*, Fibula osteocutaneous flap ready for transfer. Reproduced with permission from Friedman M. Operative techniques in otolaryngology-head and neck surgery—flaps for head and neck reconstruction. Part 1. Philadelphia: WB Saunders; 2000.

Scapular and Parascapular Free Flaps This thin, pliable flap provides a tremendous amount of flexibility as a result of the ability to harvest the soft tissue separately from the bone graft and the capacity to rotate these separate flaps in different orientations.

The circumflex scapular artery supplies the majority of the osseous branches to the scapula. The angular artery, a branch of the thoracodorsal artery, supplies the most inferior aspect of the lateral part of the scapula. The scapula receives both an endosteal and periosteal blood supply, but the periosteal component dominates, so the periosteum should be carefully preserved during osteotomies. The terminal branches of the circumflex scapular artery exit through a triangular space bound by the teres major, teres minor, and long head of the triceps lateral to the scapula. One terminal branch divides in the subdermal fascia to send a horizontal branch on which the scapular skin flap is based and an oblique branch on which a parascapular skin flap is based. The vascular pedicle can be extended up to 14 cm by dissecting the pedicle up to its takeoff from the subscapular artery and vein. Previous axillary node dissection is a contraindication for this flap.

The scapular bone is harvested from the lateral border of the scapula 1 cm below the glenoid fossa to the tip of the scapula. The addition of the angular artery and vein allows for the harvest of two bony segments or a single longer segment, up to 12 cm in

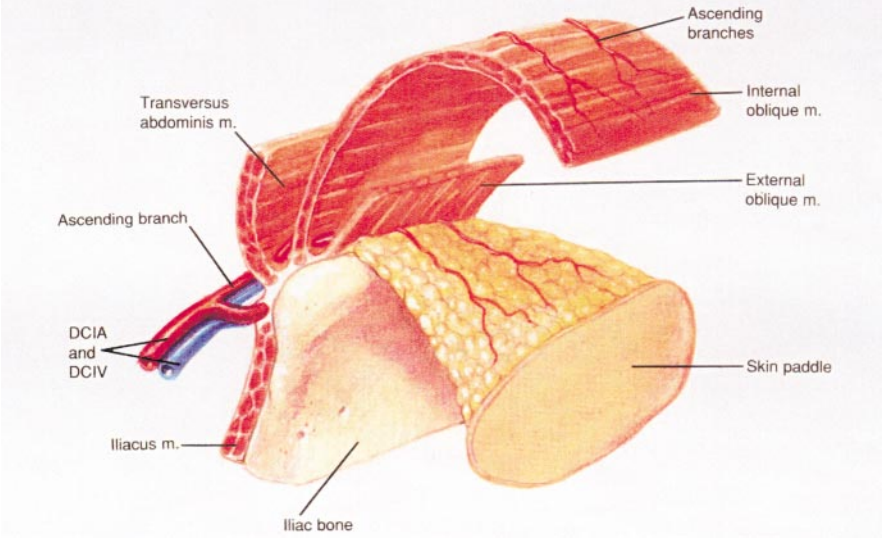


FIGURE 41–18. The tripartite iliac internal oblique osteomusculocutaneous flap. When this flap is inset into the oral cavity for oromandibular reconstruction, it is usually turned 180 degrees so that the crest forms the inferior border of the neomandible. Adapted from Urken ML, Cheney ML, Sullivan MJ, Biller HF, editors. Atlas of regional and free flaps for head and neck reconstruction. Philadelphia: Lippincott Williams and Wilkins; 1995.

length. Despite the fact that the bone is thinner and not as strong as the iliac crest or fibular bone graft, it is still a viable option for mandibular reconstruction and may permit the placement of osseointegrated dental implants. Harvesting this bone does require the detachment and reattachment of the teres major muscle to the remaining scapula to prevent donor-site morbidity. This flap provides a larger surface area of skin for reconstruction than the fibular or iliac crest composite flap. The two skin flaps can also be placed in different three-dimensional orientations since they have separate pedicles and may be used to resurface two contiguous sites (Figure 41–19).[49]

The donor-site defect, which is closed primarily, is located posteriorly, making it cosmetically advantageous. Donor-site morbidity is relatively infrequent and consists of shoulder weakness and diminished range of motion.[50]

Harvesting this flap requires special positioning and makes synchronous two-team surgery difficult.

This flap is well suited for composite defects that require a large amount of soft tissue. Its thinness and pliability also make it quite useful in palatal and orbital floor reconstruction. With de-epithelializing, this flap becomes useful for reconstruction of midface defects.[49]

Visceral Flaps

Visceral flaps provide certain advantages over cutaneous flaps. They provide a thin, pliable mucosal surface that is particularly beneficial in oral cavity

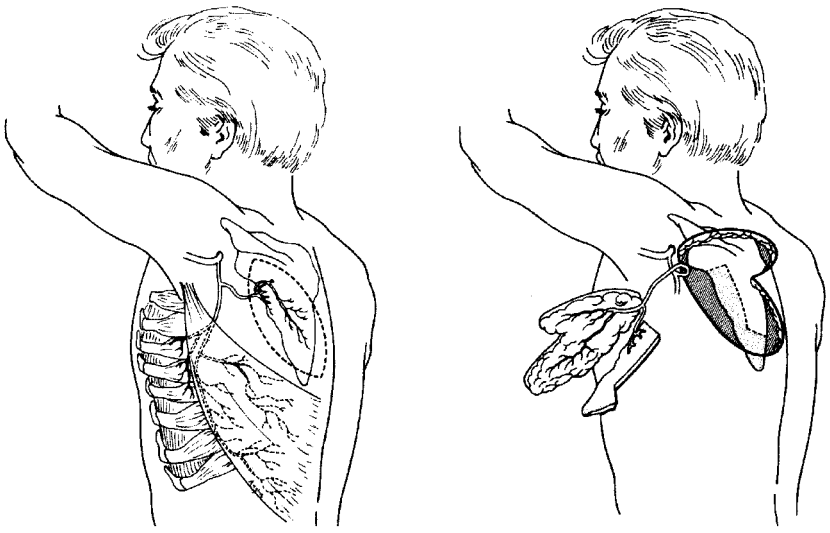


FIGURE 41–19. The lateral border of the scapula received a rich vascular supply from several vessels that leave the circumflex scapular artery near the bone. Since the vascular pedicle to the skin is separate from the pedicle to the bone, 360-degree rotation of the skin paddles around the axis of the one flap is possible. Adapted from Paparella MM, Shumrick DA, Gluckman JL, Meyerhoff WL, editors. Otolaryngology. Vol IV. 3rd ed. Philadelphia: WB Saunders; 1991.

and pharyngeal reconstruction. Because these visceral flaps also secrete mucus, the problems of xerostomia are diminished. In fact, in the early postoperative period, excess mucus production can be a problem. Despite these advantages, violation of the abdomen is required to obtain these flaps. The benefits of the visceral flaps must be balanced with potential intra-abdominal complications such as adhesions, peritonitis, and intra-abdominal abscess. The advent of laparoscopic surgery may provide an optimal method of harvesting these flaps with lessened morbidity.[51–54]

Jejunum Free jejunal transfer in head and neck reconstruction has been widely used for reconstruction of the pharyngoesophagus after total laryngopharyngectomy, and its use has been more recently expanded to encompass reconstruction of the oral cavity and oropharynx.

The jejunum is located just distal to the duodenum and constitutes 25 cm of proximal small bowel (Figure 41–20). The harvest is usually performed from the second loop of jejunum, which is perfused by the superior mesenteric artery. The superior mesenteric artery supplies the jejunum by forming a vascular arcade that divides the jejunum into segments that are based on a single mesenteric artery and vein. An appropriate-sized segment is harvested, along with the segmental vascular pedicle. If a pharyngoesophageal reconstruction is planned, the graft should be marked to differentiate the proximal from the distal end as the graft must be aligned with the jejunum in an isoperistaltic direction. The diameter of the jejunum is similar to the diameter of the proximal esophagus. Proximal anastomosis to the broader base of the tongue requires the jejunum to be bivalved, thereby increasing its circumference. Excessive length of the segment should be avoided to prevent redundancy with its associated stasis and dysphagia. To monitor flap viability during the early postoperative period, a small indicator segment of jejunum is exteriorized. Tracheoesophageal puncture can be performed but provides a wet voice (Figure 41–21).[55]

By opening the jejunum along its antimesenteric border, a patch graft is created.[56] This provides a thin, pliable flap that secretes mucus and resists contraction (Figure 41–22).

Primary contraindications for this procedure involve intra-abdominal pathology, particularly a history of ascites and chronic inflammatory bowel diseases.[55]

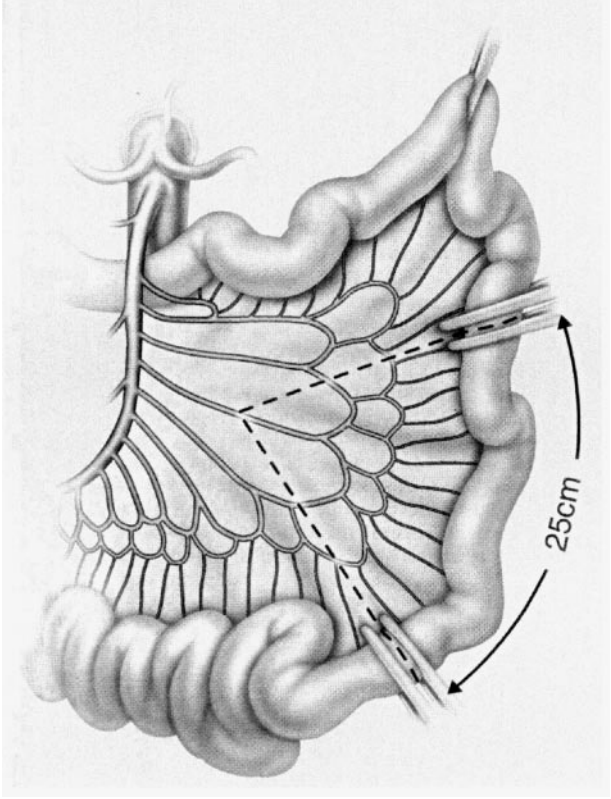


FIGURE 41–20. Jejunum is resected for use as a graft. Reproduced with permission from Friedman M. Operative techniques in Otolaryngology-head and neck surgery—flaps for head and neck reconstruction. Part 1. Philadelphia: WB Saunders; 2000.

Gastro-omental Free Flap A free omental or gastro-omental flap offers a great deal of versatility. This pliable flap may be readily tailored to provide an appropriate amount of bulk. Almost the entire omentum can be safely harvested. Conversely, the omentum can be trimmed away from the vascular system to thin the flap. The gastro-omental flap is ideal for intraoral and oropharyngeal defects. The addition of gastric mucosa provides a thin, pliable tissue with a smooth, mucus-secreting surface. The omental portion can be used to protect the carotid. The omental flap has been used for a variety of defects, taking advantage of these special properties. The omentum provides neovascularization to ischemic tissue through capillary ingrowth.[57,58] Because of this property, the omentum may be used to cover calvarial

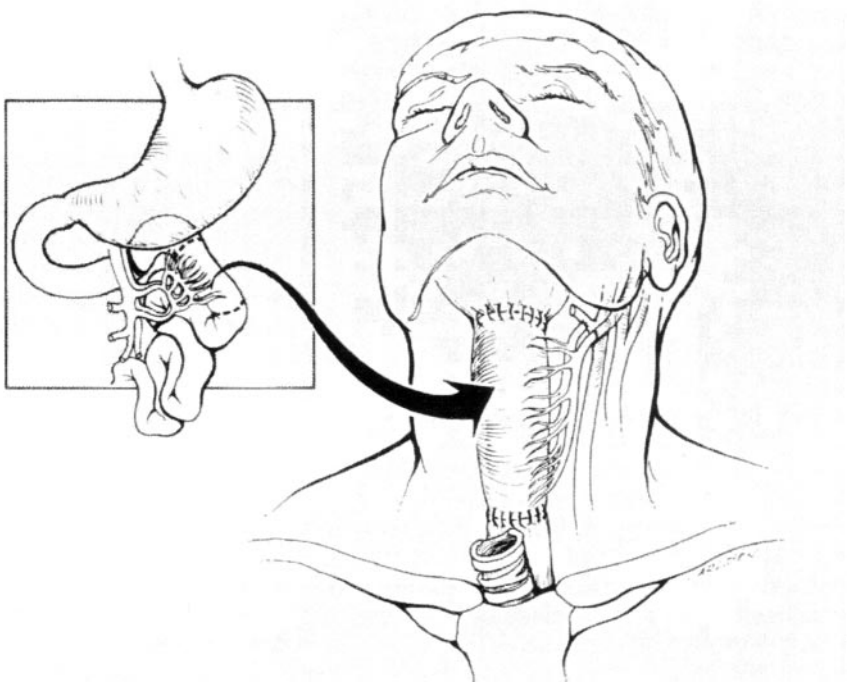


FIGURE 41–21. Technique of free jejunal transfer. Reproduced with permission from Schusterman MA, Shestak K, deVries EJ, et al. Reconstruction of the cervical esophagus: free jejunal transfer versus gastric pull-up. Plast Reconstr Surg 1990;85:17.

grafts or to revascularize areas of osteoradionecrosis and osteomyelitis. The omentum is also well suited for the augmentation of the midface, reflecting its ability to be subdivided into compartments.

The greater omentum is a large, double-layered peritoneum that hangs from its attachments to the greater curvature of the stomach and the transverse colon. The greater curvature of the stomach and the omentum are both supplied by the right and left gastroepiploic arteries. The right gastroepiploic is usually chosen as the pedicle because of its larger vessel diameter and longer pedicle length. This flap is rela-

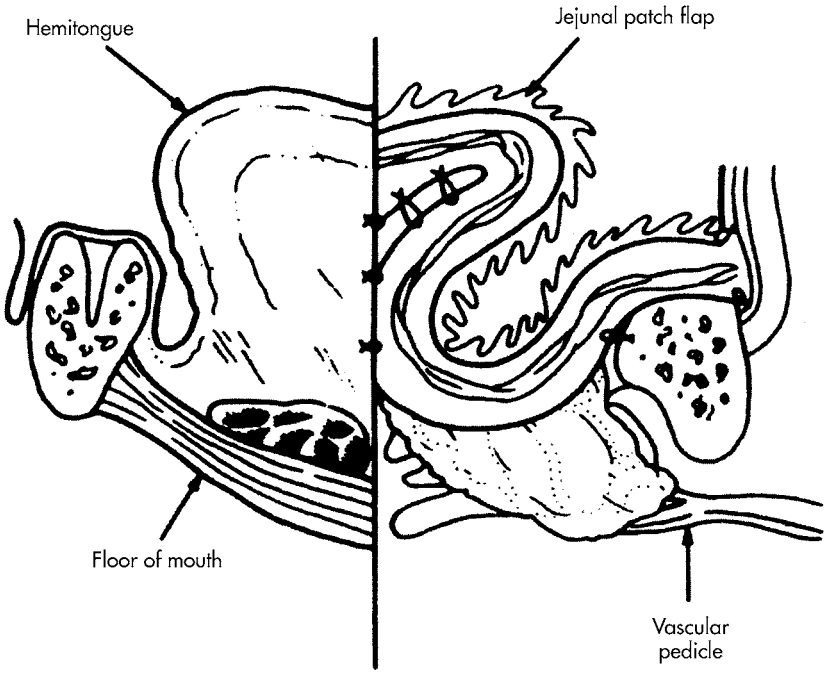


FIGURE 41–22. Schematic diagram showing free jejunal patch flap for reconstruction of hemitongue, floor of mouth, and hemi-mandibulectomy. Reproduced with permission from Cocks HC, Kumar BN, Gupta AR, Simms MH. Free jejunal patch flaps in oral and oropharyngeal reconstruction. J Laryngol Otol 1999;113:681.

tively easy to harvest but has the drawback of requiring a laparotomy (Figure 41–23).

This flap is contraindicated in patients with a history of gastric outlet obstruction and peptic ulcer disease.

RECONSTRUCTION OF REGIONAL DEFECTS

Oral Cavity and Pharynx

Reconstruction of the oral cavity and pharynx can be a challenging endeavor, and one must take into consideration the needs for normal deglutition and speech. Small intraoral defects may be allowed to heal by secondary intention. Mucous membrane grafts provide excellent coverage with minimal contracture. Local flaps are useful for the reconstruction of moderate-sized floor-of-mouth defects. Mucous membrane, platysma, and nasolabial flaps are readily available and provide dependable and adequate coverage in most patients. They are associated with few functional and cosmetic consequences and minimal donor-site morbidity.[59]

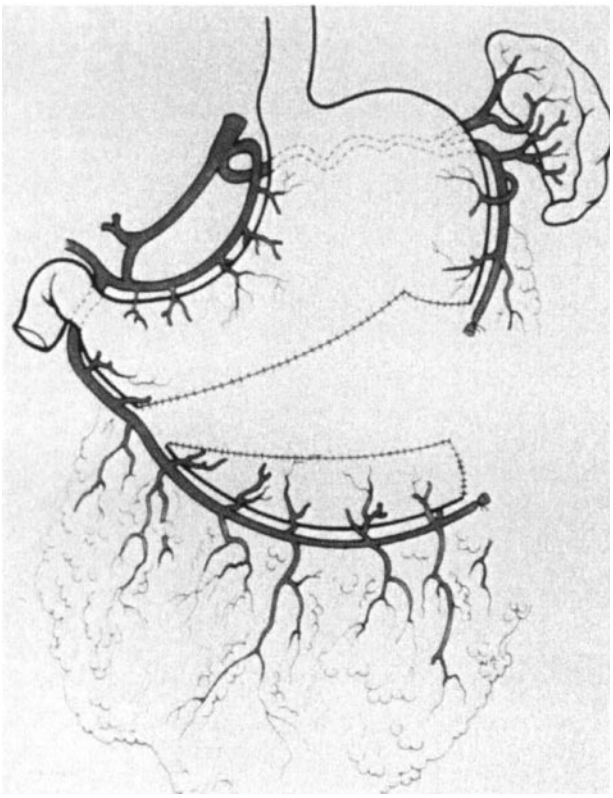


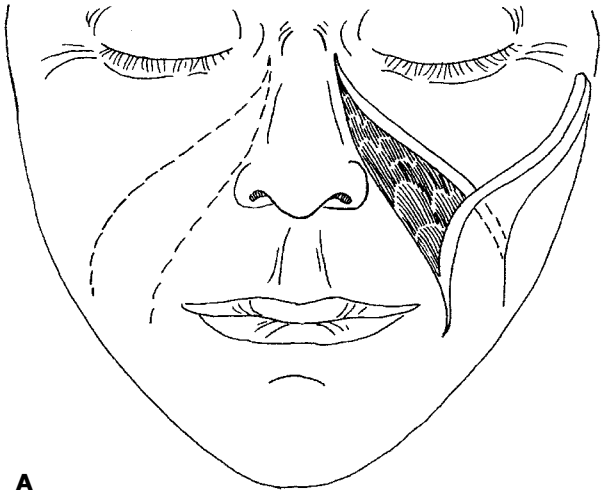
FIGURE 41–23. Schematic view of the free gastro-omental flap being harvested. The left gastroepiploic vessels have been sectioned. A stapling device has been employed to section a piece of the greater curvature from the rest of the stomach. The stomach flap with the omentum is pedicled on the right gastroepiploic vessels. Reproduced with permission from Hayden RE. Microvascular free flaps for soft-tissue defects. Otolaryngol Clin North Am 1991; 24:1343.

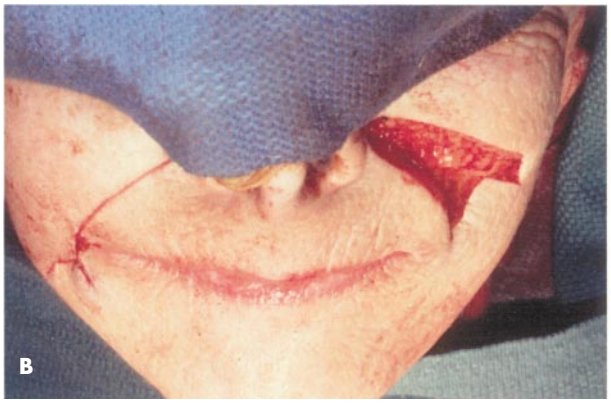
The tunneled, inferiorly based nasolabial flap is well suited for anterior intraoral defects. By using both cheeks as the donor site, up to 25 cm^2 of tissue may be provided for intraoral lining. The blood supply from the labial artery is excellent, making delay unnecessary. The flaps maintain a supple floor of the mouth, allowing for mobility of the remaining tongue. They provide excellent coverage for the exposed mandible after removal of the alveolus or lingual cortex with minimal bulk. Donor-site morbidity is minor, and cosmesis is excellent (Figure 41–24).[59]

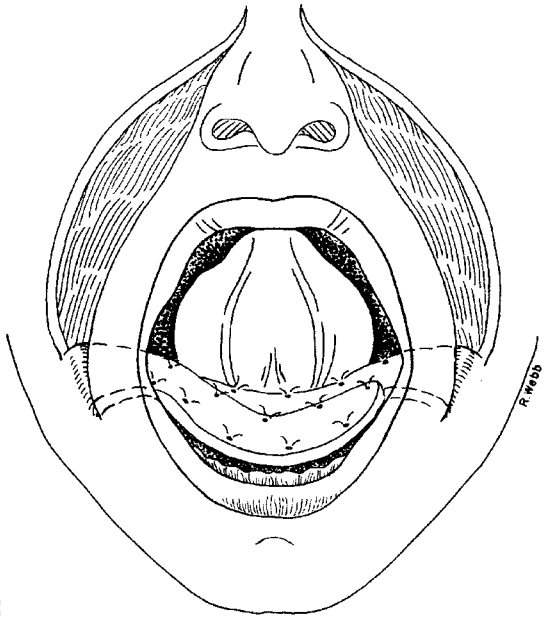
Xerostomia is a significant problem with oral cavity cancers that require surgery and radiation therapy. The visceral flaps, including the jejunal patch and gastro-omental flap, have the advantage of providing tissue that continues to secrete mucus, making them appropriate for postradiation failures. These flaps work particularly well in areas for which a thin flap is preferred, such as the floor of the mouth, alveolar ridge, pharyngeal wall, and buccal mucosa. The jejunal flap can be folded in on itself to reconstruct a soft palate defect. Gastro-omental flaps are more appropriate when the oral cavity defect is associated with a large soft tissue neck defect. These visceral flaps may be the only flaps that are thin and pliable in an obese patient.

Large defects of the floor of mouth and oropharynx can be readily reconstructed with a pedicled regional flap. The pectoralis major myocutaneous flap is the flap of choice when a quick reliable flap is desired. A fasciocutaneous modification is appropriate when a large flap with less bulk is required.

Tongue mobility, adequate bulk, and sensation are key factors in planning oral cavity reconstruction. The radial forearm free flap meets these needs and represents the preferred choice for reconstruction after partial glossectomy. In addition, this flap can be bivalved to reconstruct both the tongue and the floor of the mouth. The lateral arm flap is an alternative if the radial forearm is not available.[60,61]







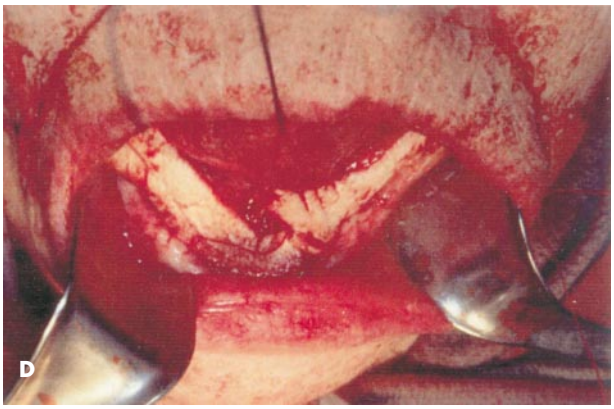


FIGURE 41–24. *A*, Nasolabial flaps outlined on the right and elevated on the left. Reproduced with permission from Atkins JP, Keane WM. Repair of the floor of the mouth with local flaps. In: Snow JB, editor. Controversy in otolaryngology. Philadelphia: WB Saunders; 1980. *B*, Intraoperative photograph of nasolabial flap harvest. *C,* Nasolabial flaps interdigitated and sutured in place. Reproduced with permission from Atkins JP, Keane WM. Repair of the floor of the mouth with local flaps. In: Snow JB, editor. Controversy in otolaryngology. Philadelphia: WB Saunders; 1980. *D*, Intraoperative photograph of the nasolabial flaps being placed onto the lingual surface of the mandible.

The ability to provide sensation to free flaps is a definite advantage over pedicled tissue.

In total glossectomy, adequate bulk is critical to allow contact of the flap with the hard palate. The latissimus dorsi and rectus abdominis free flaps are the best reconstructive alternatives for reconstruction of complete defects of the floor of the mouth and tongue. Depending on the patient's body habitus, a pectoralis major myocutaneous flap may also be appropriate.

Composite resections of the oral cavity, which include the mandible, may be reconstructed with the scapular, fibular, and iliac crest composite grafts. These problems are discussed in more detail in Chapter 40.

Free flap reconstruction of the mandible or oral cavity can be accomplished using the ipsilateral branches of the external carotid. If these are unavailable, the branches of the thyrocervical trunk, subclavian, or contralateral external carotid vessels are alternative options.[62]

Pharyngoesophagus

When more than 50% of the hypopharynx is resected, primary closure will lead to stenosis. Partial pharyngoesophageal defects are best reconstructed with either a pedicled pectoralis major myocutaneous flap or a radial forearm free flap.

Circumferential (total) defects require more options depending on the inferior and superior extent of the defect.

The simplest option is the tubed pectoralis major myocutaneous flap. This would be appropriate in patients with significant medical comorbidities. Difficulties with this choice include excessive bulk and multiple suture lines. A radial forearm free flap is the best choice for smaller circumferential defects.

When the resection extends into the cervical esophagus, the gastric pull-up is the procedure of choice. In these cases, skip lesions are a concern, and a total esophagectomy is oncologically more sound. The gastric pull-up allows primary closure with early deglutition. In some cases, this flap may not provide adequate length to reach the nasopharynx, resulting in excessive tension and subsequent wound dehiscence.

For defects that extend into the nasopharynx, free flaps such as the jejunal, radial forearm, and lateral thigh should be considered, and in obese patients, the jejunum is the flap of choice.[42,63,64] For hypopharyngeal and cervical esophageal reconstruction, the branches of the thyrocervical trunk, especially the transverse cervical vessels, are ideal for vascular reanastomosis.[38] Branches of the subclavian artery also are excellent options.[62]

Skull Base

Prior to the introduction of the microvascular free flaps, reconstructive options for skull base defects were particularly limited. These defects often required a large area of tissue that was not readily available. Pedicled flaps provided bulk, but their shorter pedicle length often failed to reach all limits of the defects and posed a high risk for partial flap necrosis. The rectus abdominis, with its bulk and long vascular pedicle, has been the flap of choice for large, laterally situated skull base defects. Although the rectus abdominis is considered the workhorse of skull base reconstruction, other flaps, such as the latissimus dorsi and omental free flaps, provide a thinner and more pliable option. The primary recipient vessels for scalp and skull base reconstruction are the superficial temporal vessels, with the posterior auricular, occipital, and transverse facial arteries as second choices.[62] Other vessels in the neck may also be used but may require an interposition vein graft.

REFERENCES

1. Cook TA, Guida RA. Soft tissue techniques. In: Papel ID, Nachlas NE, editors. Facial plastics and reconstructive surgery. St. Louis: Mosby Year Book; 1992. p. 22–44.
2. Petruzzelli GJ, Johnson JT. Skin grafts. Otolaryngol Clin North Am 1994;27:25–37.
3. Bakamjian VY. A two-stage method for pharyngoesophageal reconstruction with a primary pectoral skin flap. Plast Reconstr Surg 1965;36:173–84.
4. Ariyan S. The pectoralis major myocutaneous flap. Plast Reconstr Surg 1979;63:73–81.
5. Ariyan S, Ross DA, Sasaki CT. Reconstruction of the head and neck. Surg Oncol Clin North Am 1997;6:1–43.
6. Ijsselstein CB, Hovius SE, ten Have BL, et al. Is the pectoralis myocutaneous flap in intraoral and oropharyngeal reconstruction outdated? Am J Surg 1996;172:259–62.
7. Urken ML, Biller HF. Pectoralis flap. In: Urken ML, Cheney ML, Sullivan MJ, Biller HF, editors. Atlas of regional and free flaps for head and neck reconstruction. New York: Raven Press; 1995. p. 3–28.
8. Conley J. Use of composite flaps containing bone for major repair of the head and neck. Plast Reconstr Surg 1972;49:522–6.
9. Aviv JE, Urken ML, Lawson W, Biller HF. The superior trapezius myocutaneous flap in head and neck reconstruction. Arch Otolaryngol Head Neck Surg 1992;118:702–6.
10. Panje WR. The island (lateral) trapezius flap. Presented at the International Symposium of Plastic and Reconstructive Surgery, New Orleans, LA, April 29 to May 4, 1978.
11. Demergasso F. The lateral trapezius flap. Presented at the International Symposium of Plastic and Reconstructive Surgery, New Orleans, LA, April 29 to May 4, 1978.
12. Baek SM, Biller HF, Krespi YP, Lawson W. The lower trapezius island myocutaneous flap. Ann Plast Surg 1980;5:108–14.

13. Rosen HM. The extended trapezius musculocutaneous flap for cranio-orbital facial reconstruction. Plast Reconstr Surg 1985;75:318–24.
14. Cummings CW, Eisele DW, Coltrera MD. Lower trapezius myocutaneous island flap. Arch Otolaryngol Head Neck Surg 1989;115:1181–5.
15. Jianu J. Paralizie faciale dupa exterparea unei tumori a parotidei, tratata prin operatia Dlui Gomoiu. Bull Mem Soc Chir Bucharest 1908;22.
16. Kierner AC, Aigner M, Zelenka I, et al. The blood supply of the sternocleidomastoid muscle and its clinical implications. Arch Surg 1999;134:144–7.
17. Tiwari R. Experiences with the sternocleidomastoid muscle and myocutaneous flaps. J Laryngol Otol 1990;104:315–21.
18. Sebastian P, Cherian T, Ahamed MI, et al. The sternomastoid island myocutaneous flap for oral cancer reconstruction. Arch Otolaryngol Head Neck Surg 1994;120:629–32.
19. Kingdom TT, Singer MI. Enhanced reliability and renewed applications of the deltopectoral flap in head and neck reconstruction. Laryngoscope 1996;106:1230–3.
20. Hamaker RC. Oral cavity and oropharyngeal reconstruction. In: Cummings CW, et al, editors. Otolaryngology-head and neck surgery. St. Louis: Mosby-Year Book; 1986.
21. Verneuil AAS. De la creation d'une fausse articulation par section ou resection patielle de l'os maxillaire inferieur. Arch Gen Med 1872;15:284.
22. Clauser L, Curioni C, Spanio S. The use of the temporalis muscle flap in facial and craniofacial reconstructive surgery. A review of 182 cases. J Craniomaxillofac Surg 1995;23:203–14.
23. Shindo ML, Sullivan MJ. Muscular and myocutaneous pedicled flaps. Otolaryngol Clin North Am 1994;27:161–72.
24. Allegretti JP, Panje WR. The temporoparietal fascial flap in head and neck reconstruction: the workhorse from above. In: Friedman M, editor. Flaps for head and neck reconstruction. Part 1. Oper Tech Otolaryngol Head Neck Surg 2000;11: 143–6.
25. Verschuur HP, Dassonville O, Santini J, et al. Complications of the myocutaneous platysma flap in intraoral reconstruction. Head Neck 1998;20: 623–9.
26. McGuirt WF, Matthews BL, Brody JA, et al. Platysma myocutaneous flap: caveats reexamined. Laryngoscope 1991;101:1238–44.
27. Cannon CR, Johns ME, Atkins JP, et al. Reconstruction of the oral cavity using the platysma myocutaneous flap. Arch Otolaryngol 1982;108:491–4.
28. Persky MS, Kaufman D, Cohen NL. Platysma myocutaneous flap for intraoral defects. Arch Otolaryngol 1983;109:463–4.
29. Quillen CG, Shearin JC, Geogiade NG. Use of the latissimus dorsi myocutaneous island flap for reconstruction in the head and neck area. Plast Reconstr Surg 1978;62:113–7.
30. Haughey BH, Fredrickson JM. The latissimus dorsi donor site. Otolaryngol Head Neck Surg 1991;117: 1129–34.
31. Har-el G, Bhaya M, Sundaram K. Latissimus dorsi myocutaneous flap for secondary head and neck reconstruction. Am J Otolaryngol 1999;20:287–93.
32. Surkin MI, Lawson W, Biller HF. Analysis of the methods of pharyngoesophageal reconstruction. Head Neck 1984;6:953–70.
33. Lam KH, Wong J, Lim ST, Ong GB. Pharyngogastric anastomosis following pharyngolaryngoesophagectomy. Analysis of 157 cases. World J Surg 1981;5: 509–16.
34. Blackwell KE. Unsurpassed reliability of free flaps for head and neck reconstruction. Arch Otolaryngol Head Neck Surg 1999;125:295–9.
35. Urken ML, Weinberg H, Buchbinder D, et al. Microvascular free flaps in head and neck reconstruction. Arch Otolaryngol Head Neck Surg 1994; 120:633–40.
36. Schusterman MA, Miller MJ, Reece GP, et al. A single center's experience with 308 free flaps for repair of head and neck cancer defects. Plast Reconstr Surg 1994;93:472–8.
37. Shestak KC, Jones NF. Microsurgical free-tissue transfer in the elderly patient. Plast Reconstr Surg 1991;88:259–63.
38. Urken ML. Recipient vessel selection in free tissue transfer to the head and neck. In: Urken ML, Cheney ML, Sullivan MJ, Biller HF, editors. Atlas of regional and free flaps for head and neck reconstruction. New York: Raven Press; 1995. p. 331–7.
39. Tsue TT, Desyatnikova SS, Deleyiannis FW, et al. Comparison of cost and function in reconstruction of the posterior oral cavity and oropharynx. Arch Otolaryngol Head Neck Surg 1997;123:731–7.
40. Kroll SS, Evans GR, Goldberg D, et al. A comparison of resource costs for head and neck reconstruction with free and pectoralis flaps. Plast Reconstr Surg 1996;99:1282–6.

41. Lutz BS, Wei F, Chang SC, et al. Donor site morbidity after suprafascial elevation of the radial forearm flap: a prospective study in 95 consecutive cases. Plast Reconstr Surg 1999;103:132–7.
42. Hayden RE, Deschler DG. Lateral thigh free flap for head and neck reconstruction. Laryngoscope 1999; 109:1490–4.
43. Cordeiro PG, Santamaria E. The extended, pedicled rectus abdominis free tissue transfer for head and neck reconstruction. Ann Plast Surg 1997;39:53–9.
44. Taylor GI, Palmer JH. The vascular territories (angiosomes) of the body: experimental and clinical applications. Br J Plast Surg 1987;40:113–31.
45. O'Leary MJ, Martin PJ, Hayden RE. The neurocutaneous free fibula flap in mandibular reconstruction. Otolaryngol Clin North Am 1994;27:1081–96.
46. Urken ML, Sullivan MJ. Fibular osteocutaneous free flap. In: Urken ML, Cheney ML, Sullivan MJ, Biller HF, editors. Atlas of regional and free flaps for head and neck reconstruction. New York: Raven Press; 1995. p. 291–306.
47. Carroll WR, Esclamado R. Preoperative vascular imaging for the fibular osteocutaneous flap. Arch Otolaryngol Head Neck Surg 1996;122:708–12.
48. Turk JB, Vuillemin T, Raveh J. Revascularized bone grafts for craniofacial reconstruction. Otolaryngol Clin North Am 1994;27:955–82.
49. Hayden RE, Fredrickson JM. Microvascular surgery for head and neck reconstruction. In: Paparella MM, Shumrick DA, Gluckman JL, Meyerhoff WL, editors. Otolaryngology. 3rd ed. Vol IV. Philadelphia: WB Saunders; 1991. p. 2667–80.
50. Coleman SC, Burkey BB, Day TA, et al. Increasing use of the scapula osteocutaneous free flap. Laryngoscope 2000;110:1419–24.
51. Rosenberg MH, Sultan MR, Bessler M, Treat MR. Laparoscopic harvesting of jejunal free flaps. Ann Plast Surg 1995;34:250–3.
52. Saltz R. Endoscopic harvest of the omental and jejunal free flaps. Clin Plast Surg 1995;22:747–54.
53. Gherardini G, Gurlek A, Staley CA, et al. Laparoscopic harvesting of jejunal free flaps for esophageal reconstruction. Plast Reconstr Surg 1998;102: 473–7.
54. Kamei Y, Torii S, Hasegawa T, Nishizeki O. Endoscopic omental harvest. Plast Reconstr Surg 1998;102:2450–3.
55. Sullivan MJ, Urken ML. Free jejunal autograft. In: Urken ML, Cheney ML, Sullivan MJ, Biller HF, editors. Atlas of regional and free flaps for head and neck reconstruction. New York: Raven Press; 1995. p. 307–19.
56. Haughey BH. The jejunal free flap in oral cavity and pharyngeal reconstruction. J Laryngol Otol 1994;27: 680–2.
57. Yonekawa Y, Yasargil M. Brain vascularization by transplanted omentum: a possible treatment of cerebral ischemia. Neurosurgery 1977;1:256–9.
58. Vineberg A. Revascularization of the right and left coronary arterial systems: internal mammary artery implantation, epicardiectomy, and free omental graft operation. Am J Cardiol 1967;19:344–53.
59. Atkins JP, Keane WM. Repair of the floor of the mouth with local flaps. In: Snow JB, editor. Controversy in otolaryngology. Philadelphia: WB Saunders; 1980. p. 520–34.
60. Ross DA, Thomson JG, Restifo R, et al. The extended lateral arm free flap for head and neck reconstruction: the Yale experience. Laryngoscope 1996; 106:14–8.
61. Sullivan MJ, Carroll WR, Kuriloff DB. Lateral arm free flap in head and neck reconstruction. Arch Otolaryngol Head Neck Surg 1992;118:1095–101.
62. Takamatsu A, Harashina T, Inoue T. Selection of appropriate recipient vessels in difficult, microsurgical head and neck reconstruction. J Reconstr Microsurg 1996;12:499–507.
63. Shektman A, Silver C, Strauch B. A re-evaluation of hypopharyngeal reconstruction: pedicled flaps versus microvascular free flaps. Plast Reconstr Surg 1997;100:1691–6.
64. Schusterman MA, Shestak K, deVries EJ, et al. Reconstruction of the cervical esophagus: free jejunal transfer versus gastric pull-up. Plast Reconstr Surg 1990;85:16–21.

CHAPTER 42

Microtia, Canal Atresia, and Middle Ear Anomalies

Simon C. Parisier, MD, Jose N. Fayad, MD, Charles P. Kimmelman, MD

The child born with a malformed ear faces a lifelong hearing and communication impairment along with the social stigma of a facial deformity. Associated disturbances of the vestibular system may add the developmental hurdle of a motor delay. Frequently, there are additional anomalies, such as mandibular hypoplasia, as well as other facial and skeletal deformities. There may be dysfunction of associated neural pathways, including cranial nerves and intracranial structures. Additionally, there are psychological factors to be considered, including parental guilt, peer ridicule, and the shame of "being different." Educationally and economically, these hearing-impaired children face the prospect of limited opportunities. The appropriate management involves recognizing the problems and limitations of therapy, which need to be thoroughly understood by the parents and, when appropriate, the patient.

EMBRYOLOGY OF ATRESIA AND MICROTIA

In the 3 to 4 mm embryo (3 to 4 weeks), the first indications of aural ontogenesis are the first and second branchiomeric structures and the otic placode, an ectodermal thickening on the lateral surface of the head opposite the fourth ventricle. The placode invaginates to first form a pit and then a vesicle detached from its surface origin. This otocyst forms the inner ear membranous structures, with the endolymphatic duct developing first at the 6 mm stage, followed by the appearance of the semicircular ducts and the cochlear diverticulum at the 15 mm stage (6 weeks). By the end of the third month, the cochlea is fully coiled.

The cranial nerves entering the otocyst exert an inductive influence to produce neuroepithelium, for which retinoic acid is a potent morphogen. Retinoic acid receptors are uniquely expressed in the developing organ of Corti; medications that affect retinoic acid metabolism, such as isotretinoin (Accutane, Roche, Nutley, New Jersey), can lead to embryopathies, including inner ear malformation.[1]

The cochleovestibular ganglia develop from the otic placode epithelium. The nerve fibers themselves not only influence sensory cell development but are also directed to the developing inner ear by the sensory cells.[2]

As the membranous structures of the inner ear form, they become enveloped in a cartilaginous capsule, which eventually gives rise to the petrous portion of the temporal bone. Concurrently, the structures that originate from the first pharyngeal pouch develop separately, but adjacent to, the otic capsule derivatives. The pouch begins to form in the 3 to 4 mm embryo and expands into a tubotympanic recess, which will eventually give rise to the eustachian tube, middle ear space, and mastoid air cell system. The third branchial arch migrates superiorly to the level of the recess, and its artery (the internal carotid) comes to lie dorsal to the eustachian tube. Variations in this relationship may result in a lateralized displacement of the internal carotid artery into the middle ear space. In adults, an ectopic carotid artery can be mistaken for a middle ear mass, such as a glomus tumor.

As the pharyngeal pouches form in the 3 to 4 mm embryo, corresponding grooves develop on the external surface of the nascent cervical region. The first of these branchial clefts deepens until it approaches the tubotympanic recess, being separated only by the thin layer of mesoderm destined to become the middle fibrous layer of the tympanic membrane. Subsequently, in the 30 mm embryo (8 weeks), the primordial external canal becomes occluded by an ectodermal plug. By the twenty-first

week, this begins to resorb to form the definitive external auditory canal, replete with its hair and glandular appendages. Aberrations in the canalization process can lead to stenosis, canal tortuosity, or fibrous/osseous obliteration. Since middle ear structures develop independently, the tympanic cavity and ossicles may be normal.

Defects in the canalization process may also be associated with faulty formation of the pinna, which arises in the 8 to 11 mm embryo from six mesodermal thickenings. These hillocks surround the entrance of the first branchial cleft. The first branchial arch cartilage (Meckel's) forms the tragus and superior helical crus; the remainder of the pinna derives from the second arch cartilage (Reichert's), although some authorities posit a hyoid arch derivation for all but the tragus. The developing auricular appendage migrates from its initial position in the lower face toward the temporal area. This movement occurs along the fusion plane of the first and second branchial arches. The auricle is initially located anteriorly in a horizontal axis; with development of the branchial structures it migrates from its original position in the lower face laterally, and as its axis rotates, it assumes a more vertical angulation. Branchial cleft dysmorphogenesis can impede this migration and leave the pinna in a low, transverse orientation (Figure 42–1).

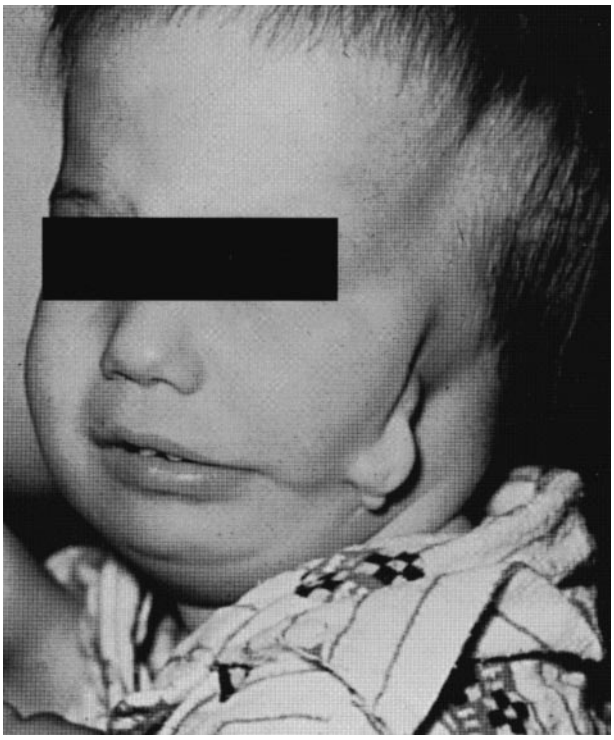


FIGURE 42–1. This child's auricle has not migrated from its embryonic low, transverse position.

As the middle ear forms, the separation between the first pharyngeal pouch and cleft is filled in by mesenchyme. In the 8 mm embryo (6 weeks), part of the connective tissue condenses to form the malleus handle; the subsequent expansion of the tympanic cavity superiorly is delayed until the cartilaginous otic capsule fully forms. The expansion of the first pharyngeal pouch results in the envelopment of the ossicles in an endodermal epithelium. The ossicles predominantly originate from mesenchymal visceral bars of the first and second arches. The first arch forms the head of the malleus and body of the incus, with the second arch giving rise to the manubrium, long process of the incus, stapes superstructure, and lateral portion of the footplate. The medial lamina of the footplate is derived from the otic capsule. The obturator foramen of the stapes forms around the stapedial artery, which usually remains diminutive while the stapes enlarges. If variations in vascular development cause enlargement of the artery, a conductive hearing loss may result from impairment of stapes motion. As the derivatives of the first pharyngeal pouch continue to extend into the developing temporal bone, the antrum, mastoid air cells, and petrous pyramid cells begin to form. Most mastoid development is postnatal; abnormalities that arrest middle ear formation result in a poorly pneumatized bone.

The facial nerve is intimately related to the development of the middle and inner ear structures outlined above. The blastema of the stapes is adjacent to the seventh nerve, which divides the second (hyoid) visceral mesenchymal bar into a laterohyale, stapes blastema, and interhyale. The interhyale forms the stapedius tendon, whereas the laterohyale forms part of the bony fallopian canal and pyramid. Thus, the development of the stapes is closely related to that of the facial nerve. Abnormalities in the development of the stapes are frequently associated with facial nerve anomalies. This relationship of facial nerve to developing middle ear structures increases the likelihood of an anomalous course of the nerve in the malformed middle ear.[3,4]

ETIOLOGY OF AURAL ATRESIA

Atresia and microtia are part of several known syndromes associated with inherited defects or acquired embryopathies owing to intrauterine infection (ru-

bella, syphilis), ischemic injury (hemifacial microsomia), or toxin exposure (thalidomide, isotretinoin).

Although inner ear abnormalities, such as Usher's syndrome, Waardenburg's syndrome, and the neurofibromatoses, are becoming understood on a molecular biologic basis, the genetic basis of external and middle ear anomalies generally remains poorly characterized.[5] Aural atresia occurs in approximately 1 in 20,000 live births. Although the inner and middle ears develop separately, inner ear abnormalities coexist in 12 to 50% of cases. Atresia is bilateral in 30% of cases, occurring more commonly in males and in the right ear.[6]

It is not surprising that an embryonic insult severe enough to cause aural atresia would also affect other organ systems. The following organs or systems may be anomalous in patients with atresia: neurocranium defects (Crouzon's disease or craniofacial dysostosis), central nervous system (mental retardation), oral cavity (first and second branchial arch syndromes), the eye (Goldenhar's syndrome), the neck (branchial fistulae), the CHARGE association (coloboma, heart defect, choanal atresia, retarded growth, genitourinary defects, and ear anomalies), Treacher Collins syndrome (mandibulofacial dysostosis), Duane's syndrome (abducens palsy with retracted globe), VATER complex (probable disorganization of the primitive streak with impairment of early mesodermal migration causing vertebral defects, anal atresia, tracheoesophageal fistula, renal defects, and genital anomalies), and Pierre Robin syndrome. Chromosomal anomalies affecting the external and middle ears include Turner's syndrome and trisomy 13 to 15, 18, 21, and 22 syndromes.[7]

Anomalies of the ear in the absence of syndromes are usually not familial. From the above information, it is obvious that other congenital anomalies should be assiduously sought; some, such as renal dysgenesis, may not be readily apparent. A chromosomal analysis may be indicated. As progress is made in the identification of genes and their products in the recently mapped human genome, the genetic basis for many of these disorders may provide a means of treatment or prevention.[8]

DIAGNOSIS AND EVALUATION

In the more severe cases of microtia, the diagnosis is apparent on inspection of the external ear. Depending on the degree of the abnormality, the microtic ear may be classified into three grades. In grade I, the auricle is developed and, though misshapen, has a readily recognizable, characteristic anatomy (Figure 42–2). In grade II, the helix is rudimentary and the lobule developed (Figure 42–3). In grade III, an amorphous skin tag is present[9] (Figure 42–4). In all stages, wide variations of morphology exist. The pinna may be fully formed with a transverse and low-set orientation. There may be accessory appendages of the pinna (pretragal tags with or without cartilage) and preauricular sinus tracts. The external canal may be stenotic or atretic to varying degrees[10] (Figure 42–5). In cases of stenosis, entrapped squamous epithelium may lead to a retention cholesteatoma with bone destruction.

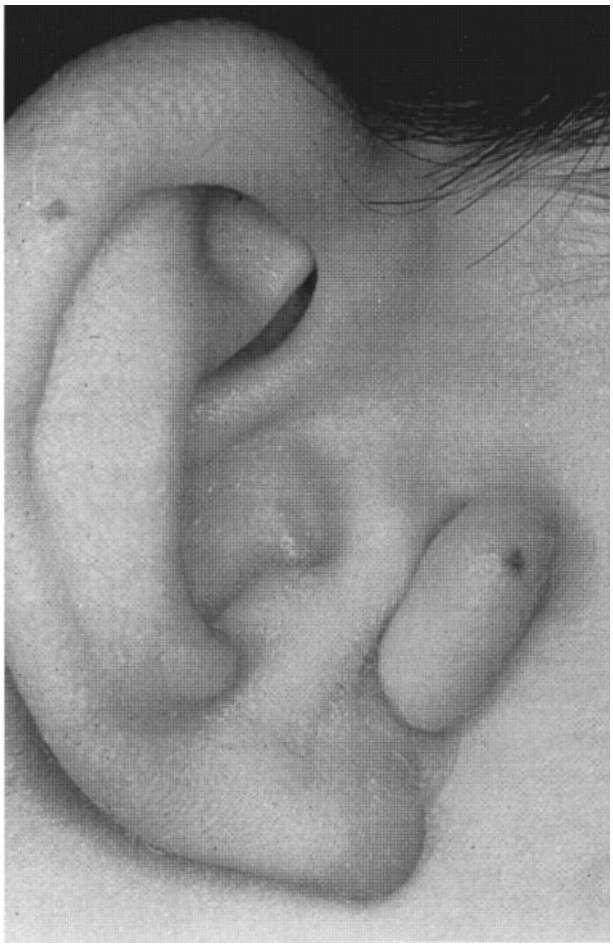


FIGURE 42–2. This pinna exhibits a grade I microtia with canal atresia. The size of the auricle and the characteristic anatomic landmarks are fairly normal. A pretragal skin tag is present.

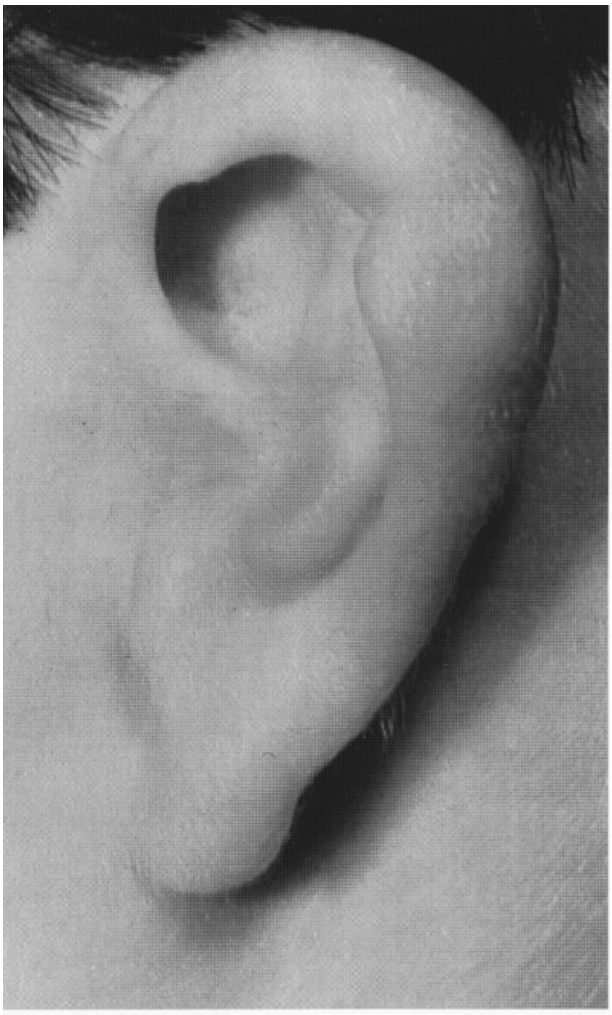


FIGURE 42–3. An example of a grade II microtia. The auricle is reduced in size and has a characteristic recognizable shape. The inferior and superior crura have not developed, although the helix is well preserved.

TABLE 42–1. Minor Congenital Middle Ear Abnormalities with Patent Ear Canal, Mobile Tympanic Membrane, and Conductive Loss

I. Ossicular Fixation
  A. Stapes Fixation
    1. Deficient annular ligament
    2. Elongated pyramidal process, ossified stapedial tendon
  B. Malleolar-Incudal Fixation
    1. Lateral epitympanic ankylosis
    2. Medial epitympanic ankylosis of body of incus, head of malleus
    3. Ossified anterior malleolar ligament
II. Ossicular Discontinuity
  A. Absent stapes arch
  B. Deficient lenticular process of incus
  C. Deficient long process of incus
III. Cochlear Capsule and Facial Nerve Anomalies
  A. Aplasia of oval or round window
  B. Facial nerve anomaly occluding oval window
  C. High jugular bulb occluding the round window niche

Schuknecht observed that in 7 ears with congenital meatal stenosis, all had cholesteatomas, whereas 3 of 11 ears with partial atresia and narrowed canals had cholesteatoma.[11] In 50 ears with complete atresia, only 2 had cholesteatomas.

Abnormalities of the tympanic cavity alone may occur with a normal tympanic membrane and external ear. In ears with conductive hearing losses and normal otoscopic examinations, isolated ossicular anomalies should be suspected (Table 42–1). Ossicular fixation can be produced by a variety of abnormalities and may involve the stapes, incus, or malleus. Alternatively, there may be a failure of bone development producing an ossicular discontinuity, which generally involves the incus and/or stapes arch. Ossicular malformations caused by abnormalities related to first or second branchial cartilaginous derivatives can frequently be surgically repaired (see Table 42–1, groups I and II). However, ossicular fixations owing to cochlear capsule abnormalities, especially when associated with an aberrant facial nerve, may not be surgically correctable (see Table 42–1, group III). Otic capsule abnormalities producing stapedial fixation are frequently associated with abnormal communication between the inner ear and the subarachnoid space. In these instances, manipulation of the stapes results in a persistent gusher of cerebrospinal fluid. Vascular malformations also occur in the middle ear, such as high jugular bulb, persistent stapedial artery, and anomalous course of the carotid artery.[12]

Further delineation of structural abnormalities requires computed tomography (CT).[13] Younger children may require sedation if they are unable to cooperate. The images are processed with a bone algorithm image enhancement using 1.5 mm slices at either 1.5 or 1 mm intervals. Optimally, both axial and coronal scans are obtained, although reformatted coronal or sagittal images may be adequate. Three-dimensional reconstruction[13,14] may be useful in visualizing the temporal bone topography in the

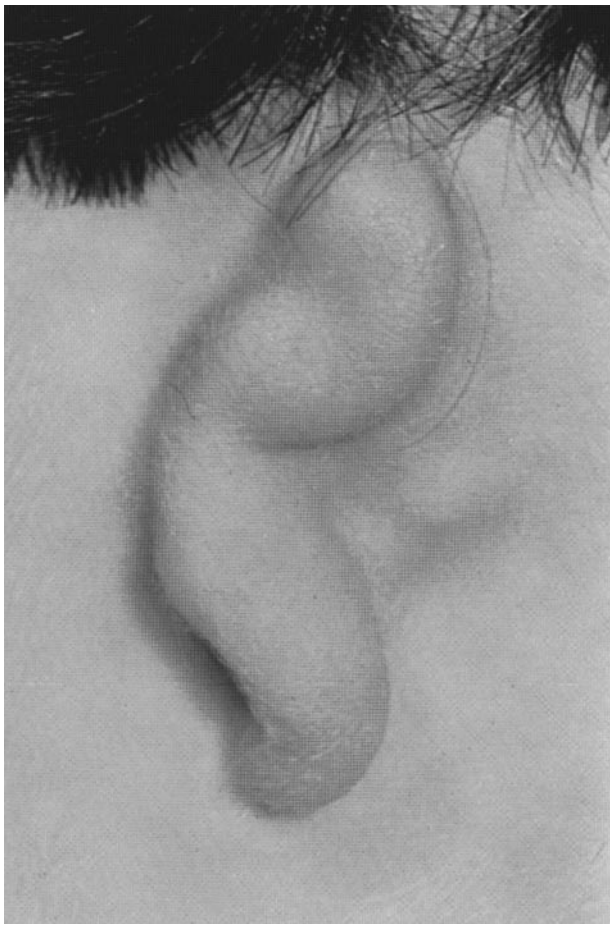


FIGURE 42–4. Auricle showing a grade III microtia. An amorphous ridge of skin and nubbins of cartilage are present in place of a recognizable, well-developed auricle.

more extensive malformations occurring in Treacher Collins syndrome and hemifacial microsomia. Computed tomography studies allow determination of the degree of the canal atresia, the thickness of the atresia plate (Figure 42–6, A), the extent of pneumatization of the middle ear and mastoid, the distance between the glenoid fossa and mastoid, the intratemporal course of the facial nerve, the status of the malleus and incus, and, in some instances, the presence or absence of the stapes and oval window (Figure 42–6, B). However, subtleties of oval and round window anatomy are often not revealed. An unsuspected cholesteatoma may also be apparent. The normalcy of the osseous inner ear structures is also revealed by CT.

Brainstem response audiometry should be performed neonatally in all children born with either microtia or atresia. In cases of microtia with fairly normal canals, an ossicular malformation producing a conductive hearing loss may be present. In patients with unilateral atresia, it is not unusual for the seemingly normal contralateral ear to have a hearing loss. Bilateral involvement may cause masking dilemmas; such is the case when one ear has a conductive loss and the other a sensorineural loss. These dilemmas may be minimized by using multichannel analysis of ipsilateral versus contralateral responses to determine the laterality of wave I.[15] As the infant matures, behavioral audiometric evaluations must be obtained to confirm the neurophysiologic tests of hearing. In children older than 1 year, conditioned free field play audiometry should help to quantify the overall hearing levels. Eventually, pure-tone and speech testing with masking should be obtained in children with fairly normal canals. Impedance and stapedial reflex measures in a seemingly normal ear can provide valuable information as well.

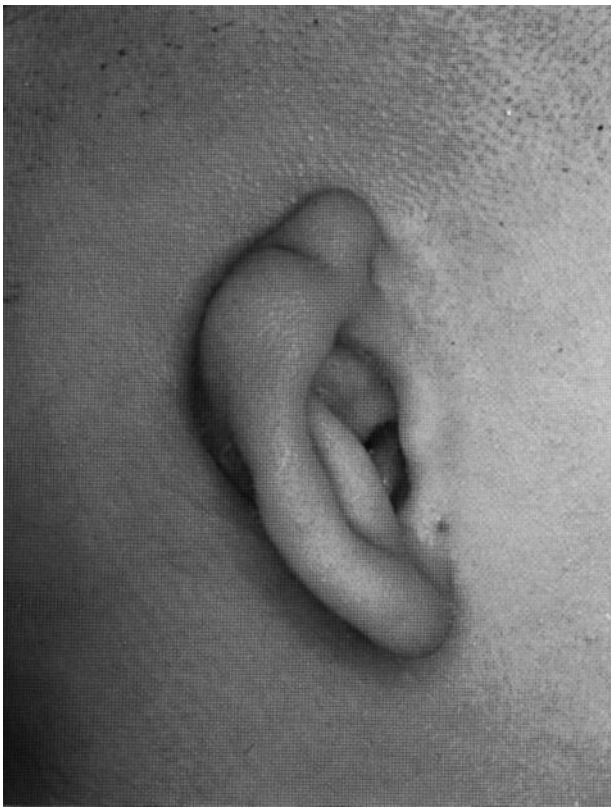
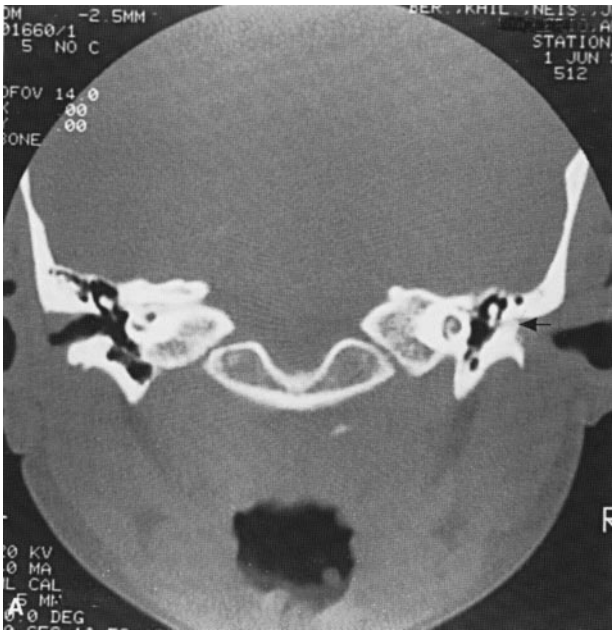


FIGURE 42–5. A grade II to III microtia with poor formation of the superior third of the auricle and a preauricular pit, possibly because of a first branchial arch dysmorphism. The lower portion of the ear, which is derived from the second branchial arch, has a relatively normal appearance. The canal is stenotic, leading to the subsequent development of a retention cholesteatoma.



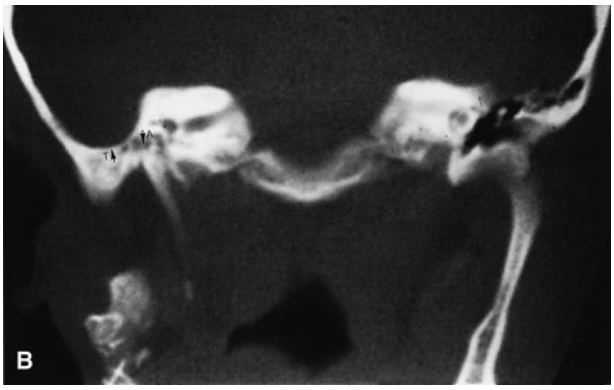


FIGURE 42–6. *A*, A coronal computed tomographic (CT) scan showing a bony atresia plate (*arrow*). The middle ear space is normally pneumatized. *B*, Coronal CT scan showing a low-lying tegmen (T), which precludes construction of an external canal. The middle ear cleft has not developed. A rudimentary antrum is unaerated (A). Atresia surgery is contraindicated in this case.

Evaluation of the function of other organ systems should be considered by the otologist. For example, some of the syndromal associations of atresia involve mandibular and laryngeal anomalies that can lead to airway obstruction; cardiac, renal, endocrine, and immune function should be ascertained in selected instances.

MANAGEMENT OF AURAL ATRESIA AND MICROTIA

Nonsurgical

The prime concern in the young child is the assessment and improvement of hearing. When ear deformities are present, the auditory brainstem response techniques outlined above are helpful in determining the type and severity of hearing loss and should be performed as early in life as possible. Early amplification, auditory training, and speech therapy can improve speech and language skills. In children with bilateral atresia, amplification with bone-conduction aids should be provided as soon as possible, preferably within the first few months of life.[16–18] In infants with a unilateral atresia and conductive hearing loss in the seemingly normal ear, an air-conduction aid should be fitted to the ear with a canal.

Surgical

In the unilateral case with normal contralateral hearing, repair of either the microtia or the atresia is considered elective, and there is less urgency to intervene during childhood. When the atresia is bilateral with acceptable-appearing auricles, reconstructive surgery can be performed at a fairly young age. However, in cases of microtia that require reconstruction, atresia repair should be deferred until the initial stages of the auricular repair are completed. Generally, microtia surgery requires a cartilage graft that is obtained from the lower costochondral region. An adequate graft requires sufficient growth and fusion at the donor site, which has usually occurred by 5 to 6 years of age. Additionally, in bilateral cases, many authorities recommend postponing surgery for the restoration of hearing until at least age 5 so that pneumatization can develop.[11] In children with frequent upper respiratory infections and suspected eustachian tube dysfunction, it may also be necessary to delay hearing reconstructive surgery.

Surgical Treatment of Congenital Conductive Hearing Loss The success of the surgical correction of congenital conductive hearing losses is related to the abnormality since a wide range of malformations is possible.[10] Surgery in an ear that has an isolated ossicular malformation with an otherwise

TABLE 42–2. Congenital Atresia Classification

Class I.	A. Aerated normal middle ear space
	B. Developed oval window with mobile stapes
	C. Oval window not obstructed by facial nerve
Class II.	A. Narrowed, but aerated, middle ear space
	B. Fixed stapes, oval window aplasia
	C. Oval window not obstructed by facial nerve
Class III.	A. Nonaerated, hypoplastic middle ear space
	B. Oval window obstructed by facial nerve
	C. Tegmen low hanging, obstructs access to middle ear space

patent canal, intact tympanic membrane, aerated middle ear cleft, normal facial nerve, and mobile stapes has the potential for excellent hearing (Table 42–2, class I). Conversely, an ear with canal atresia, a narrowed and poorly aerated middle ear space, and an anomalously positioned facial nerve occluding the oval window is destined to have poor postoperative hearing (see Table 42–2, class III). Between these two extremes is an array of anomalies with variable surgical outcomes. Even with sophisticated imaging techniques, the preoperative prediction of what awaits the otologist is not always accurate.[19]

Surgeons wishing to correct congenital conductive hearing losses are embarking on a procedure with significant potential complications. Generally, the initial attempt is the procedure most likely to improve the hearing. Unplanned revisions are often more complex and less successful. Additionally, if a significant sensorineural hearing loss occurs after surgery in a patient with bilateral malformations, the opposite ear is effectively excluded from surgical correction since it becomes the better hearing ear. Before assuming this responsibility, especially in children, the otologic surgeon must have sufficient experience to maximize the likelihood of a successful outcome.

The presence of a conductive hearing loss in an ear with a normal canal and mobile drumhead generally indicates that the ossicular chain is not transmitting sound energy to the cochlea. Isolated ossicular malformations involving the stapes arch and long process of the incus may not be appreciated on CT studies. In these cases, the middle ear can be explored by elevating a tympanomeatal flap and assessing the normalcy of the ossicular transduction mechanism. If the canal is small, a postauricular approach will improve operative exposure. Frequently, even the young child's ear canal has a sufficient diameter to admit an adequately sized ear speculum, even though the length of the canal is shorter.

Once the middle ear is entered, the surgeon needs to decide if the chain is fixed. If so, what is the cause of the fixation (see Table 42–1). First, the motility of the incus and malleus is ascertained. Is the fixation in the epitympanum and, if not, where (Figure 42–7)? Drilling away the bone over the malleus head and body of the incus provides the required exposure to explore this area. Is the stapedial tendon ossified? If the stapes is fixed, mobilization may produce a sustained improvement. Alternatively, a stapedectomy using a replacement prosthesis with an oval window tissue graft to prevent a perilymphatic leak may be performed. In some cases, the surgeon may elect to defer stapes surgery until the child is grown.

A gap interrupting the ossicular chain's continuity occurs most commonly at the incudostapedial joint, either owing to absence of the lenticular process and/or long process of the incus or the stapes arch. The presence of a mobile stapes facilitates the surgery and greatly improves the operative result. A variety of techniques have proven useful in the restoration of ossicular chain continuity by means of bridging an existing gap. If there is a small distance between the stapes capitulum and the long process of the incus, a tragal cartilage graft can successfully negotiate the gap. Deficiencies of the long process of the incus can be corrected by either interposing the reshaped incus between the mobile stapes and the malleus handle or using an alloplastic prosthesis designed for this purpose (type III tympanoplasty). When the stapes arch is absent, similar procedures can be performed to bridge the gap from either the malleus handle or the undersurface of the drumhead to the mobile footplate (type IV tympanoplasty). A coexisting fixation of the stapes or obstruction of the oval window by the facial nerve increases the technical difficulty, yields poorer hearing results, and increases the chance of complications, such as sensorineural hearing loss or facial nerve injury.[20]

Surgery for Microtia and Canal Stenosis or Atresia Congenital malformations of the auricle, external canal, and middle ear may occur as isolated abnormalities or in various combinations. Recon-

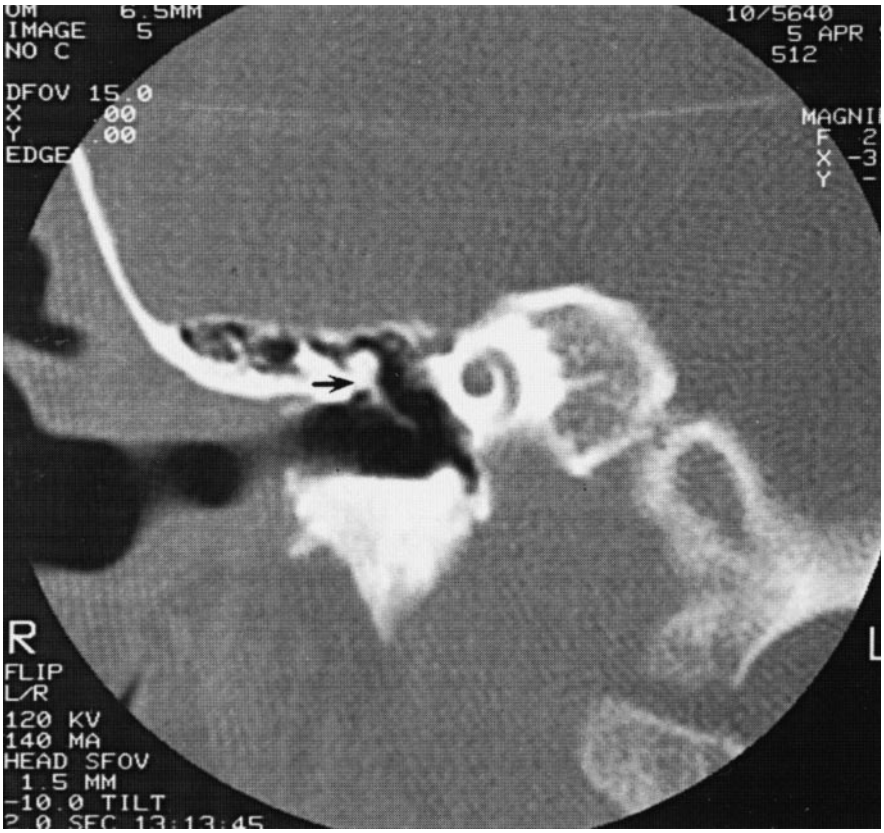


FIGURE 42–7. A coronal computed tomographic scan in a patient with a congenital conductive hearing loss caused by osseous fixation of the malleus head to the lateral epitympanic wall (*arrow*).

struction of the external ear is usually performed by a plastic surgeon, whereas external auditory canal and middle ear defects are corrected by the otologist; both work as a team to achieve the optimal result.

Microtia surgery is technically difficult, and not infrequently, the results are somewhat disappointing. The surgery should be performed by individuals with special expertise.[21] Auricular reconstruction is an elective procedure. The deformity can usually be masked by a longer hair style. Generally, the slight deformity of a grade I microtia may be cosmetically acceptable. Reconstruction of a moderately deformed grade II microtia must be individualized. Correction of a severe grade III microtia requires several staged procedures. During the first stage, a segment of cartilage is obtained from the lower costochondral skeleton. It is then carved to the appropriate shape to form a scaffold for the new auricle. This process requires an artistic appreciation of the external ear folds. The graft is implanted into a subcutaneous pocket at the appropriate position. A portion of the microtic skin tag is maintained to construct the lobule. The second stage is performed after several months. The auricle is freed from the scalp, and a split-thickness skin graft is applied to create a postauricular sulcus. At a third stage, several months later, the tragus is created. An alternative technique for the treatment of severe grade III microtia is to use a lifelike prosthetic ear, which is attached to surgically implanted titanium posts that securely hold the auricle in a normal anatomic position.

The timing of the microtia and atresia repairs requires close communication between the surgeons. Because scarring interferes with the auricular repair, reconstruction of the atresia is deferred until the auricular cartilaginous scaffolding is completed. However, it is important to position the reconstructed auricle correctly so that it aligns with the meatal opening and the middle ear. The meatal, canal, and middle ear reconstruction is frequently performed as part of the second or third stage of the auricular reconstruction.[13,22]

A postauricular approach is used, with great care being taken to avoid exposing the scaffold (Figure 42–8). Temporalis fascia is harvested for reconstruction of the tympanic membrane. The

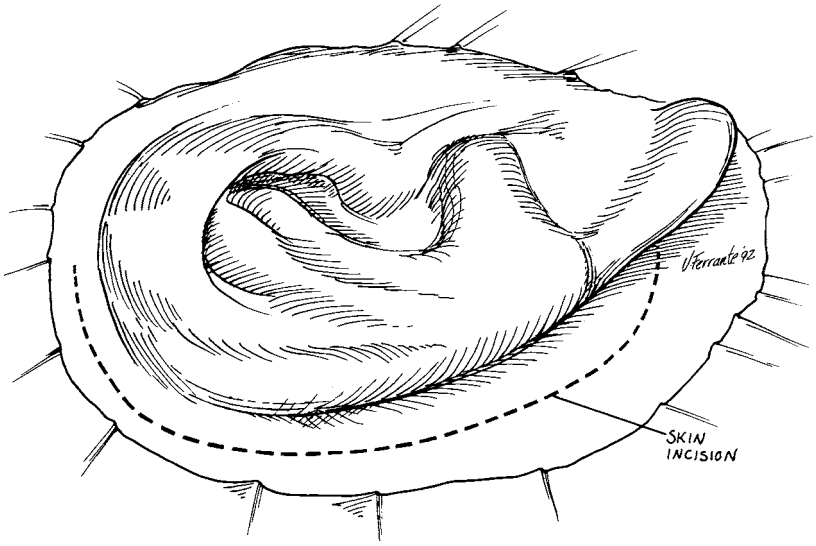


FIGURE 42–8. Surgery for canal atresia. A postauricular incision is made. While exposing the mastoid cortex and dysmorphic tympanic bone, the previously implanted cartilaginous auricular scaffold should be protected during retraction.

periosteum of the lateral surface of the temporal bone is elevated. A bony circular canal is created by drilling between the glenoid fossa anteriorly and the tegmen plate superiorly while trying to avoid entering mastoid air cells posteriorly (Figure 42–9). Continuous suction irrigation cools the bone and clears away blood and bone dust while the water enhances the translucency of the wet bone. This permits visualization of structures such as dura, facial nerve, and glenoid fossa through the intact thinned bone. The surgeon should anticipate that the facial nerve will be located more anteriorly and more laterally than in the normal temporal bone. The nerve frequently follows a C-shaped path in its course between the geniculate ganglion and its point of emergence from the temporal bone. The purposeful identification of the facial nerve will prevent inadvertent injury. A facial nerve monitor may facilitate identification of the nerve in atretic ears.

Visualization of the facial nerve as a landmark permits the opening of the atresia plate to approximate the size of a normal middle ear opening. While drilling on this plate, it must be remembered that the malleus handle may be fused to it. Precautions must be taken to avoid drilling on the mobilized atresia remnant since this may transmit traumatic vibratory energy through the ossicular chain to the inner ear and result in a permanent sensorineural hearing loss.

The status of the ossicular chain is assessed (Figure 42–10). If the stapes is mobile, an appropriate ossicular reconstruction is performed (Figure

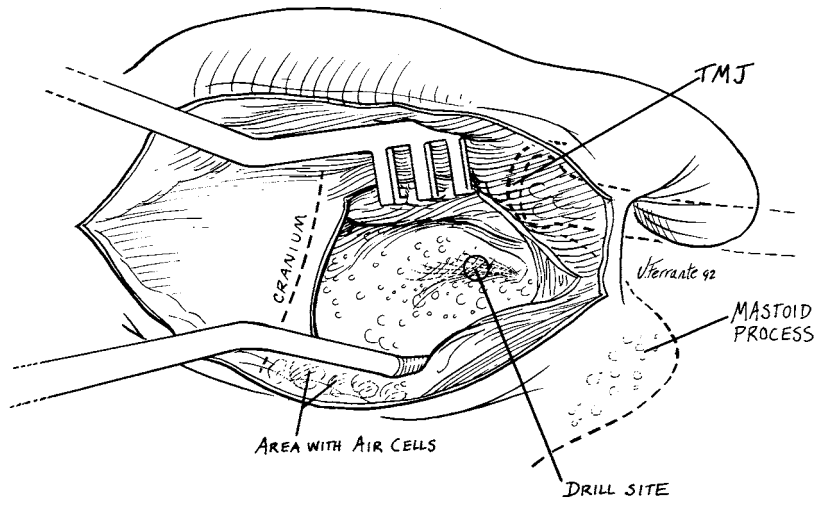


FIGURE 42–9. Construction of a canal is begun by drilling between the glenoid fossa anteriorly and the superior temporal line. The latter roughly corresponds to the tegmen (floor of the middle cranial fossa). When possible, entry into the mastoid air cells is avoided. TMJ = temporomandibular joint.

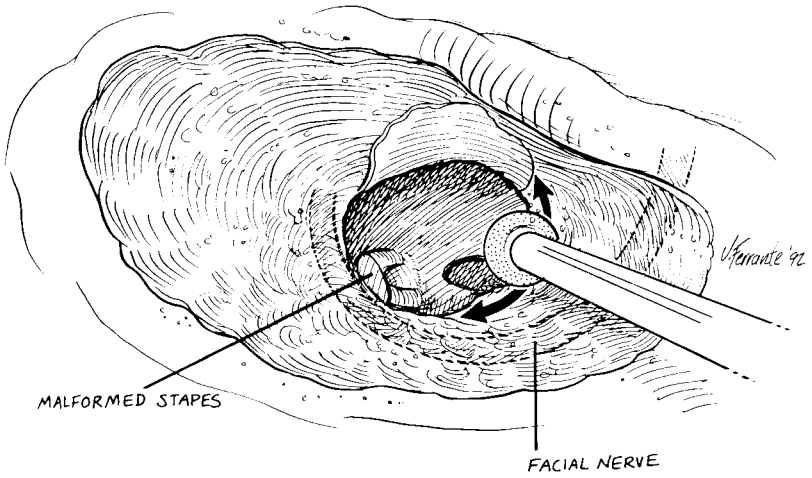


FIGURE 42–10. Drilling is continued until a normal-sized middle ear opening has been created. Great care is taken not to drill on a mobile, deformed, but intact ossicular chain. The anteriorly coursing facial nerve is identified.

42–11). If it is fixed or the oval window is obscured by the facial nerve, the ossicular reconstruction should be deferred. Temporalis fascia is used to fabricate a new tympanic membrane (Figure 42–12). A meatal opening is made in the area of the imperforate concha, incising the skin, to create a rectangularly shaped, anteriorly based conchal skin flap. The underlying cartilage and soft tissue are removed to debulk the flap and create an opening that communicates with the drilled out external canal (Figure 42–13). The conchal flap is rotated to resurface the anterior third of the external canal and is stabilized by suturing to the adjacent soft tissues (Figure 42–14). A thin split-thickness skin graft is harvested from the lower abdomen and is used to resurface the remaining ear canal. The canal is packed snugly with Gelfoam (Upjohn, Kalamazoo, Michigan) to compress the skin grafts against the underlying bone and soft tissue. The postauricular incision is reapproximated using absorbable sutures.

Risks and Benefits of Atresia Surgery Atresia surgery is technically demanding. The results achieved are directly related to the experience and skill of the operating surgeon. Generally, hearing improvement to a serviceable level can be achieved in approximately 65 to 75% of selected patients.[23,24] Complications of surgery include the small risk of facial paralysis, a severe to profound sensorineural hearing loss, stenosis requiring additional surgery, persistent otorrhea, and tympanic membrane graft failure with perforation.

In children with a unilateral atresia and a normal contralateral ear, surgery may not be routinely

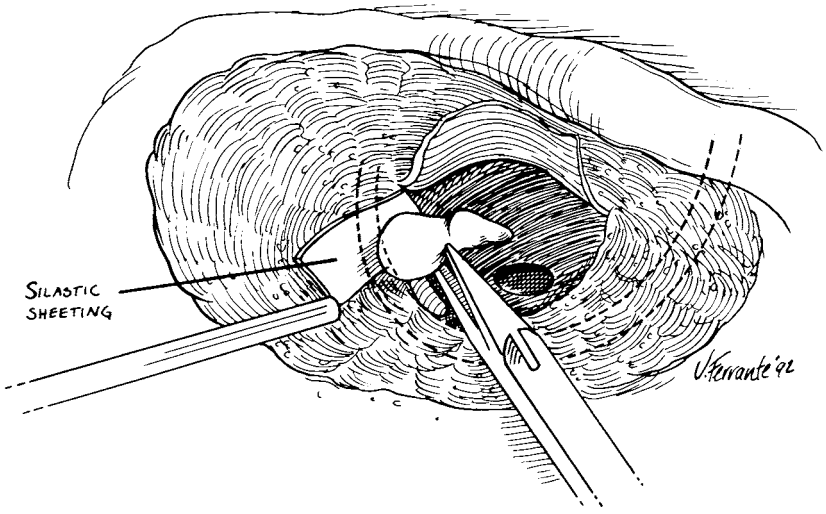


FIGURE 42–11. The ossicular chain is reconstructed by interposing the malformed, sculpted ossicle onto the stapes. Silastic sheeting, 0.005 inches thick, is positioned to prevent bony ankylosis.

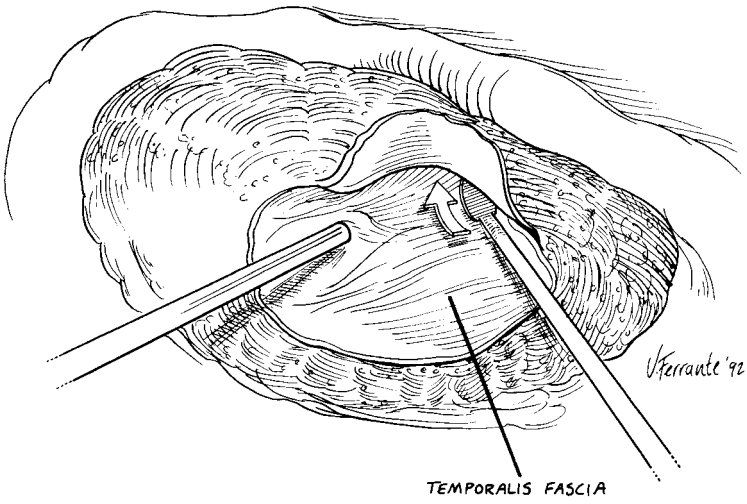


FIGURE 42–12. Temporalis fascia is used to construct a tympanic membrane.

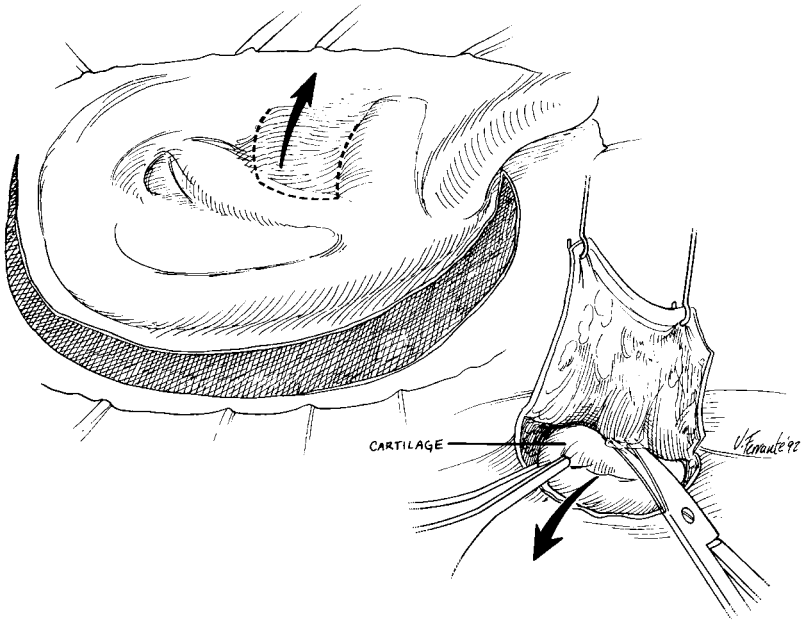


FIGURE 42–13. A meatal opening is made through the imperforate conchal skin. A rectangular, anteriorly based skin flap is debulked and rotated to line the anterior lateral third of the new canal.

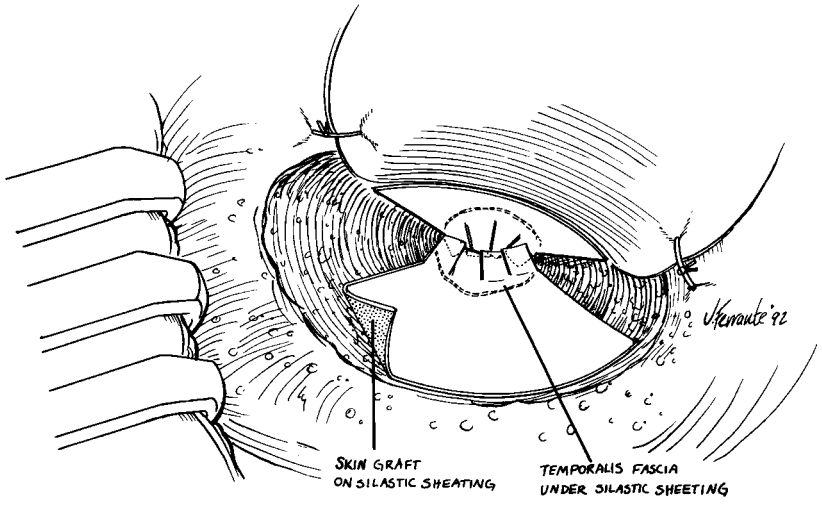


FIGURE 42–14. A thin, split-thickness skin graft harvested from the lower part of the abdomen is used to resurface the remainder of the canal. Handling and placement of the graft are facilitated by gluing the epidermal surface to thin (0.005 inch) Silastic sheeting. The raw surface of the graft is compressed against the bone by filing the canal lumen with Gelfoam.

indicated.[19] The potential benefit of this sophisticated surgery, that is, achieving binaural hearing, may not justify the risk of complications. In binaural cases, the successful creation of an external canal, even when hearing cannot be improved, will allow the use of an ear-level air-conduction hearing aid to restore hearing without using a cumbersome bone-conduction aid. Thus, in bilateral cases in which the facial nerve obstructs the oval window or when a hearing improvement cannot be achieved, providing the patient with a stable, skin-lined canal is a worthy goal in and of itself.

REFERENCES

1. Lefebvre PP, Malgrange B, Staecker H, et al. Retinoic acid stimulates regeneration of mammalian auditory hair cells. Science 1993;260:692–5.
2. Anson BJ, Davies J, Duckert LG. Embryology of the ear. In: Paparella MM, Shumrick DA, Gluckman JL, et al, editors. Otolaryngology. 4th ed. Philadelphia: JB Lippincott; 1991. p. 3–21.
3. Williams GH. Developmental anatomy of the ear. In: English GM, editor. Otolaryngology. Philadelphia: JB Lippincott; 1990.
4. Nager GT, Proctor B. Anatomic variations and anomalies involving the facial canal. Otolaryngol Clin North Am 1991;24:531–53.
5. Snow JB. Preface. Otolaryngol Clin North Am 1992;25:xv–i.
6. Jafek BW, Nager GT, Strife J. Congenital aural atresia: analysis of 311 cases. Trans Am Acad Ophthalmol Otolaryngol 1975;80:580–95.
7. Sando IS, Shibahara Y, Wood RP. Congenital anomalies of the external and middle ear. In: Bluestone CD, Stool SE, editors. Pediatric otolaryngology. Philadelphia: WB Saunders; 1990. p. 271–302.
8. Bergstrom LB. Anomalies of the ear. In: English GM, editor. Otolaryngology. Philadelphia: JB Lippincott; 1990.
9. Teunissen EB, Cremers CWRJ. Classification of congenital middle ear anomalies. Report on 144 ears. Ann Otol Rhinol Laryngol 1993;102:606–12.
10. Jahrsdoerfer RA, Yeakley JW, Aguilar EA, et al. Grading system for the selection of patients with congenital aural atresia. Am J Otol 1992;13:6–12.
11. Schuknecht HF. Congenital aural atresia. Laryngoscope 1989;99:908–17.
12. Curtin HD. Congential malformations of the ear. Otolaryngol Clin North Am 1988;21:317–36.
13. Jahrsdoerfer RA, Garcia ET, Yeakley JW, et al. Surface contour three-dimensional imaging in congenital aural atresia. Arch Otolaryngol Head Neck Surg 1993;119:95–9.
14. Andrews JC, Anzai Y, Mankovich NJ, et al. Three-dimensional CT scan reconstruction for the assessment of congenital aural atresia. Am J Otol 1992; 13:236–40.
15. Jahrsdoerfer RA, Yeakley JW, Hall JW, et al. High-resolution CT scanning and auditory brain stem response in congenital aural atresia: patient selection and surgical correlation. Otolaryngol Head Neck Surg 1985;93:292–8.
16. Granstrom G, Bergstrom K, Tjellstrom A. The bone-anchored hearing aid and bone-anchored epithesis for congenital ear malformations. Otolaryngol Head Neck Surg 1993;109:46–53.
17. van der Pouw KT, Snik AF, Cremers CW. Audiometric results of bilateral bone-anchored hearing aid application in patients with bilateral congenital aural atresia. Laryngoscope 1998;108:548–53.
18. Tjellstrom A, Hakansson B. The bone anchored hearing aid (BAHA) design principles, indications and long-term clinical results. Otolaryngol Clin North Am 1995;115:1–20.
19. Trigg DJ, Applebaum EL. Indications for the surgical repair of unilateral aural atresia in children. Am J Otol 1998;19:679–84.
20. Jahrsdoerfer RA, Lambert PR. Facial nerve injury in congenital aural atresia surgery. Am J Otol 1998;19: 283–7.
21. Chandrasekhar SS, De la Cruz A, Garrido E. Surgery of congenital aural atresia. Am J Otol 1995; 16:713–7.
22. Bellucci RJ. Congenital aural malformations: diagnosis and treatment. Otolaryngol Clin North Am 1981;14:95–124.
23. Shih L, Crabtree JA. Long-term surgical results for congenital aural atresia. Laryngoscope 1993;103: 1097–102.
24. Lambert PR. Congenital aural atresia: stability of surgical results. Laryngoscope 1998;108:1801–5.

CHAPTER 43

Anatomy and Physiology of the Oral Cavity and Pharynx

Louis D. Lowry, MD, Selcuk Onart, MD

The mouth or oral cavity is divided into two compartments: the vestibule or external compartment (vestibulum oris) and the internal compartment (cavum oris). The vestibule is the space external to the maxillary and mandibular alveolar ridges and teeth and within the lips and buccal mucosa. Internal to the alveolar ridges and teeth, the oral cavity proper is the space bounded by the hard and soft parts of the palate superiorly (ie, the roof of the mouth) and the lingual mucosa inferiorly (ie, the floor of the mouth and covering over the genioglossus, geniohyoid, and mylohyoid, which lie under the tongue). The anterior two-thirds of the tongue, bound posteriorly by the line of circumvallate papillae, are also part of the oral cavity proper.

The anterior pillars of the palatine tonsils separate the oral cavity from the oropharynx.

The pharynx is a fibromuscular tube that serves as a conduit for both the respiratory and digestive tracts. It is divided regionally into three segments: Superiorly, the nasopharynx is a posterior extension of the nasal cavities and ends at the level of the soft palate. The oropharynx is the region between the soft palate and the base of the tongue, down to the superior tip of the epiglottis, bounded laterally by the palatoglossal and palatopharyngeal arches. The laryngopharynx is posterior to the larynx and extends from the superior border of the epiglottis to the inferior aspect of the cricoid cartilage.

TONGUE

The mobile anterior two-thirds of the tongue are divided into four parts: tip, lateral borders, dorsum, and ventral surface. The dorsal surface faces the hard palate and the ventral surface faces the floor of the mouth. The lingual muscles are extrinsic and intrinsic. The extrinsic muscles are attached to the mandibular, styloid, and hyoid bones. They are the genioglossus, styloglossus, and hyoglossus and are able to move the tongue in various directions. The intrinsic muscles are the superior and inferior longitudinal, transverse, and vertical and, because of their orientation, impart great diversity to the movements of the tongue. The tongue is attached to the floor of the mouth ventrally and anteriorly and centrally by the frenulum.

The posterior third of the tongue is divided from the anterior two-thirds by the sulcus terminalis, which is a V-shaped groove, with its apex being posterior (Figure 43–1). The foramen cecum is at the apex of the sulcus and marks the origin of the thyroglossal duct. The lingual tonsil constitutes most of the posterior one-third of the tongue and varies in size. Anteriorly, the tongue is covered by a thin mucous membrane closely attached to the underly-

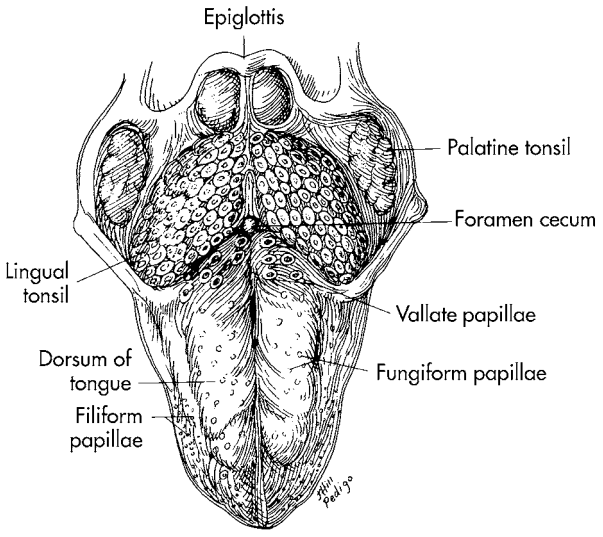


FIGURE 43–1. The dorsal surface of the tongue. Filiform papillae cover the dorsum. The lingual tonsil and palatine tonsils are part of Waldeyer's ring.

ing muscles. Posteriorly, the mucosa is thick and more freely moveable.

Innervation of the Tongue

The motor nerve to the tongue is the hypoglossal (cranial nerve [CN] XII), except for the palatoglossus muscle, which receives its innervation from the glossopharyngeal (CN IX) nerve. The hypoglossal nerve runs forward just above the greater cornu of the hyoid bone and has a course along the outer surface of the hyoglossus muscle, where it is accompanied by the lingual vein. The two major sensory nerves of the tongue are the lingual (CN V3) for the anterior two-thirds and the glossopharyngeal for the posterior one-third. The lingual nerve, a branch of mandibular division of the fifth cranial nerve the also receives the chorda tympani from the facial nerve (CN VII) (Table 43–1). The lingual nerve runs along the lateral border of the tongue above the submandibular gland and the hypoglossal nerve. The glossopharyngeal nerve passes around the posterior border and lateral surface of the stylopharyngeal muscle to reach the posterior part of the tongue deep to the hypoglossus muscle.

The trigeminal nerve (CN V) provides the general sensation to the anterior two-thirds of the tongue through the lingual nerve, whereas taste is subserved by CN VII (Figure 43–2). The glossopharyngeal nerve provides both taste and general sensation to the posterior third of the tongue and the circumvallate papillae.

Taste Buds

Taste buds in the human tongue, soft palate, pharynx, larynx, epiglottis, and esophagus are not distributed evenly. Most taste buds are found in the tongue, and in humans, there appear to be fewer extralingual taste buds than in laboratory animals.

TABLE 43–1. Relationship of the Branchial Arches and Their Nerves to the Sensation of the Tongue

Neural Branch	*Cranial Nerve*	*Branchial Arch*	*Sensory Modality Served*
Lingual nerve	V	First arch	Sensation to anterior two-thirds of tongue
Chorda tympani	VII	Second arch	Taste to anterior two-thirds of tongue
Glossopharyngeal	IX	Third arch	Taste and sensation to posterior one-third of tongue

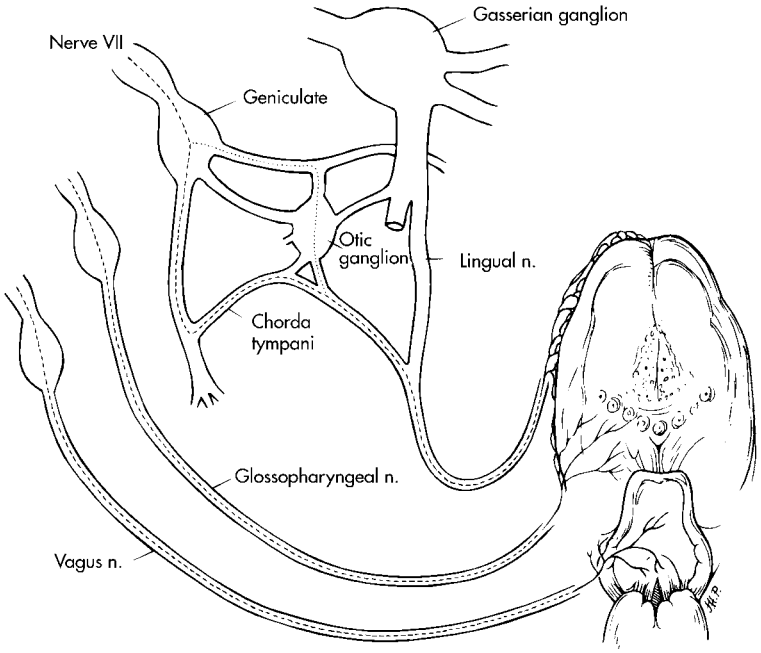


FIGURE 43–2. Diagram of the nerves serving the sense of taste on the tongue and surrounding regions.

In the tongue, taste buds are associated with three types of gustatory papillae, whereas extralingually, they are found on smooth epithelial surfaces. The number of taste buds varies among individuals, ranging between 500 and 20,000, with most in the 2,000 to 5,000 range.[1]

There are four types of lingual papillae: filiform, fungiform, foliate, and circumvallate. Filiform papillae are distributed throughout the dorsum of the tongue but do not have taste buds and function in the tactile aspects of feeding. Fungiform papillae are located on the anterior two-thirds of the tongue, number about 200, and usually contain one taste bud each.[2] They are shaped like mushrooms—hence their name, with the taste bud(s) found on their dorsal surface epithelium. The foliate papillae are found on the posterolateral sides of the tongue. There is only one foliate papilla on each side of the tongue, but each is a large structure arranged in a series of clefts that have grown beneath the tongue surface. The more rostral clefts, or the lateral rugae, contain no secretory ducts or taste buds. The more caudal clefts are called the foliate papillae folds and have ducts from lingual salivary glands (Ebner's glands) between the folds as well as a thinner epithelial lining that contains taste buds. Unlike taste buds on fungiform papillae, these taste buds are not found on the dorsal surface epithelium but instead face a cleft. In humans, there is an average of 600 taste buds in each foliate papilla, and each papilla is composed of between two and nine folds. The circumvallate papillae are found on the tongue just anterior to the sulcus terminalis. Humans have between 8 and 12 of them arranged in a V-shaped formation. Circumvallate papillae are large mushroom-shaped structures surrounded by a circular sulcus (vallum) and contain an average of 250 taste buds each. The taste buds are located along the walls of the sulcus, both on the papillae and on the adjacent epithelium. Secretions from Ebner's glands are released from ducts located at the base of the sulcus.[1]

The function of Ebner's glands has been under some debate recently. Traditionally, they were thought of as ancillary to taste buds and were mainly involved in simply washing the vallum, but some studies have suggested that they serve some additional purposes. One of these is the production of digestive enzymes, in particular lipase, which is important in the neonatal period when the pancreas is still immature. Also, it has been shown that Ebner's glands produce proteins that are involved in pheromone reception. For these functions to occur, it has been hypothesized that the glands must have access to detailed information about the chemical composition of the oral cavity. Supposedly, taste buds in the circumvallate papillae provide this information and help to control the activity of Ebner's glands.[3]

As stated above, extralingual taste buds are not found on papillae and are instead located within the epithelium. Extralingual taste buds tend not to be as well described as those on the tongue but still have some pattern of distribution that may be important functionally. Palatal taste buds are often located near the opening of ducts of mucus-secreting glands.[4]

Taste buds are made up of 50 to 150 neuroepithelial cells arranged in a compact gourd-shaped structure (Figure 43–3). Each receptor cell extends from the basement membrane of the epithelium to the epithelial surface, and the cells are arranged like sections of an orange. At the epithelial surface of each taste bud, there is a small opening called the taste pore, a channel that allows the microvilli of taste receptor cells access to gustatory stimuli. Tight junctions between taste receptor cells restrict the access of gustatory stimuli to the microvilli at the apices. There are generally four types of cells within each taste bud, and there is no firm agreement about the function or the origin of each type. The four types of cells are basal cells, type I (dark cells), type II (light cells), and type III (intermediate cells).[1]

Basal cells are small, undifferentiated, round cells near the basement membrane, from which most of the other taste bud cell types are thought to be derived. The other three types of cells are more elongated and bipolar, with microvilli at the apices and synapses with afferent fibers at the bases. These are thought to be the actual taste receptor cells. Over 50% of the taste receptor cells are type I, or dark cells. They have long microvilli and dense, membrane-bound granules near the cell apex. Type II, or light, cells have shorter microvilli, a lighter cytoplasm, and membranous vesicles near the apex. Type III, or intermediate, cells have characteristics of both type I and II cells.[1]

It is unclear whether the existence of four different cell types means that they have different functions or that they come from different cell lines. Some studies suggest that the different cell types are simply cells in different stages of development, with

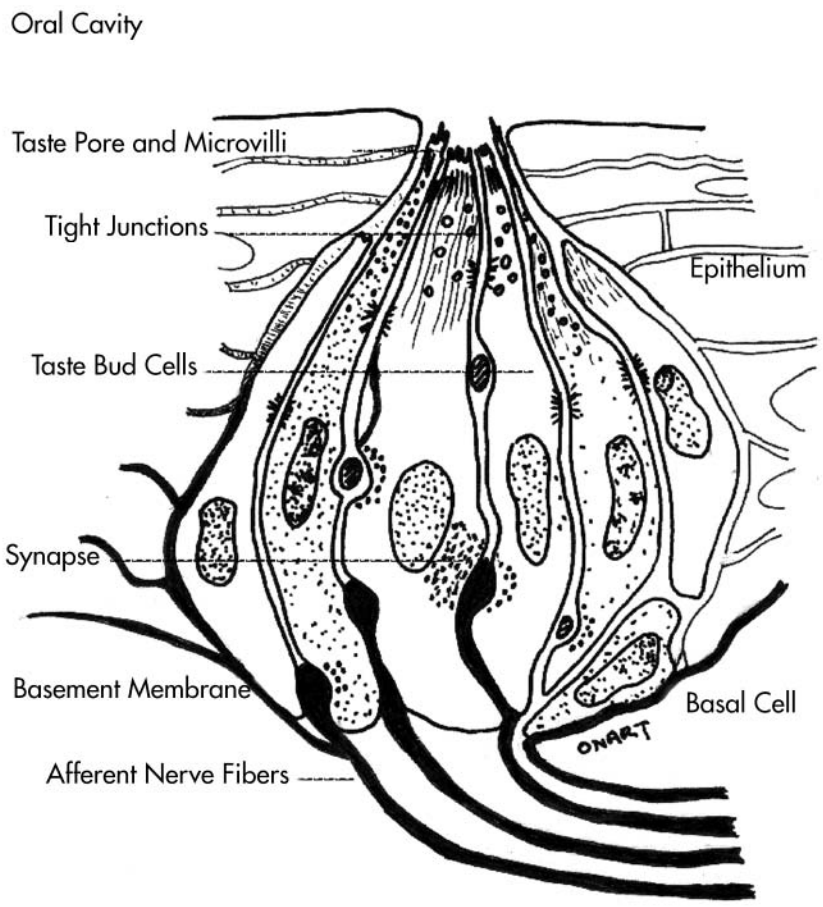


FIGURE 43–3. Taste bud demonstrating the relationship of the taste pore to the oral cavity. Afferent nerve fibers go to several different types of cells.

the basal cells turning into dark cells, which turn into intermediate cells, which turn into light cells. Other studies suggest that the light cells come from a separate cell line. Nevertheless, all three receptor cell types are known to synapse with afferent neurons, which make the arrangement complex when cells are turning over. The life span of an individual taste cell is only 10 to 14 days, and the cells are constantly being replaced, with the new cells constantly re-establishing connections with afferent nerve fibers to maintain taste sensation.[1]

Taste

The world of taste perception and the explanation of its mechanism of transduction is one convoluted at several different levels. Most importantly, taste should not be confused with flavor as the latter is actually a combined perception of gustation and olfaction. Additionally, one must understand that the study of taste relies on psychophysical perception and is thus prone to the error and complication of subjective experience. Therefore, it is often necessary to use the study of animal models, from which, for several reasons, the results are not always generalizable to humans. Further, it is necessary to qualify the concept of primary taste modalities. According to Brand, the criteria for the designation of a primary taste quality include

> 1) psychophysical and descriptive data that tend to isolate one primary taste from another on the basis of statistical criteria, 2) electrophysiologic evidence that reports unique neural transduction features of the putative taste modality, and 3) biochemical and molecular biological evidence that identifies and localizes unique receptors and cellular responses to the candidate primary modality.[5]

Generally, these criteria have been fulfilled for sweet, salty, sour, and bitter tastes, with evidence gradually proving for umami as well. Umami is the most recently discovered taste sensation that with continued research is slowly proving to be the fifth primary taste. Since the major prototypical stimulus for this taste has been shown to be the sodium salt of

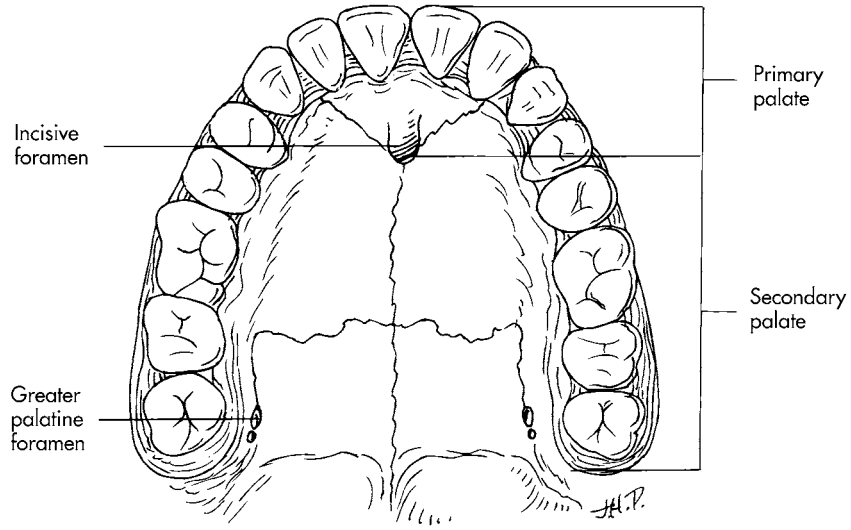


FIGURE 43–4. The hard palate and its divisions. Permanent teeth 1 through 16 are shown.

L-glutamic acid (monosodium glutamate), it is not surprising that umami means delicious in Japanese. It is a naturally occurring ingredient in many Asian foods. Given that the central nervous system harbors a multitude of receptors for glutamate, umami research has focused on the similarities and differences between receptors in the central nervous system and taste cells. For more information on the senses of smell and taste, see Chapter 27.

FLOOR OF THE MOUTH

The floor of the mouth contains the submandibular and sublingual salivary glands and the extrinsic musculature. It is supported by the paired mylohyoid muscles. The two muscles arise from the mandible and insert into a median raphe that extends posteriorly to the anterior surface of the hyoid bone. The muscles are so firm that swelling owing to infection above them can encroach on the airway.

BONES OF THE ORAL CAVITY

The *mandibles* form the skeleton of the lower jaw and the inferior part of the face They are the largest and strongest facial bones. Each mandible consists of two parts: the horizontal body and the vertical oblong ramus. The ramus ascends vertically from the posterior aspect of the body at the angle of the mandible. The superior part of the ramus has two processes: the posterior condylar process with a head, or condyle, and a neck and the anterior coronoid process. The condylar process is separated from the coronoid process by the mandibular notch, which forms the concave superior border of the mandible. On the internal aspect of the ramus is the large mandibular foramen, which is the entrance to the mandibular canal. The mandibular canal transmits the alveolar vessels and nerves. Branches of these vessels and the mental nerve emerge from the mental foramen, inferior to the second premolar tooth on the lateral side of each mandible. Inferiorly and slightly anteriorly from the mandibular foramen on the internal surface of the mandible runs the mylohyoid groove. This groove indicates the course taken by the mylohyoid nerve and vessels. Extending superiorly and posteriorly from the symphysis of the mandibular bodies is the mylohyoid line. Just superior to the anterior end of the mylohyoid line are two small sharp mental spines, which serve as attachments for the genioglossus muscle.

The palatine processes of the *maxillae* and the horizontal plates of the *palatine bones* form the anterior bony part of the palate (Figure 43–4). In hard palates of young people, a suture line, the palatine raphe, is visible between the premaxillary part of the maxilla and palatine processes of the maxillae. This represents the site of fusion of the median and lateral palatine processes during the twelfth week of prenatal development. In older people, this suture line is often absent. Contained within the hard palate are several canals. The incisive canal transmits the nasopalatine nerve and the terminal branch of the sphenopalatine artery. The incisive foramen is

located posterior to the maxillary central incisor teeth. Medial to the third molar tooth, the greater palatine foramen opens on the lateral border of the bony palate. It is the inferior orifice of the greater palatine canal. The greater palatine vessels and nerve emerge from this foramen and run anteriorly in two grooves on the palate. The lesser palatine foramina transmit the lesser palatine nerves and vessels to the soft palate and adjacent structures.

Fixed in the alveoli of the mandibular and maxillary bones are teeth. Although not bones, they will be described here for completeness. Deciduous teeth begin to develop before birth. The first tooth usually erupts at 6 to 8 months and the last by 24 months of age. The deciduous teeth are shed between 6 and 12 years of age and are replaced by permanent teeth. Eruption of permanent teeth is usually complete by 18 years of age, except for the third molars. If they are malposed or impacted, they may not erupt. There are 10 deciduous teeth in each jaw (2 medial incisors, 2 lateral incisors, 2 canines, 2 first molars, and 2 second molars). There are 16 permanent teeth in each jaw (2 medial incisors, 2 lateral incisors, 2 canines, 2 first premolars, 2 second premolars, 2 first molars, 2 second molars, and 2 third molars). The permanent teeth are numbered 1 to 32, starting with the right maxillary third molar as #1 and continuing to the mandible, with the left third molar as #17 and the right third mandibular molar as #32.

BLOOD SUPPLY TO THE ORAL CAVITY

The blood supply to the oral cavity is derived from branches of the external carotid artery (ECA). Most of the blood to this region comes from the ECA's first branches in the head, the lingual and external maxillary (facial) arteries, and their smaller branches. The ECA's two terminal branches, the internal maxillary and superficial temporal arteries, also contribute blood to the oral cavity. As in many regions of the head and neck, anastomoses between contributing arteries are important to the region's vascular integrity. Most of the blood drains to the jugular veins. The lingual artery, its branches, and veins of the same name provide circulation to the tongue and the floor of the mouth. Arising from the ECA opposite the greater horn of the hyoid bone, the lingual artery enters the tongue and supplies its muscles via its branching dorsal lingual artery. The lingual artery continues anteriorly and divides, sending the deep lingual artery superoanteriorly to supply the tip of the tongue and sending the sublingual branch to the sublingual gland and the floor of the mouth.

BLOOD SUPPLY OF THE PHARNYX AND TONSIL

The chief blood supply to the pharynx is derived from the ascending pharyngeal and tonsillar branches of the external maxillary artery and from the descending palatine branch of the internal maxillary artery (Figure 43–5). The tonsillar arteries, branches of the external maxillary artery, are the chief vessels to the tonsil, although the ascending palatine sometimes takes their place. The tonsillar arteries (or ascending pharyngeal artery) pass upward through the superior constrictor muscle, giving branches to the soft palate and tonsil. The dorsal lingual artery ascends to the base of the tongue and sends branches to the tonsil and tonsillar pillars. The descending palatine artery supplies the tonsil and soft palate from above, forming an anastomosis with the ascending palatine.

WALDEYER'S RING

Waldeyer's ring is a "ring" of lymphoid tissue, which provides the first contact of ingested or inhaled pathogens with the lymphoid system intimately related to epithelium. It is composed of the (1) ade-

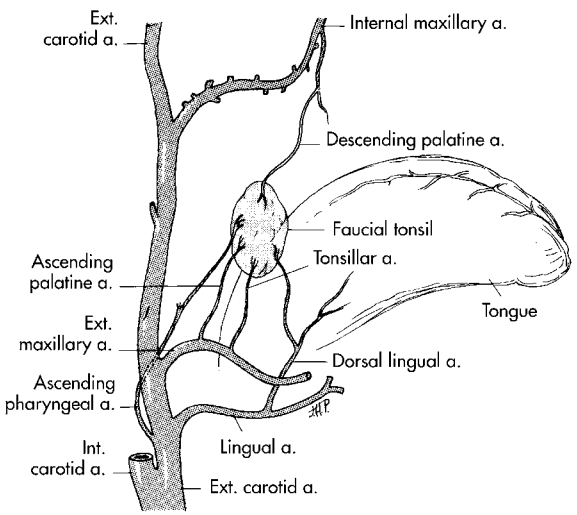


FIGURE 43–5. The blood supply of the palatine tonsil.

noid, (2) palatine tonsils, (3) lingual tonsil, (4) lateral pharyngeal bands, (5) scattered lymphoid follicles, and (6) nodules near the eustachian tube. This lymphoid tissue has only efferent lymphatics, much as Peyer's patches in the bowel, and is strategically located to sample everything that enters the aerodigestive tract. Waldeyer's ring is involved in the development of non–thymus-related lymphocytes or B cells, particularly in the first few years of life. Production of major classes of immunoglobulins and T lymphocytes with intact effector function of cell-mediated or "delayed" immunity can be attributed to elements of the ring.

Adenoid

The adenoid (also known as the pharyngeal or Luschka's tonsil) is a lobulated mass of lymphoid tissue found on the superior and posterior walls of the nasopharynx. The adenoid has no crypts but vertical folds lined by respiratory epithelium. It is composed of lymphoid tissue in a delicate reticulum of fibers and acts as a peripherally placed lymph node from which efferent lymph ducts pass to the cervical chain. The exposed surface of the adenoid is covered by stratified and pseudostratified epithelium. Groups of lymphoid cells are differentiated in the form of more or less rounded or oval areas with pale centers and darker margins. These lymphoid follicles or germ centers have the pseudonym germ centers of Goodsir. Unlike the palatine tonsil, the adenoid has no capsule. The incoming air from nasal breathing contacts the adenoid, and foreign substances initiate immune responses. The adenoids are capable of considerable hyperplasia and can obstruct the airway under adverse conditions (see Chapter 44).

In the midline of the nasopharynx, surrounded by the adenoid, is a depressed structure, the pharyngeal bursa, which represents the remnants of the notochord. Thornwaldt's disease is an infection of this bursa.

Palatine Tonsils

The palatine tonsils, also known as the faucial tonsils, are grape-like masses of lymphoid tissue lying between the palatoglossus muscle (anterior pillar) and the palatopharyngeus muscle (posterior pillar). The lateral surface of each tonsil is covered by pharyngeal fascia and attached to the superior pharyngeal constrictor muscle. Condensations of the fascia form a capsule. From the tonsillar capsule, trabeculae extend into the parenchyma of the tonsil and support blood vessels, nerves, and efferent lymphatic vessels. Contractions of the superior constrictor, palatoglossus, and palatopharyngeus (as in swallowing) cause compression of the tonsil. The free surface of the tonsil is covered by a closely adherent stratified squamous epithelium that extends into blind pouches or crypts. The epithelium lining the crypts is thin and, in fact, may be a semipermeable membrane that "lends" itself to sampling ingested material. The crypts, 8 to 10 in number, are usually compound, extend deep into the substance of the tonsil, and come into intimate contact with the lymphatic germinating follicles, suggesting permeability. With swelling of the tonsil, the bottom of the crypts remains relatively fixed; thus, the crypts become longer. The germinating follicles are centers in which mother cells of the leukocytic group form young lymphoid cells. The interfollicular tissue is made up of lymphoid cells in various stages of development. Above the tonsil and between the pillars is the supratonsillar fossa.

Lingual Tonsil

The lingual tonsil constitutes part of the base of the tongue and extends from the foramen cecum to the epiglottis. This sessile midline lymphoid accumulation is covered by stratified squamous epithelium. It is poorly separated from the tongue musculature by a layer of fibrous tissue. It consists of numerous rounded or crater-like elevations of lymphoid tissue in the center of which a duct of a mucous gland opens.

LYMPHATICS OF THE ORAL CAVITY

The lymphatic drainage of the oral cavity is composed of both superficial and deep vessels that ultimately drain into the submental, submandibular, upper jugular, and, occasionally, spinal accessory lymph nodes. Cancers arising in the buccal mucosa primarily drain into the preglandular lymph nodes of the submandibular triangle and secondarily drain into the upper jugular lymph nodes of the deep jugular chain. Contralateral metastases are infrequent.

The lymphatic network overlying the gingiva of maxillary and mandibular bones joins the lingual and buccal surfaces. The preglandular submandibu-

lar lymph nodes represent the primary drainage route for both systems.

The superficial lymphatic vessels of the floor of the mouth connect with the gingival system before draining to the preglandular submandibular lymph nodes (Figure 43–6). The superficial system located in the anterior part of the floor of the mouth crosses the midline and drains into either the ipsilateral or contralateral preglandular lymph nodes. The deep collecting system drains into the same preglandular lymph nodes before draining into the subdigastric lymph nodes of the internal jugular chain.

A second pathway bypasses the submandibular lymph nodes and drains into the jugular omohyoid lymph nodes. Additionally, an efferent upper jugular pathway crosses the midline and bypasses the submandibular lymph nodes (Figure 43–7).

The tongue possesses both a superficial and a deep lymphatic system. No midline separation exists, so crossover into contralateral lymphatics is not uncommon. Lymphatic capillaries in the anterior portion of the tongue tend to drain into the submental and midjugular lymph nodes, the lymphatics from the middle third to submandibular or digastric lymph nodes, and those from the base of the tongue to the upper jugular lymph nodes (Figure 43–8). The normal anatomic configuration of the lymph flow can be altered by obstruction caused by infection, neoplastic involvement, previous surgery, and irradiation.

Minor Salivary Glands

There are 600 to 1,000 minor salivary glands throughout the oral cavity and oropharynx, with the greatest concentration found in the palate. The hard palate has 250 minor salivary glands and the soft palate has 150, with their parasympathetic innervation coming from the sphenopalatine ganglion. The remaining minor salivary glands of the oral cavity and oropharynx receive their parasympathetic innervation from the lingual and glossopharyngeal nerves. Each gland has its own excretory duct emptying directly into the oral cavity or oropharynx. The vascular supply, venous outflow, and lymphatic drainage correspond to the region of the oral cavity and oropharynx where the glands are located.

PHYSIOLOGY

Saliva is the secretory product of the three paired major salivary glands and the 600 to 1,000 minor salivary glands. It has a myriad of functions, including aiding in digestion, gustation, protection against infection, dental caries and cancer, mastication, de-

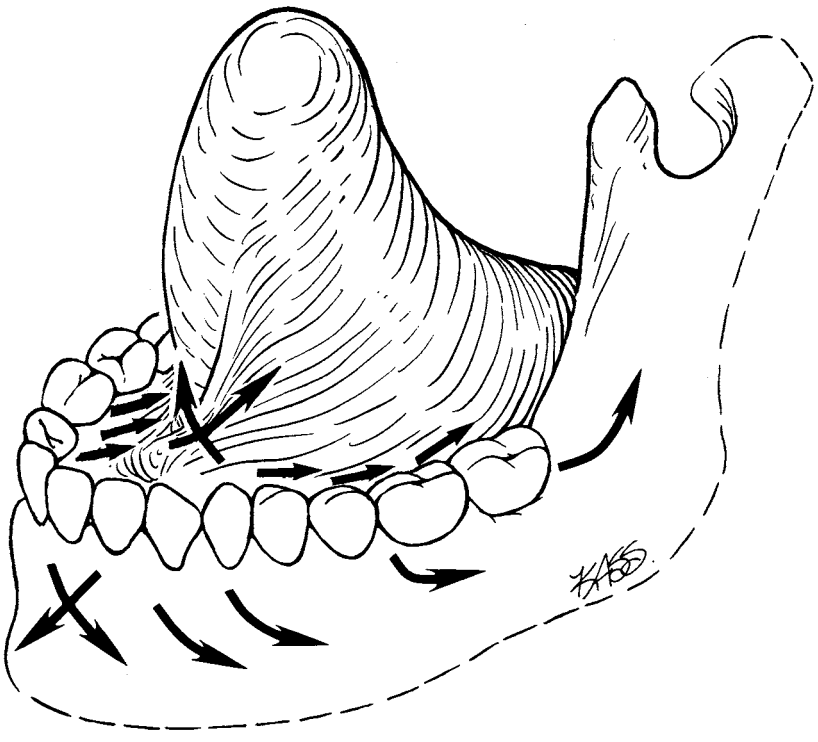


FIGURE 43–6. Drawing of the superficial lymphatic vessels of the floor of the mouth.

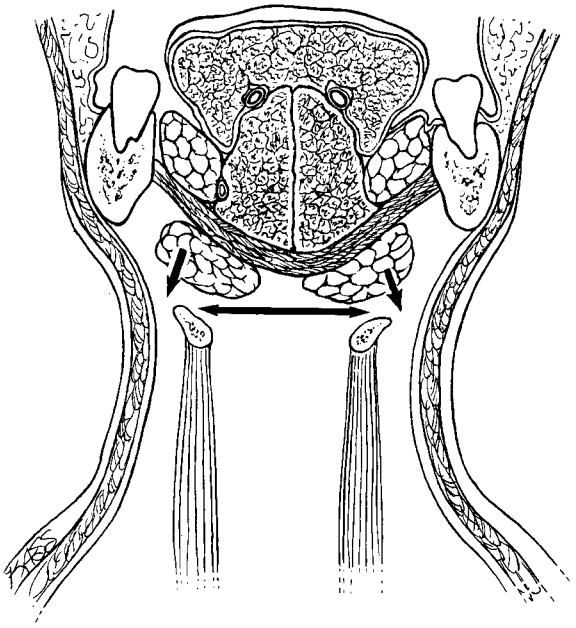


FIGURE 43–7. Drawing demonstrating the upper jugular efferent lymphatic communicating pathway when the submandibular group of lymph nodes has been bypassed.

glutition, lubrication, protective coating, and homeostasis, as well as hormonal interaction. Saliva has numerous components, including salts, gases, and organic substances, including immunoglobulin A, and its production is influenced by several factors, including psychological states, circadian rhythms, hydration, sleep, age, hormones, diet, and drugs. The actual production of saliva is a complex process involving both secretion and reabsorption. Saliva has many functions, many of which are intertwined. Saliva aids in digestion through lubrication that allows for the smooth passage of food. Saliva also allows for the dissolution of soluble food products and the active digestion of carbohydrates by amylase. The dissolution of soluble food products also enables their chemical interaction with taste receptors. The importance of salivary amylase has been debated since a lack of salivary amylase does not seem to precipitate nutritional deficits. Its activity, though, is not disputed, and it can be active in the stomach for several hours inside a food bolus.

Saliva serves an important protective role through numerous mechanisms. Saliva dilutes and dissolves retained food products, cellular debris, and bacteria. Glycoproteins in saliva coat the mucosal surface and serve as barriers to chemical irritants or desiccation. Lysozyme, peroxidase or thiocyanate-dependent factor, bacteriolysin, and immunoglobulin A all confer an antibacterial effect.

Saliva is composed of salts, namely sodium chloride, potassium chloride, sodium bicarbonate, bisodium and monosodium phosphates, calcium carbonate, calcium phosphate, and potassium sulfocyanate; gases, namely carbon dioxide, nitrogen, and oxygen; and organic substances, namely amylase, maltase, lysozyme, seralbumin, seroglobulin, urea, uric acid, creatine, amino acids, and mucin. The bicarbonates and phosphates serve to buffer the pH around 6.8. At alkaline pH levels, calcium salts precipitate and may form stones within the salivary ducts and cause the deposition of tartar (dental calculus).

Regulation of Saliva

One to 1½ L of saliva are produced daily. The rate of flow varies during the day but averages 1 mL/min. During sleep, practically no saliva is produced by the major salivary glands. Resting flow is estimated to be around 0.33 to 0.65 mL/min, which can increase to

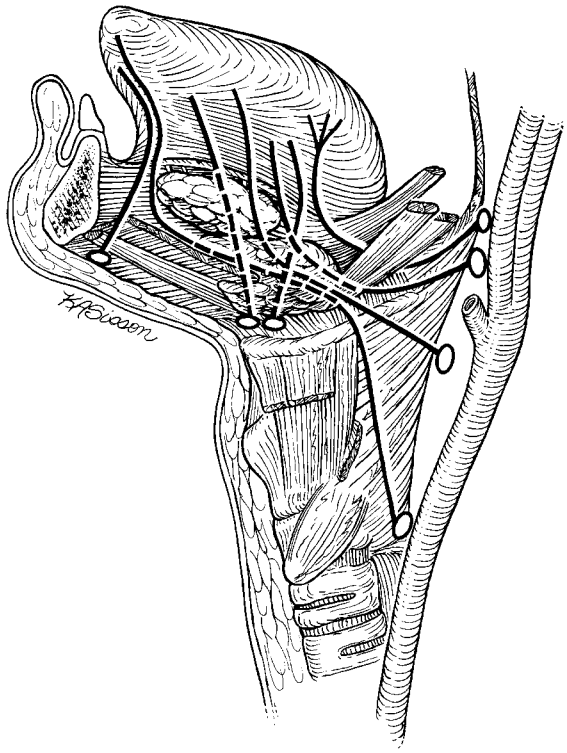


FIGURE 43–8. Lymphatic drainage routes from the tongue.

1.7 mL/min when stimulated. In the unstimulated state, saliva from the major salivary glands is produced predominantly by the submandibular glands (69%), followed by the parotid glands (26%) and the sublingual glands (5%). During stimulation, the parotid glands significantly increase saliva production and supplant the submandibular glands with 67% of saliva production. The minor salivary glands remain constant in their production of saliva at 7 to 8%.

Speech

The production of speech requires a source of air flow (lungs), a sound generator (vocal folds), and a resonator (pharynx and oral cavity; the nose is included in the resonating chamber when the palate is lowered). Articulators change the form of the vocal tract. The structures that affect articulation are the lips, tongue, mandible, teeth, palate, pharynx, and larynx; specifically, the gestures of these structures in articulation include opening, pursing, and tensing of the lips; position, configuration, and dynamics of the tip, margins, dorsum, and base of the tongue; opening and closure of the velopharyngeal valve by the palate and pharynx; movement of the side walls of the pharynx; and raising and lowering of the larynx. Articulators play the dominant role in producing speech sounds represented by consonants.

The dimensions of the vocal tract affect resonance. With resonance, formants or dominant frequencies emerge that can be altered or tuned by changing the shape of the vocal tract.[6–9] Formants are used to produce a voice quality unique to the individual. They determine vowel quality and provide a unique personal voice timbre. The two lowest formants determine most of the vowel color, and the third, fourth, and fifth determine the personal voice timbre (see Chapters 47 and 49).

Swallowing

Swallowing is a complex and complicated function and is divided into four overlapping phases: oral preparatory stage, oral stage, pharyngeal stage, and esophageal stage.[10] The first two are voluntary, and the third and fourth are involuntary. In the oral preparatory stage, the lips are sealed, and the cheeks tense to hold the muscles close to the teeth. The food is chewed while the tongue constantly repositions it to keep it between the teeth. The soft palate is in contact with the tongue, and the tongue holds the food in a ball ready to begin the oral stage of swallowing. In the oral stage of the swallow, the tongue and, to some extent, the buccal musculature move the food bolus from the front of the oral cavity to the anterior faucial arch, where the swallowing reflex is initiated.

The pharyngeal stage has four defined components: velopharyngeal closure; peristalsis of the superior, middle, and inferior pharyngeal constrictor muscles; airway protection; and cricopharyngeal relaxation. These are reflexive and are mediated in the brainstem in the reticular formation immediately adjacent to the respiratory center. There is coordination between the two centers. The larynx is elevated and closed (see Chapter 47), protecting the airway,[11] while the piriform sinuses contract and the cricopharyngeus muscle relaxes.[12] These reflex movements are in sequence and overlap one another.[13] This happens very rapidly as shown on videofluoroscopic analysis, where in a single frame (1/24 of a second), the barium from the piriform sinuses empties into the esophagus.

In the final esophageal stage, the bolus moves down the esophagus assisted by gravity, positive external pressure, peristalsis of the esophageal musculature, and negative intraesophageal pressure.

ACKNOWLEDGMENT

Dr. Lowry expresses appreciation to the following medical students who assisted in the development of this chapter: Esmael N. Amjad, Jennifer G. Andrus, Patrick C. Barth, Lisa Grunebaum, Brian Kung, Spencer Payne, and Joseph L. Smith.

REFERENCES

1. Linden RWA. Taste. Br Dent J 1993;175:243–52.
2. Slavkin HC. Reflections of taste for oral health professionals. J Am Dent Assoc 1997;128:1697–701.
3. Sbarbati A, Crescimanno C, Osculati F. The anatomy and functional role of the circumvallate papilla/von Ebner gland complex. Med Hypotheses 1999;53: 40–4.
4. Miller IJ, Bartoshuk LM. Taste perception, taste bud distribution, and spatial relationships. In: Getchell TV, Doty RL, Bartoshuk LM, Snow JB, editors. Smell

and taste in health and disease. New York: Raven Press; 1991. p. 205–33.
5. Brand JG. Receptor and transduction processes for umami taste. J Nutr 2000;130 Suppl 4:942–5.
6. Sundberg J. Vocal tract resonance in professional voice. In: Sataloff RD, editor. The science and art of clinical care. 2nd ed. San Diego (CA): Singular Publishing Group; 1997. p. 47–64.
7. Sundberg J, Lindblom B, Liljencrants J. Formant frequency estimates for abruptly changing area functions: a comparison between calculations and measurements. J Acoust Soc Am 1992;91:3478–82.
8. Dang J, Honda K. Acoustic characteristics of the piriform fossa in models and humans. J Acoust Soc Am 1997;101:456–65.
9. Baer T, Gore JC, Gracco LC, Nye PW. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. J Acoust Soc Am 1991;90:799–828.
10. Logeman JE. Mechanisms of normal and abnormal swallowing. In: Schuller DE, editor. Otolaryngology-head and neck surgery. 2nd ed. St. Louis: CV Mosby; 1986. p. 1741–9.
11. Donner MW, Bosma JF, Robertson DL. Anatomy and physiology of the pharynx. Gastrointest Radiol 1985;10:196–212.
12. Cook IJ. Cricopharyngeal function and dysfunction. Dysphagia 1993;8:244–51.
13. Palmer JB, Tanaka E, Siebens AA. Motions of the posterior pharyngeal walls in swallowing. Laryngoscope 1988;98:414–7.

CHAPTER 44

Diseases of the Oral Cavity, Oropharynx, and Nasopharynx

Lawrence W. C. Tom, MD, Ian N. Jacobs, MD

The oral cavity and oropharynx are contiguous anatomic regions, and diseases of these two areas should be considered together. A careful and systematic examination of the oral cavity will include the oropharynx. These regions are frequently exposed to the same external factors. Infectious, inflammatory, and traumatic conditions frequently affect both regions.

Although the nasopharynx is distinct from the oral cavity and oropharynx, the most common pediatric conditions affecting this area involve the adenoid tissue either through infection, hyperplasia, or both. In most cases, the tonsils are also involved, and it is only logical that the adenoids and tonsils be discussed together. Congenital anomalies and neoplasms of the nasopharynx are discussed in other chapters.

ORAL MANIFESTATIONS OF SYSTEMIC DISEASES

Lesions of the oral cavity and oropharynx may be presenting signs of an undetected systemic condition. These regions are readily accessible to the otolaryngologist, and with the proper equipment and a little patience, a thorough examination can usually be performed. Awareness and familiarity of the oral and oropharyngeal manifestations of systemic diseases are important for their early diagnosis. Some diseases can be diagnosed because of a characteristic presentation. Many other signs, however, are nonspecific, but their presence, especially when accompanied by symptoms and signs of disease in other regions of the body, should alert the otolaryngologist to the possibility of a systemic condition.

Infections

Several common viral childhood infections have oropharyngeal manifestations. Rubeola (measles) is characterized by cough, coryza, conjunctivitis, a generalized maculopapular rash, and Koplik's spots. Koplik's spots are red lesions with a pale blue center and are found on the buccal mucosa, most commonly opposite the lower molars. They appear 24 to 48 hours before the rash. With varicella (chickenpox), a generalized pruritic vesicular eruption usually begins on the trunk, spreading to the head and extremities, and is accompanied by mild systemic symptoms. Vesicles can affect any mucous membrane, including the oral cavity and pharynx. The vesicles rupture, progressing to shallow ulcers that resemble aphthous ulcers.

Inflammatory Disorders

Systemic inflammatory diseases frequently affect the oral cavity and oropharynx. Significant sequelae may develop if there is a delay in diagnosis and treatment.

Behçet's syndrome is rare in children. It is an inflammatory disorder of unknown etiology characterized by recurrent oral and genital ulcers and ocular inflammation such as uveitis and iridocyclitis. The initial complaint is usually that of a painful oral ulcer indistinguishable from an aphthous ulcer. The diagnosis should be suspected when the triad of findings is present but may be difficult because multifocal involvement may not occur simultaneously. The most common significant complication is scarring about the eye, leading to visual loss and even blindness. Corticosteroids, chlorambucil, and acyclovir have been used to treat this condition.

Reiter's syndrome is another condition of unknown cause that is uncommon in children. It occurs primarily in young men and classically presents with urethritis, uveitis, conjunctivitis, and arthritis. Oral lesions often occur and are characterized by superficial, erythematous ulcers, which are surrounded by white annular lines.[1] When the ulcers

occur on the tongue, the condition may be mistaken for geographic tongue. Although the ulcers usually present after the onset of the disease, their appearance helps to establish the diagnosis. Most children recover without complications. Treatment is supportive.

Wegener's granulomatosis may affect the oral cavity. Gingival hyperplasia is the most common early oral lesion.[2] Lingual, palatal, and buccal ulcers are also common. Characteristically, they are painful and surrounded by an erythematous rim.

Erythema multiforme is an acute inflammatory disorder of unknown etiology involving the skin and mucous membranes. Its onset has been associated with infections, autoimmune disease, emotional stress, and medications, including penicillins, cephalosporins, and sulfonamides. It affects males more than females and usually occurs in patients 10 to 30 years old.[3] There are two forms of the disease. Erythema multiforme minor causes minimal skin and mucosal damage and is self-limited, lasting 2 to 3 weeks. Erythema multiforme major (Stevens-Johnson syndrome) is associated with systemic manifestations, purulent conjunctivitis, and hemorrhagic, necrotic skin and mucosal lesions. It is common with drug ingestion and may result in death, primarily from pulmonary involvement.

The skin lesions usually begin on the palms and feet and are often accompanied by fever, malaise, and fatigue. The characteristic appearance is the "target" or "iris" lesion, a central bulla surrounded by an erythematous rim. These lesions may merge and typically spread symmetrically over the extremities. Oral lesions occur in 50% of cases, usually simultaneously with the skin lesion, but in 25% of patients, skin involvement is absent, making diagnosis difficult. The lips, tongue, and floor of the mouth are most often affected. The lesions begin as vesicles and bullae that rupture, producing superficial ulcers and hemorrhagic crusts, especially about the lips (Figure 44–1). The diagnosis is made clinically. Treatment of the minor form consists of local care and antihistamines. Treatment of Stevens-Johnson's syndrome requires systemic corticosteroids and, if possible, eliminating the causative agent.

Pemphigus vulgaris, pemphigoid, and cicatricial pemphigoid (benign mucous membrane pemphigoid) are autoimmune vesiculobullous disorders of the skin and mucous membranes. They occur primarily in patients between 40 and 60 years but may affect children. Differentiation among these conditions is based on clinical, histologic, and immunologic criteria. All conditions are characterized by a loss of cohesion among epidermal cells (acantholysis), with a resulting cleft and accumulation of intradermal fluid and blisters. In pemphigus vulgaris, the cleft is in the suprabasal layer and in pemphigoid at the basement membrane zone. Loss of epidermis by rubbing skin and mucous membranes (Nikolsky's sign) occurs in pemphigus vulgaris but not in pemphigoid.

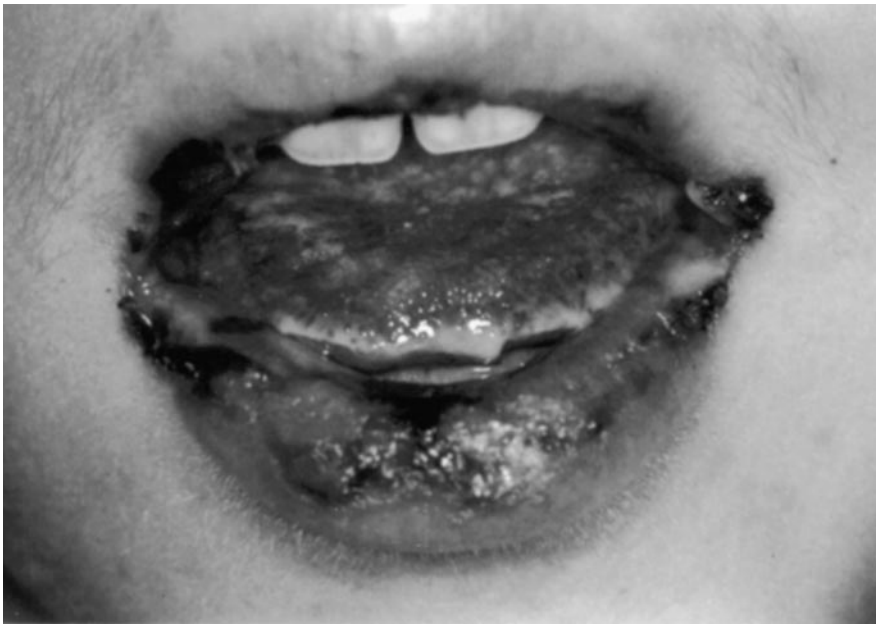


FIGURE 44–1. Erythema multiforme with multiple erosions of the lips and tongue. Courtesy of George Lasarkis. Reproduced with permission of Lasarkis G.[9]

Pemphigus vulgaris is the most common form of pemphigus, representing over 90% of cases. It presents initially in the oral cavity in 60% of cases, and in over 90%, there will be oral cavity involvement. In children, almost all cases begin in the oral cavity.[4] Vesicles and bullae form and rupture, leaving painful, superficial ulcers. These lesions are nonspecific and are often indistinguishable from erythema multiforme. Cases limited to the oral cavity are difficult to diagnose. The presence of lesions on the skin and other mucosal surfaces should raise the suspicion for this disease. Patients are treated with corticosteroids and immunosuppressants. Although the prognosis is generally good, serious cases can be life threatening.

Cicatricial pemphigoid begins in the oral cavity in 95% of cases and eventually involves the oral mucosa in all cases. Desquamative gingivitis is the most common lesion. The mucosal lesions are characterized by recurrent vesicles or small bullae that rupture, causing erosions. The skin and other mucosal surfaces are less frequently involved. Bullous pemphigoid is primarily a cutaneous disease characterized by bullous lesions that rupture, leaving denuded areas. The oral mucosal lesions develop in 40% of cases. Cicatricial and bullous pemphigoid are treated with systemic corticosteroids.

Kawasaki's disease (mucocutaneous lymph node syndrome) is an acute multisystem vasculitis of unknown etiology. It is a disease occurring primarily in children under 5 years of age, with a peak incidence in children 1 to 2 years.[5] It is characterized by 4 to 5 days of high fever, bilateral conjunctivitis, polymorphous erythematous rash, erythema and edema of the extremities (Figure 44–2), cervical lymphadenopathy, and oropharyngeal changes, including strawberry tongue; dry, fissured lips; pharyngeal injection; and necrotic pharyngitis. Early diagnosis and treatment are paramount. Kawasaki disease is the most common cause of acquired cardiac disease in children under 5 years of age. Fifteen to 20% of untreated children will develop cardiac complications, with coronary artery aneurysm being the most serious. Patients are treated with intravenous gammaglobulin and aspirin.

Sjögren's syndrome is a collagen vascular disease characterized by the triad of keratoconjunctivitis, xerostomia, and arthritis. The xerostomia results from the lymphocytic infiltration of the salivary glands and causes dental caries and fissures and ulcers of the tongue, lips, and buccal mucosa. A labial biopsy will establish the diagnosis by demonstrating lymphocytic infiltration of the minor salivary glands.

Systemic lupus erythematosus is a generalized vasculitis affecting many organ systems. It primarily affects adolescents and young women. Twenty-five percent of patients with systemic lupus erythematosus have oral lesions. Xerostomia and its associated conditions are common. Other findings include petechiae, hemorrhagic bullae, and white keratotic plaques on the tongue and oral mucosa.[6]

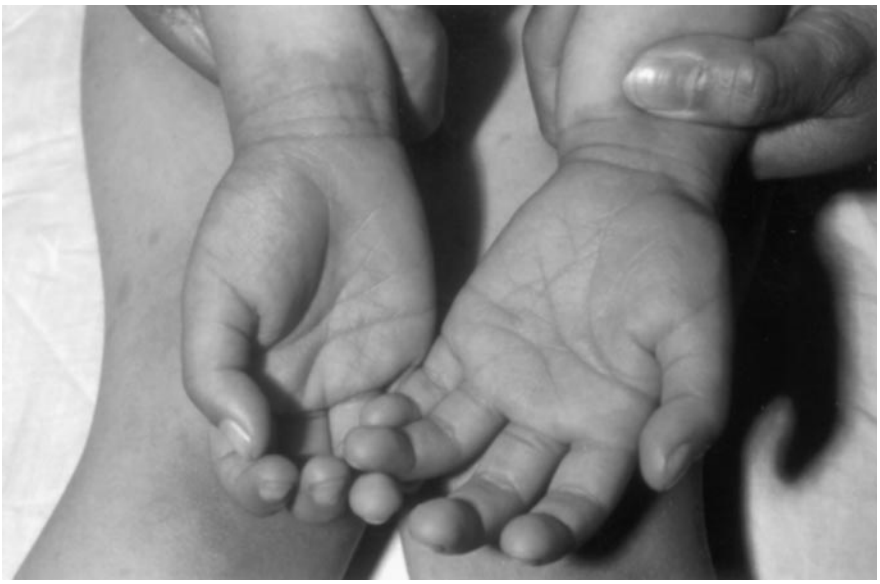


FIGURE 44–2. Bilateral palmar erythema in a child with Kawasaki's disease.

Miscellaneous Disorders

Several vitamin and mineral disorders have oral manifestations. Iron deficiency anemia is common in children. Its oral manifestations include burning of the tongue, pallor of the mucous membranes and gingiva, atrophic smooth erythematous tongue, angular cheilitis, and, occasionally, hyperkeratotic lesions on the oral mucosa. Plummer-Vinson syndrome usually occurs in middle-aged women, with iron deficiency accompanied by dysphagia. The dysphagia is secondary to erosions and webs in the esophagus. This syndrome has been associated with development of oropharyngeal and esophageal malignancies.

Blue-black lines along the gingival margins and spots on the buccal mucosa, palate, and tongue may occur with lead, mercury, arsenic, and bismuth poisoning. Fluoride may cause staining of the teeth, whereas granulomas of the oral cavity have resulted from iodides and bromides.

Vitamin C deficiency, scurvy, is associated with swollen red gingiva, which frequently progresses to bleeding ulcerative gingivitis, oral petechiae and hemorrhages, and stomatitis. Glossitis occurs in the early stages of riboflavin (vitamin B_2) deficiency. As the condition continues, the tongue may appear smooth red. Gingival and labial pallor, angular cheliosis, and burning of the lips, mouth, and tongue may develop. Both vitamin B_{12} and folic acid deficiencies may cause a megaloblastic anemia and present with recurrent oral ulcers and painful atrophy of the oral mucosa and tongue. The signs and symptoms of niacin deficiency, pellegra, include pain, erythema, and edema of the oral mucosa and tongue, necrotizing ulcerative gingivitis, and angular cheilitis. Since multiple vitamin deficiencies are common, many of these findings may be noted together.

Lesions of the oral cavity often are among the initial signs of acute hematologic disorders. Petechiae, ecchymoses, and hemorrhages appear with thrombocytopenia, whereas pallor of the mucous membrane may occur with anemia. Patients with polycythemia may present with a purple-blue discoloration of the tongue and gums. In addition to being anemic and thrombocytopenic, patients with leukemia are often neutropenic, and secondary infections of the oral cavity and oropharynx commonly occur, resulting in necrosis and ulceration of the mucous membranes. Leukemic cells may also infiltrate the soft tissues of the oral cavity. Leukemic infiltration of the gums is common, causing gingival hypertrophy, edema, inflammation, and bleeding.

INFECTIONS

Infections of the oral cavity and oropharynx may be isolated local infections or manifestations of systemic infection. Acute viral pharyngitis constitutes one of the most common reasons for seeking outpatient medical care. The otolaryngologist must recognize the clinical pattern of infections, distinguish viral infections from other diseases, and institute proper treatment.

Viral Infections

Viral infections in the oral cavity may often be misdiagnosed as bacterial infections. At this time, there is no definitive therapy for most viral lesions. Fortunately, most are self-limited in duration. Many can be treated for symptomatic relief with mouthwashes. Occasionally, a superimposed bacterial infection may occur, and antibiotic coverage may be of additional benefit. With some viral infection, the agent may lie dormant for periods of time only to be reactivated, causing clinical exacerbations of the disease.

Herpangina Herpangina is a relatively common viral infection in young children and may mimic bacterial pharyngitis. It is usually caused by coxsackie A virus, but coxsackie B virus and other enteric pathogens may also cause herpangina. The peak incidence is during the summer and fall. Symptoms include rapid onset of severe sore throat, pyrexia, and malaise. Physical examination reveals severe erythema of the oral cavity and multiple small vesicles and superficial ulcers of the mucosa. The most common area of involvement is the posterior aspect of the oropharynx, including the soft palate, tonsils, uvula, and posterior pharyngeal wall. Usually, the gingiva is not affected. The disease is usually self-limited, and treatment is aimed at symptomatic relief.

Hand-Foot-and-Mouth Disease Hand-foot-and-mouth disease is a viral illness of the oral cavity mucosa caused by a coxsackie A virus. It usually affects young children. Its presentation is similar to herpangina, with the exception that it affects the anterior part of the oral cavity more frequently. The tongue, palate, and buccal mucosa are more likely to

be involved. In addition, vesiculopapular lesions are found on the palms and soles.[7] Associated pain and fever may occur. This disease is self-limited, and the treatment is supportive.

Herpes Simplex Herpes simplex virus (HSV) infection may present in two ways in the oropharynx: a primary infection, primary herpetic gingivostomatitis, and recurrent or secondary herpes. Primary herpetic gingivostomatitis is a common illness in the preschool child. Initially, it begins with a prodrome of nonspecific symptoms, including fever and malaise. Within a day or two, the oral mucosa becomes painful and erythematous. Vesicles form and rupture within 24 hours, and the ulcers heal in about 2 weeks. The ulcers form a gray pseudomembrane, may coalesce, and cover most of the oral cavity (Figure 44–3). Diffuse bilateral lymphadenopathy is common. In addition to the oral mucosa, lips, and tongue, the gingiva is almost always affected but often without the formation of vesicles. Diagnosis is made by viral studies and biopsy, which will reveal intranuclear inclusion bodies.

The HSV will remain dormant in the neuroganglion cells for life. Activation of the virus will result in recurrent herpes. Precipitating factors include sunlight, fever, bacterial infections, stress, immunodeficiency, and other systemic illness. Clinically, there may be an initial period of burning in the mouth. Multiple vesicles appear on the oral mucosa and near the vermilion borders, forming typical "cold sores," and rupture over a course of 10 to 14 days. Herpes labialis develops when the lips are involved. In most cases, treatment is symptomatic, but acyclovir may be beneficial.

Herpes Zoster Herpes zoster results from reactivation of the dormant varicella virus in the sensory ganglia. Oral and skin lesions develop along the distribution of the second and third branches of the trigeminal nerve. A classic trigeminal skin pattern establishes the diagnosis. Herpes zoster is uncommon in infants and children and usually occurs in elderly patients. Treatment is with topical antiviral agents such as acyclovir.

Cytomegalovirus Infection Cytomegalovirus (CMV) is a member of the herpes family, and CMV infections are becoming more prevalent, especially with the increasing number of immunosuppressed and immunodeficient patients. In the oral cavity, these infections are characterized by large ulcers of the mucosa and exudative pharyngotonsillitis, which is similar to that of infectious mononucleosis but less severe. Frequently, there is an associated cervical lymphadenopathy. Treatment is primarily symptomatic.

Infectious Mononucleosis Oropharyngeal signs are prominent in infectious mononucleosis. Children, adolescents, and young adults frequently present with a sore throat, grossly enlarged tonsils with a

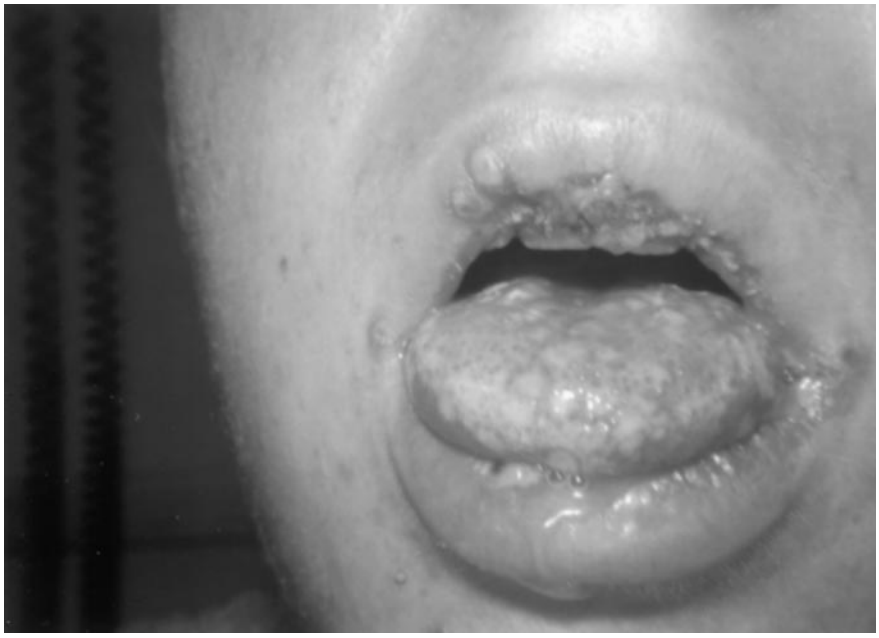


FIGURE 44–3. Herpetic stomatitis with ulcers of the lips and tongue. Courtesy of Dr. Steven Handler.

gray membraneous exudate, and edema and petechiae of the soft palate (Figure 44–4). Other manifestations include cervical lymphadenopathy, which can be marked, hepatosplenomegaly, fever, malaise, and fatigue. Treatment is usually supportive, but when the tonsillar hypertrophy and pharyngeal edema cause respiratory distress, corticosteroids are used.

Human Immunodeficiency Virus Infection There are at least 40 oral manifestations of human immunodeficiency virus (HIV) infection, and the most common sites of involvement are the palate, tongue, lips, and buccal mucosa.[8,9] Signs that suggest the possibility of HIV infection include oral candidiasis; recurrent herpetic and aphthous ulcers; angular cheilitis; hairy leukoplakia, a white corrugated lesion found at the lateral borders of the tongue; periodontitis; gingivitis, especially bleeding ulcerative gingivitis; papillomas; and diffuse parotitis. Kaposi's sarcoma is highly suggestive of HIV infection. In the oral cavity, it may present early as a pigmented macule, papule, or patch or as a hemorrhagic lesion. These lesions may progress to necrosis and ulceration. Later, Kaposi's sarcoma may appear as a lobulated tumor with or without ulceration. In children, the most common signs of HIV infection are oral candidiasis, diffuse parotitis, and oral ulcers.[10]

Recurrent Respiratory Papillomatosis The human papillomavirus causes recurrent respiratory papillomatosis of the airway and may also involve the oral cavity and oropharynx. The virus causes the formation of exophytic, cauliflower-like lesions. Papillomas are found most commonly on the soft palate and tongue. Treatment involves surgical excision with a laser, electrocautery blade, or knife. Histologic analysis establishes the diagnosis and should always be performed because of the possibility of malignant transformation. Recurrences of papilloma in the oral cavity and oropharynx are relatively uncommon after complete excision.

Bacterial Infections

Streptococcal Infections Group A beta-hemolytic streptococcus may cause severe pharyngitis. Patients complain of fever, malaise, sore throat, odynophagia, dysphagia, otalgia, and halitosis. The oropharynx and, to a lesser extent, the oral cavity are swollen and red, and the tonsils are covered with exudates. Strawberry tongue, small red spots on the tongue, is a common finding. Examination of the neck may reveal swollen, tender lymph nodes. Since other bacterial pathogens may account for a similar presentation, cultures are required to make a definitive diagnosis. Scarlet fever produces a diffuse erythematous rash that is caused by an erythrogenic toxin produced by the group A beta-hemolytic streptococcus. Rheumatic fever and glomerulonephritis are potential complications of streptococcal pharyngitis. When streptococcus is suspected, antibiotics should be started before culture results have been reported. Penicillin and amoxicillin are initial choices.

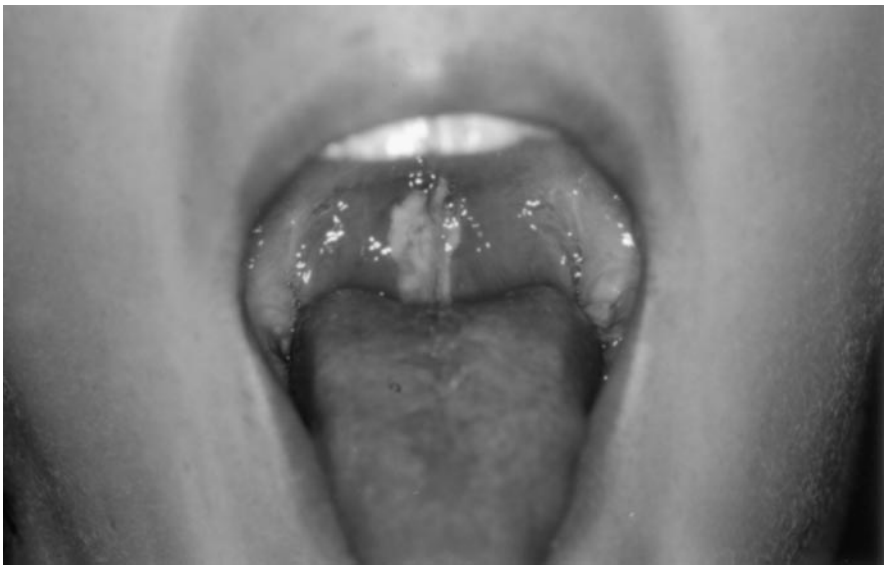


FIGURE 44–4. Exudative tonsillitis in a patient with infectious mononucleosis. Courtesy of Dr. Steven Handler.

Diphtheria Diphtheria, which is caused by *Corynebacterium diphtheriae*, is rare in the United States but not uncommon in other parts of the world. The primary site of involvement is the oropharynx, where it presents as a necrotizing pharyngitis associated with severe erythema and a thick pseudomembrane. The disease may progress to involve the larynx and trachea, leading to airway obstruction. The bacteria produce exotoxins, which may cause systemtic symptoms, including cardiac and neurologic sequelae. *Corynebacterium diphtheriae* may be identified on stains from the pseudomembrane, but cultures are required for definitive diagnosis. Treatment for diphtheria consists of airway support, antitoxin, and antibiotics (see Chapter 52).

Syphilis and Gonorrhea Both syphilis and gonorrhea may involve the oral cavity and oropharynx. A classic feature of congenital syphilis is Hutchinson's teeth, hypoplastic teeth with notching of central incisors. Rhagades, atrophic glossitis, and a high arched palate are other common oral signs. In acquired syphilis, the primary lesion is a chancre, an enlarging ulcer surrounded by induration. If untreated, this resolves spontaneously, but secondary mucosal lesions, most commonly superficial ulcers covered by a grayish-white membrane, will develop. The oral lesions of tertiary syphilis include gummas of the palate, granulomatous lesions that may destroy and perforate the palate, and atrophic glossitis. Although it is uncommon, gonorrhea may affect this region, primarily the oropharynx. A gonococcal infection must be considered in any patient presenting with a membranous pharyngitis (see Chapter 30).

Fungal Infections

Candidiasis is the most common fungal infection of the oral cavity. *Candida albicans* is an opportunistic microorganism and part of the normal oral flora. When there is a reduction in the normal oral flora or a decrease in the host resistance, it becomes increasingly prevalent and causes an overt infection. Predisposing conditions include long-term antibiotic use, diabetes, corticosteroid therapy, and any condition causing immunosuppression. Three presentations are common in children.

Pseudomembranous candidiasis (thrush) is the most common candidal infection in the oral cavity. It occurs in all age groups but most often in infants. Patients complain of burning, dryness, and pain in the mouth. Thrush is characterized by the presence of fine white lacey plaques covering the mucosa of the oral cavity, usually the tongue, hard palate, and buccal mucosa. When the plaque is disrupted, a raw surface remains. In children with diabetes or immunodeficiency, severe invasive oropharyngeal and esophageal candidiasis may develop. The diagnosis is usually made on clinical grounds. Smears may demonstrate hyphae, but cultures are required for a definitive diagnosis.

Acute atrophic candidiasis occurs primarily as a result of long-term antibiotics, long-term corticosteroids, or immunosuppressive agents and is much less frequent than thrush. Patients complain of sore throat, burning of the mouth, and foul taste. Examination reveals red, ulcerated mucosa rather than a pseudomembrane.

Angular cheilitis is manifest by pain and erythema and microfissures at the oral commissure. As the condition continues, erosions of the skin and desquamation of epithelium surrounded by hyperkeratosis develop.

Before treating a candidal infection, a search for an underlying condition should be made. Most children can be treated with a topical, antifungal agent, and if this fails, a systemic agent such as ketoconazole, fluconazole, and even amphotericin can be used.

Other fungal agents may occasionally involve the oral cavity. Histoplasmosis may cause ulcers in the oral cavity. In immunosuppressed and chronically ill patients, mucormycosis may manifest as an ulcer or granuloma, usually on the hard palate.

Miscellaneous Conditions of the Oral Cavity

Leukoplakia Leukoplakia is characterized by the presence of a white plaque on the oral mucosa. It is primarily a condition of adults and has a prevalence of 0.1 to 5%. It is considered a precancerous lesion, with malignant transformation occurring in 2 to 6% of patients. If the plaque adheres to the mucosa and keratosis is present, the chance of malignancy increases. Although it is uncommon in children, it has been noted in children with candidiasis and some viral infections.

Hairy leukoplakia is characterized by slightly elevated, corrugated white plaque usually found on the lateral borders of the tongue. It suggests the

development of acquired immune deficiency syndrome (AIDS) in patients with HIV infection.

Aphthous Stomatitis Recurrent aphthous stomatitis is the most common condition affecting the oral mucosa. The exact etiology is unknown, but recent studies suggest that an altered cellular immune response is a predisposing factor. It occurs in all age groups but usually appears initially in older children and adolescents. The condition is more common in females. Twenty-four to 48 hours before the appearance of the ulcers, there is a prodrome of a tingling, burning feeling and erythema in the oral cavity. There are three forms of ulcers based on their size, number, and duration. Minor aphthous ulcers occur in 80% of the cases. They are painful, shallow, ovoid ulcers less than 0.6 cm in size. The ulcers are covered by a gray membrane and surrounded by a red halo. They may be found alone or in groups, most often on nonkeratinized oral mucosa, such as the buccal mucosa, lips, and tongue. They resolve within 2 weeks without scarring and usually recur in 1- to 5-month intervals. Major aphthous ulcers are a more severe form of the condition. They are characterized by one to five ulcers that are larger than 0.6 cm and more painful than minor ulcers (Figure 44–5). In addition to the sites of minor ulcers, major ulcers also occur on the soft palate, floor of the mouth, and peritonsillar area. They persist for 3 to 6 weeks, may leave a scar, and recur in 1- to 3-month intervals. Herpetiform ulcers, the least common form of the condition, are characterized by crops of painful, multiple, pinpoint ulcers. These ulcers may coalesce, leaving large, irregular ulcers. Herpetiform lesions are found in all locations in the oral cavity, persist for 1 to 2 weeks, and recur at 1- to 3-year intervals. Treatment for all forms of aphthous stomatitis is symptomatic. Topical corticosteroids, topical lidocaine, and mouthwashes may reduce the pain. In severe cases, intralesional or systemic corticosteroids and acyclovir may help.

Lichen Planus Lichen planus is a common inflammatory disorder of the skin and mucous membranes. It primarily affects middle-aged adults but may occur in older children and adolescents. Its cause is unknown, but the disease is a cell-mediated immune response that has been associated with several autoimmune diseases. Oral lichen planus is usually found on the tongue, buccal mucosa, and gingiva. It may follow a course of exacerbations and remissions. There are two variants. The reticular form, the most common type, is characterized by small white papules, which may be isolated or may coalesce to form interlacing white lines (Wickham's striae). They frequently occur with bilateral symmetry. The reticular form causes no symptoms. The erosive form is manifested by painful ulcers with papules or lines at the periphery. Malignant transformation may potentially occur with the erosive form. Diagnosis is established by a biopsy of the edges of the lesions. Asymptomatic lesions require no treatment, whereas painful ones may be treated with topical, injectable, or systemic corticosteroids.

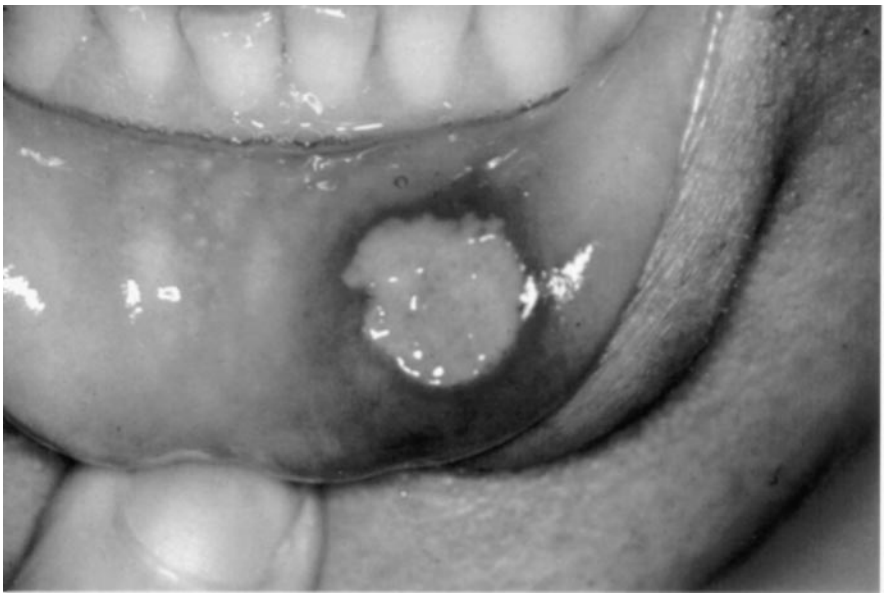


FIGURE 44–5. Major aphthous ulcer. Courtesy of George Lasarkis. Reproduced with permission from Lasarkis G.[9]

Burning Mouth Syndrome Burning mouth syndrome is a disorder of peri- and postmenopausal women. Patients complain of burning on the tip and lateral border of the tongue, which appears normal. It is caused by a local neuropathy. Treatment includes amitriptyline and clonazepam.

Median Rhomboid Glossitis Median rhomboid glossitis is a congenital disorder caused by the persistence of the tuberculum impar. As a result, a rhomboid area free of papillae appears on the dorsal surface of the tongue. It may present as a smooth, red plaque or a firm, raised, smooth, red mass (Figure 44–6). Some cases have been associated with *Candida* infections. In most cases, no treatment is necessary.

Geographic Tongue Geographic tongue (benign migratory glossitis) occurs in all age groups, with a prevalence of 1 to 2%. Although the exact etiology is unknown, it may be an inherited condition. The condition results from a loss of filiform papillae. It is characterized by a white margin of desquamating epithelium of various sizes and shapes on the dorsum of the tongue surrounding a central, red, atrophic area (Figure 44–7). These lesions disappear and migrate to other areas of the tongue. Similar lesions may occur on the oral mucosa. The condition is chronic and may persist for years. Treatment is not necessary.

Fissured Tongue This condition is characterized by multiple fissures or grooves on the dorsum of the tongue. The fissured tongue is usually asymptomatic but may become irritated if food particles, bacteria, or fungi become trapped in the fissures. It is believed to be an inherited disorder and has been associated with the genetic disorders Melkersson-Rosenthal syndrome and Down syndrome. Melkersson-Rosenthal syndrome is characterized by recurrent facial nerve palsy and edema of the face and mouth. Fissuring of the tongue is not pathognomic of Melkersson-Rosenthal syndrome but occurs frequently in it. No treatment is required.

Hairy Tongue This is a common disorder that is characterized by the hypertrophy and growth of the filiform papillae on the dorsal surface of the tongue, giving the tongue a hairy appearance. The papillae may become colonized with pigmented bacteria, producing multiple colors. Black hairy tongue is a common presentation. The cause remains unknown, although certain predisposing factors have been identified, including tobacco abuse, oral antibiotics, poor oral hygiene, and radiation therapy. The disorder is benign and usually asymptomatic. Long papillae may cause discomfort and may be treated with brushing.

Macroglossia The tongue may appear enlarged in many conditions. Relative macroglossia occurs when

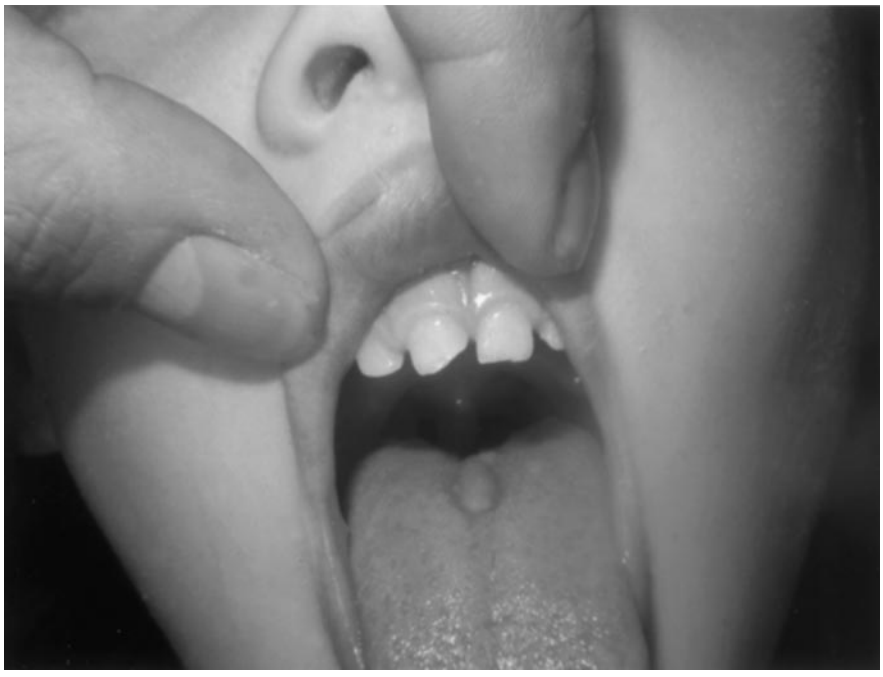


FIGURE 44–6. Median rhomboid glossitis. Courtesy of Dr. Steven Handler.

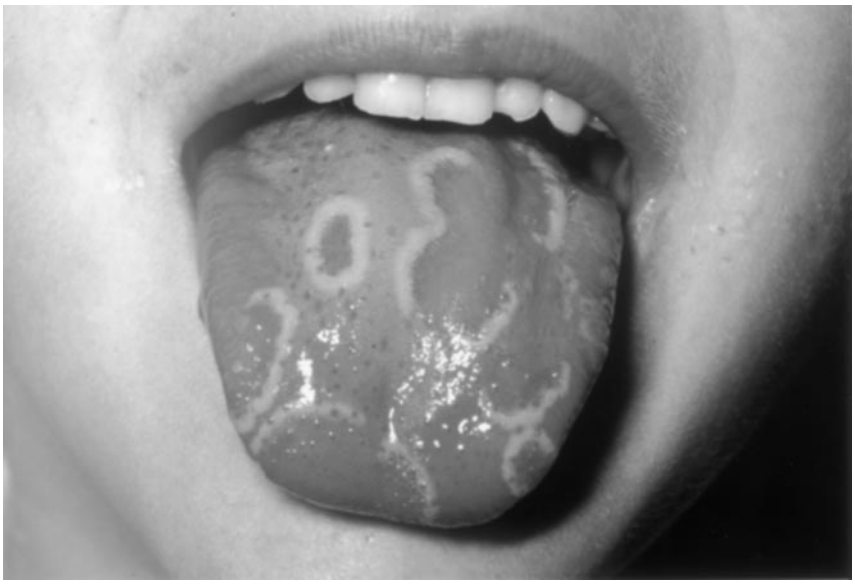


FIGURE 44–7. Geographic tongue.

the tongue is normal in size, but anomalies of the surrounding structures make it appear large. Pierre Robin syndrome and Down syndrome are two such examples. Primary macroglossia is hypertrophy of the tongue musculature, whereas secondary macroglossia is caused by infiltration of tongue tissue.[11] Causes of primary macroglossia include hypothyroidism, acromegaly, and Beckwith-Wiedemann syndrome. Beckwith-Wiedemann syndrome is an autosomal genetic disorder characterized by organomegaly, omphalocele, hypoglycemia, and macroglossia. Many children with this syndrome are at risk for developing airway obstruction. The causes of secondary macroglossia are varied and include lymphangioma, hemangioma, rhabdomyosarcoma, angioneurotic edema, amyloidosis, glycogen storage diseases, neurofibromatosis, actinomycosis, syphilis, and tuberculosis.

Ranulas There are two types of ranulas, and they occur in all age groups. The simple ranula is a cystic lesion of the floor of the mouth representing either a mucocele from obstruction of the sublingual gland or a mucus extravasation pseudocyst from the sublingual gland.[12] It appears as a blue, superficial, nontender, fluctuant mass and is usually found on one side of the floor of the mouth (Figure 44–8). Small ranulas are asymptomatic, whereas larger ones may cause speech and swallowing problems. The plunging ranula is less common. It arises from a mucus extravasation pseudocyst that has extended through the floor of the mouth and presents as a soft, painless, ballotable, submandibular mass. It may present with or without intraoral involvement, with the diagnosis being more difficult if there is no visible intraoral lesion. Simple ranulas may be treated with marsupialization or excision of the sublingual gland. Excision of the sublingual gland with evacuation of the ranula is the recommended treatment for plunging ranulas.

DENTAL INFECTIONS AND CYSTS

Dental infections arise from dental caries, trauma to the teeth, and periodontal disease. The pulp of the tooth becomes inflamed and then necroses, leading to low-grade chronic inflammation or abscess formation. The infected tooth will be painful, tender, sensitive to changes in pressure and temperature, and slightly mobile and will feel raised. The surrounding gingiva will be raised and inflamed, and a draining fistulous tract may develop, indicating pulpal necrosis and abscess in the root canal. A tract may even develop to the face. Electrical vitality testing may demonstrate diminished or absent responses. Dental infections may remain localized to the dentoalveolar process; spread along fascial planes to other regions of the face, head, and neck; or become disseminated. As the infection progresses, radiographic findings appear and include widening of the periodontal membrane space, radiolucency of the periapical bone, and osteitis and destruction of the alveolar

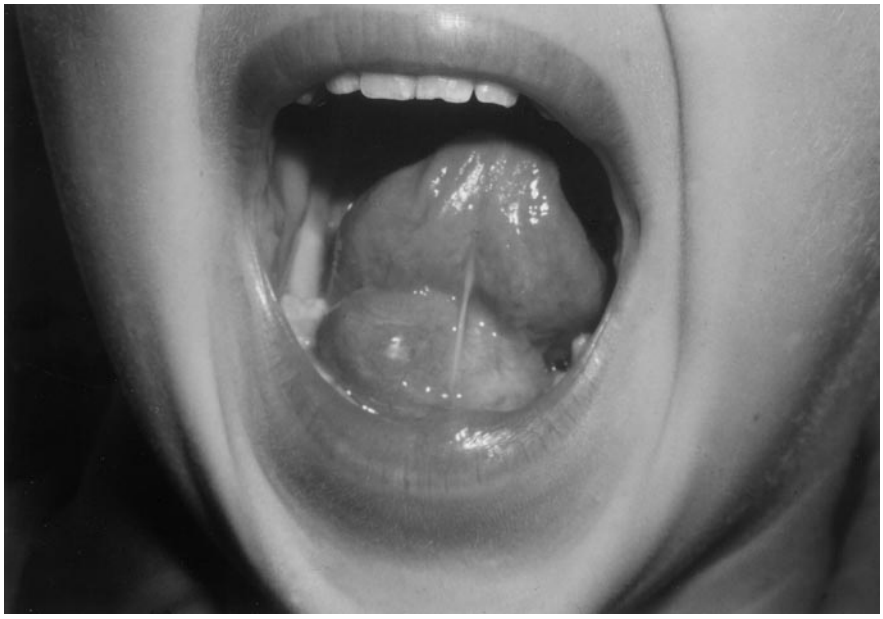


FIGURE 44–8. Ranula. Courtesy of Dr. Steven Handler.

bone. The extent and severity of the infection do not correlate with the radiographic findings.[13] Treatment of dental infections consists of antibiotics and root canal therapy or tooth extraction.

Inflammation of the tooth pulp, gingiva, and periodontal structures may cause epithelial proliferation, leading to the development of certain odontogenic cysts. The periapical cyst accounts for at least 65% of odontogenic cysts. Necrosis of the tooth pulp causes chronic inflammation, leading to granuloma formation. Further inflammation within the granuloma stimulates cyst formation. The most common symptom is discomfort with biting. Large cysts may present as compressible masses in the palate or vestibule. Radiographically, the periapical cyst appears as a well-defined periapical lucency. Dental cysts, which develop from the gingiva, and inflammatory follicular cysts, which develop from dental follicles of unerupted permanent teeth, occur infrequently.

The dentigerous cyst, the second most common odontogenic cyst, develops from expansion of an impacted tooth follicle. It occurs most often in the mandibular third molar. Most of these cysts are asymptomatic and are found on dental radiographs. Large cysts may expand, causing displacement of adjacent structures, infections, and pathologic jaw fractures. Radiographically, it appears as a unilocular, pericoronal radiolucency. The odontogenic keratocyst is associated with bone destruction and frequent recurrences, despite surgery. Other odontogenic cysts include the developmental lateral periodontal cyst and botryoid odontogenic cyst.

BENIGN LESIONS OF THE GINGIVA, MANDIBLE, AND MAXILLA

Many benign soft tissue lesions arise from the gingiva. The pyogenic granuloma is a smooth, red, soft, pedunculated lesion that bleeds easily. It develops as a response to trauma and is treated by local excision. The peripheral giant cell reparative granuloma also develops from trauma. It presents as a sessile reddish mass of vascular tissue that bleeds easily and is most often found on the gingiva anterior to the premolars. Treatment is with excision or curettage. Epstein's pearls are small keratin cysts found on the alveoli or palate of the newborn. They are transient, exfoliating within a few weeks. The congenital epulis is usually found on the anterior maxillary gingiva, presenting as a pink, round, lobulated, pedunculated, mucosa-covered mass. Excision is curative. The melanotic ectodermal tumor of infancy occurs in the first 6 months of life and usually on the anterior maxilla near the junction of the globular and maxillary processes. It presents as a firm, well-circumscribed mass that may be pigmented and may cause facial swelling and nasal obstruction. Treatment is complete excision.

A variety of benign tumors develop from the mandible and maxilla. Odontogenic tumors arise

from the precursors of tooth formation. The odontoma, the most common odontogenic tumor, is a hamartoma arising from a mixture of dentin, enamel, and cementum. It is often associated with dentigerous cysts. It grows by expansion but rarely becomes large enough to be symptomatic. Treatment is with enucleation. The ameloblastoma arises from tissue involved in enamel formation. Initially, it presents as a painless, hard mass, but cystic degeneration and softening develop. Ameloblastomas may be locally aggressive, infiltrating into adjacent tissue and causing significant deformity. Radiographs demonstrate a multicystic, osteolytic, and expansile lesion.[14] Treatment is complete excision. Myxomas occur most often in children over 10 years of age and young adults. They also present as painless masses, often associated with loose teeth. Myxomas are locally aggressive, infiltrating into neighboring bone. They require wide resection, but recurrence is not uncommon. Adenoameloblastoma, cementoma, and cementoblastoma are rare odontogenic tumors.

Nonodontogenic tumors of the jaws arise from tissues other than dental precursors. The most common lesion is the torus, an exophytic, mucosa-covered, bony overgrowth. It is usually noted after puberty. It occurs in the midline of the palate in 20% of the population. In 8% of the population, a torus occurs near the midline of the lingual surface of the mandible, and 80% of these mandibular tori are bilateral. Most are asymptomatic and require no treatment.

Fibrous dysplasia of the jaws is found more often in the maxilla and is usually monostotic. It frequently presents as painless swelling in the canine fossa or zygoma. These lesions grow rapidly from early childhood to early adolescence, causing facial swelling and even proptosis. With puberty, fibrous dysplasia tends to stabilize. Conservative subtotal resection is recommended for lesions causing cosmetic or functional problems.

The giant cell reparative granuloma occurs primarily in patients 10 to 20 years of age and more often involves the mandible. It presents as a nontender intraoral mass with an accompanying mandibular deformity. This lesion is treated with curettage or excision. Other causes of nonodontogenic tumors of the jaws include osteoma, eosinophilic granuloma, aneurysmal bone cysts, hemorrhagic bone cysts, and cherubism. Malignant neoplasmas of the oral cavity and oropharynx are discussed in Chapter 61.

TONSILS AND ADENOID

Diseases of the tonsils and adenoid are some of the most common problems seen in children. Adenotonsillar disease can be broadly classified as infectious, obstructive, and miscellaneous. The "miscellaneous" group includes unilateral tonsillar hypertrophy and tonsilloliths.

INFECTIONS

Tonsillitis Acute tonsillitis is defined as an acute infection of the tonsils with symptoms of sore throat, fever, odynophagia, and general malaise. Physical findings include tonsillar hypertrophy and erythema and exudates on the tonsillar surfaces. There may be associated cervical lymphadenopathy, skin rashes, and fever. The disease is usually self-limited, lasting from 7 to 14 days, but during this period there may be significant loss of time from school or work.

An episode of acute tonsillitis may progress to recurrent acute tonsillitis, repeated episodes of acute tonsillitis followed by periods in which the patient is asymptomatic. During each acute episode, the patient may develop symptoms of acute tonsillitis. In addition to the signs of acute tonsillitis, patients experiencing recurrent acute tonsillitis may develop enlarged tonsillar crypts that accumulate debris, persistent erythema of the tonsils, and dilated blood vessels on the surface of the tonsils.

Many pathogens may be cultured from the surface and the core of the tonsils and may cause acute tonsillitis. Of greatest concern is group A beta-hemolytic streptococcus because it may lead to the development of rheumatic cardiac disease and glomerulonephritis. Group A beta-hemolytic streptococcus may be isolated on the tonsil surface but in up to 60% of cases may be found deep within the tonsillar crypts.[15] Therefore, cultures of the tonsil surface may not detect group A beta-hemolytic streptococcus as the cause of acute tonsillitis.

Treatment of acute tonsillitis should be based on clinical judgment, taking into consideration the entire clinical picture rather than just cultures of the tonsil surface. Antimicrobial therapy may be started even in the absence of a positive culture for group A beta-hemolytic streptococcus. In most cases, penicillin and amoxicillin are the initial drugs of choice. If there is a history of treated recurrent acute tonsillitis, the responsible pathogens are likely to be β-lactamase-producing bacteria, and a β-lactamase-stable

antibiotic, such as amoxicillin-clavulanate, should be used.[16]

Some patients may be carriers of group A beta-hemolytic streptococcus. If the patient is asymptomatic, no treatment is necessary. Patients who experience frequent episodes of acute tonsillitis, infect other patients, or develop complications should be treated. An initial course of antibiotics may eradicate the bacteria and carrier state. Adenotonsillectomy is indicated for treatment failures.

Children who continue to experience recurrent acute tonsillitis, despite adequate antibiotic therapy, should be considered for tonsillectomy and adenoidectomy. In over 70% of cases, core tonsil and core adenoid tissue harbor the same pathogens, and both tonsillectomy and adenoidectomy lead to their eradication.[17] Patients with three or more episodes of acute tonsillitis within a year are candidates for tonsillectomy and adenoidectomy.[18] It is important to document the dates of the last two infections, degree of fever, severity of disease, results of any throat cultures, and response to antibiotics.

Chronic tonsillitis is defined as persistent tonsillar infection and usually occurs in older children and young adults. Patients complain of constant throat pain, halitosis, and fatigue. Examination of the tonsils reveals hypertrophy, erythema, and enlarged crypts filled with debris. Chronic tonsillitis is treated with tonsillectomy if the symptoms are severe or the persistent infection interferes with normal activities.

Peritonsillar Infections Peritonsillar infections occur in all age groups but more often in adolescents and young adults. It is the most frequent head and neck space infection in children and the most common complication of acute tonsillitis. Infections of the tonsillar crypts extend beyond the tonsillar capsule to involve the peritonsillar space. Most infections occur in the superior pole of the tonsil, but some involve the midtonsillar area and inferior pole. The infection begins as a cellulitis and progresses to an abscess.

Patients complain of fever, unilateral sore throat, odynophagia, and trismus. The examination may be difficult because of the inability to open the mouth. The classic signs include a muffled voice; drooling, unilateral swelling, and erythema of the superior tonsillar pole; deviation of the uvula to the opposite side; and bulging of the posterolateral part of the soft palate. The oral airway may be compromised. The diagnosis of a peritonsillar space infection is usually made clinically, but it may be difficult to distinguish between cellulitis and an early abscess.

Patients with a suspected peritonsillar infection should be hydrated and given intravenous antibiotics and analgesics. If there is no improvement after 48 hours, airway compromise, or definite abscess formation, drainage should be considered. There is no consensus regarding the best technique. Options include needle aspiration, incision and drainage, and immediate tonsillectomy. Needle aspiration and incision and drainage have success rates greater than 90%, but there is a 10 to 15% risk of abscess recurrence.[19] Immediate tonsillectomy is curative and prevents recurrence, but in many patients, tonsillectomy may not be necessary. Tonsillectomy should be performed in patients with a history of recurrent infections or previous abscess.

Adenoiditis The adenoid tissue may become infected during an upper respiratory tract infection, causing fever, nasal obstruction, purulent rhinorrhea, postnasal discharge, and cough. These signs and symptoms are nonspecific and can occur with respiratory tract infections and sinusitis. Examination with a mirror or telescope may demonstrate inflamed, swollen adenoid covered with exudates. Treatment is with antibiotics, and if the condition becomes recurrent, antibiotics with activity against β-lactamase-producing microorganisms are recommended. If the condition persists, adenoidectomy is performed.

The relationship between adenoid infections and sinusitis has not been clearly established. Since the presentation of these two conditions is similar, adenoiditis may be misdiagnosed as sinusitis, and adenoiditis may be a causal factor in pediatric sinusitis. In a child with chronic sinusitis, it has been suggested that an adenoidectomy be performed a minimum of 3 months prior to pediatric endoscopic sinus surgery.[20] If there is significant improvement of symptoms, sinus surgery may be avoided.

Obstruction

Upper airway obstruction is the second most common indication for tonsillectomy and adenoidectomy in children and may present in an acute or chronic fashion. Infections, such as mononucleosis or peri-

tonsillar abscess, may cause acute upper airway obstruction. With a history of sore throat and dysphagia, examination may reveal an acute infection with signs of airway obstruction. Some children may be managed as outpatients with antibiotics and corticosteroids. Others may require hospitalization, especially to monitor and secure the airway. If symptoms persist, an emergency tonsillectomy and adenoidectomy may be required. These patients may need to recover in a carefully monitored environment such as a pediatric intensive care unit. Sudden relief of upper airway obstruction may cause postobstructive pulmonary edema. Fluid may extravasate into the alveoli with a rapid change in intrathoracic pressure. Immediate recognition of this problem is essential. Treatment consists of intubation with the administration of positive end-expiratory pressure and diuretics.

Chronic upper airway obstruction most often occurs when both tonsils and adenoid are enlarged but may occur when either is enlarged. A thorough history is important to establish the diagnosis. Questions regarding mouth breathing, snoring, snorting, gasping, and apnea should be elicited. Parents should be asked about a child's sleep habits. Restless sleep, extension of the neck during sleep, enuresis, frequent waking, growth disturbances, and failure to thrive suggest airway obstruction. Dysphagia and choking on solid foods indicate obstructing tonsils, whereas hyponasal speech suggests enlarged adenoid tissue.

The obstruction is related to the relative size of the tonsils and adenoid. An oral examination will usually reveal enlarged, obstructing tonsils. Nasopharyngeal examination, which is usually performed with a nasopharyngoscope, may be performed in cooperative children and reveals the size of the adenoid and presence of obstruction. Collapse of the pharyngeal musculature and lymphoid structures into the airway during sleep increases the obstruction. Children with poor neuromuscular tone may be more susceptible to airway collapse during sleep, whereas children with craniofacial anomalies may be at greater risk for obstruction from lymphoid hypertrophy. In these children, the tonsils and adenoid may not appear enlarged but may contribute to the obstruction. Depending on the obstructing tissue, tonsillectomy and/or adenoidectomy can be performed.

Children with obstructing tonsils and adenoid may have obstructive sleep apnea (OSA). Sleep apnea is defined as the intermittent cessation of air flow at both the mouth and nose during sleep. There is a repetitive hypopnea (reduction in breathing) or apnea (cessation of breathing). Obstructive sleep apnea is caused by an obstruction in the upper airway, and there is a vigorous effort to overcome this obstruction. Obstructive sleep apnea must be distinguished from the less common central sleep apnea in which the cause is in the brainstem. With central sleep apnea (CSA), the neural drive to the respiratory muscles is temporarily abolished, and there is little effort to breathe.

If the diagnosis is equivocal, a lateral neck radiograph can be performed to assess the size of the tonsils, adenoid, and nasopharyngeal airway. Either polysomnography or sleep sonography will provide objective data regarding the presence and degree of obstruction. During polysomnography, the frequency of apneas per hour (apnea index), apneas and hypopneas per hour (hypopnea-apnea index), and arousals per hour (arousal index) is recorded. The severity of oxygen desaturation, degree of hypercapnia, and frequency and grade of any cardiac arrhythmia should also be recorded. Although polysomnography provides the best objective measurement for obstruction, it is expensive and inconvenient for most children and their families. For normal, nonsyndromic children, there is a close correlation between the results obtained from polysomnography and sleep sonography.[21]

Miscellaneous Conditions

Tonsilloliths Tonsilloliths are concretions of epithelial debris that develop within the depth of the tonsillar crypts. Prior tonsil infections have sealed the opening of the crypts, allowing the collection of the the debris. Tonsilloliths appear as gritty, caseous white or yellow calculi up to 1 cm in size and have a foul odor. They may be removed with a blunt instrument or water irrigation but frequently recur. If the patient is concerned about the condition, tonsillectomy will eliminate the condition.

Unilateral Tonsillar Enlargement The finding of a unilaterally enlarged tonsil often raises the concern of a neoplasm within the tonsil or tonsillar fossa. In most cases, the larger-appearing tonsil is situated more medially within the tonsillar fossa, giving the impression of enlargement. In some cases, the enlargement is caused by a unilateral infection. Only

rarely is the enlargement caused by a malignancy. In these cases, the tonsil is not only enlarged, but the surface appears abnormal, with areas of necrosis and ulceration. If a malignancy is suspected, the tonsil is removed and sent for appropriate studies.

Otitis Media with Effusion The adenoid affects eustachian tube function in some patients. Regardless of the size, removal of adenoid tissue may reduce the extrinsic mechanical obstruction of the eustachian tube, improving the ventilatory function of the eustachian tube and allowing for equalization of pressure.[22] In addition, adenoidectomy has been shown to decrease the recurrence rates of otitis media.[23] Adenoidectomy should be considered in any child with evidence of nasal obstruction or adenoiditis and who is undergoing tympanostomy tube placement. It should also be considered in any child in whom tympanostomy tubes are being replaced (see Chapter 9).

Technique of Tonsillectomy and Adenoidectomy

There are a variety of techniques for removal of the tonsils and adenoids. For a tonsillectomy, dissection in the subcapsular plane can be performed with a knife, protected electrocautery blade, harmonic scalpel, or laser. Hemostasis can be obtained with electrocautery, suture ligation, and topical thrombin. The adenoid can be removed with a curette, adenotome, or powered instrument.

Today, most procedures are performed under general anesthesia. The patient is intubated and placed supine in the Rose position with a roll under the shoulders. A mouth gag is placed to expose the oropharynx. The palate is palpated to detect the presence of a submucous or frank palatal cleft. A complete adenoidectomy should be considered with great caution if a cleft is detected because there is an increased likelihood of developing velopharyngeal insufficiency. If a cleft exists, superior adenoid tissue may be removed, leaving the base to participate in velopharyngeal closure.

Two urinary catheters are used to retract the soft palate, and the nasopharynx is visualized with a mirror. A curette is placed at the base of the vomer and advanced, removing the bulk of the adenoid. Any additional pieces of tissue can be removed in a similar fashion. When the adenoid has been removed, sponges are placed into the nasopharynx to help with hemostasis. The sponges are removed individually, and hemostasis is obtained with suction electrocautery. Cauterization should not be performed near the eustachian tube orifice to prevent development of scarring and subsequent eustachian tube dysfunction.

Tonsillectomy is performed next. The superior pole of the tonsil is grasped with an Allis clamp. The mucosa of the anterior tonsillar pillar is incised with an electrocautery blade (Figure 44–9). The plane of the tonsil capsule is identified. As the tonsil is retracted medially, dissection continues inferiorly in this plane (Figure 44–10). The inferior attachment of the tonsil is severed with cautery. Hemostasis is achieved using a suction cautery. A Hurd retractor may be used to expose areas of the tonsillar fossa to obtain better hemostasis. After hemostasis is achieved, the catheters and mouth gag are removed.

Complications Bleeding is the most serious complication of tonsillectomy and occurs in up to 3% of

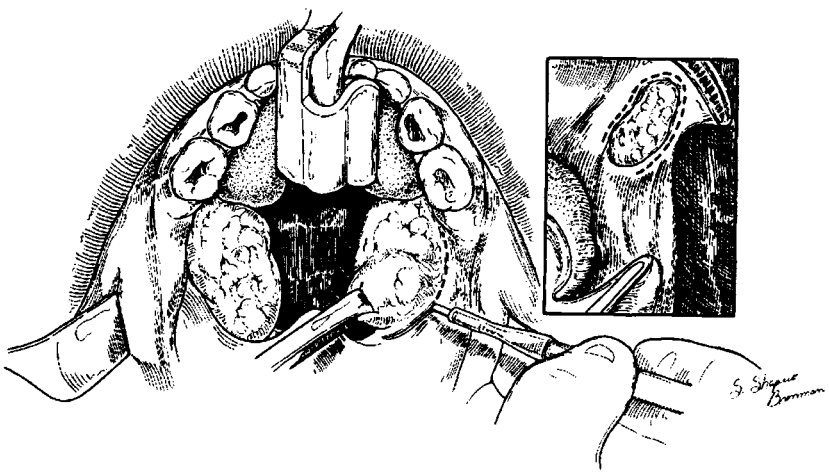


FIGURE 44–9. The tonsil is grasped with an Allis clamp, and the initial incision is made with an electrocautery blade. Courtesy of William P. Potsic, Robin T. Cotton, Steven Handler. Surgical pediatric otolaryngology. New York: Thieme Medical Publishers; 1997.

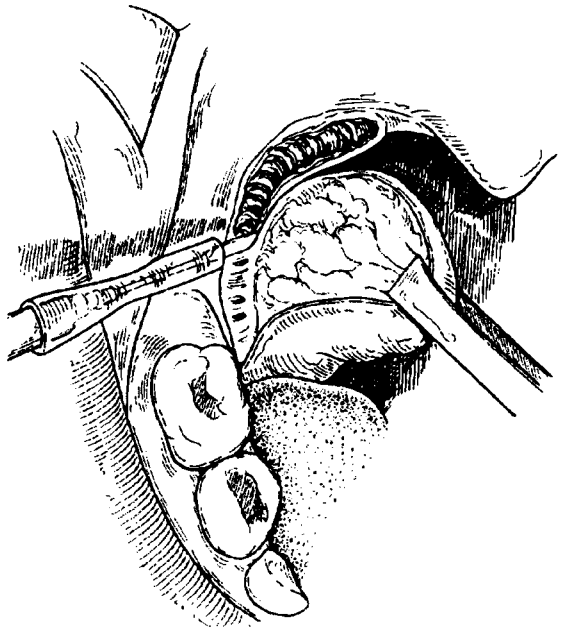


FIGURE 44–10. Dissection continues inferiorly in the plane of the tonsil. Courtesy of William P. Potsic, Robin T. Cotton, Steven Handler. Surgical pediatric otolaryngology. New York: Thieme Medical Publishers; 1997.

patients. Many of these patients must return to the operating room to control the bleeding, whereas others need only observation. Although it is rare, death from tonsillectomy still occurs and is usually a result of aspiration of blood into the airway rather than loss of blood volume.

Dehydration is a common complication in children. The children have severe sore throats and will not swallow because of the pain. These patients require intravenous hydration either on an inpatient or outpatient basis until they can maintain an adequate fluid intake.

Hypernasal speech is common immediately following the procedure because the pain causes limitation of the movement of the tonsillar pillars and soft palate. This usually resolves in several weeks. Permanent velopharyngeal insufficiency is usually caused by resection of the posterior pillars or an unrecognized palatal cleft. This may result in rhinolalia aperta and reflux of liquid through the nose. Treatment includes speech therapy and placement of a pharyngeal flap.

Torticollis is not uncommon and is usually caused by inflammation of the paraspinus muscles. This usually resolves spontaneously but may require a cervical collar or even cervical traction.

Nasopharyngeal and oropharyngeal stenosis, rare complications, are caused by resection of the posterior tonsillar pillars and cauterization of the base of the tongue, respectively. Nasopharyngeal stenosis is more common and is discussed in more detail later in this chapter. Oropharyngeal stenosis is usually caused by cauterization of the base of the tongue and posterior tonsillar pillar. Adhesions form between these structures, decreasing the size of the oropharynx.[24] Treatment includes injection with corticosteroids alone or combined with mucosal flaps.

HEAD AND NECK SPACE INFECTIONS

Antibiotics have reduced the incidence of infections of the fascial spaces of the head and neck, but these infections continue to occur. Atypical presentations are not uncommon in immunocompromised patients.

Fascia envelops the muscles, vessels, and viscera of the neck. Fascial planes form where adjacent fascia condenses. Fascial spaces are potential spaces between these planes.

The cervical fascia consists of two layers, the superficial or investing layer and deep fascia (Figure 44–11). The superficial layer lies just below the skin, completely encircles the neck, and is continuous with fascia of the face and the superficial fascia of the muscles of the back. Within this layer lie the platysma muscles, external jugular vein, and lymph nodes. The deep fascia comprises three layers. The anterior layer envelops the trapezius, sternocleidomastoid, omohyoid, and strap muscles and the parotid and submandibular glands. The middle or visceral layer surrounds the pharynx, esophagus, larynx, trachea, and thyroid gland. The prevertebral layer covers the vertebral bodies and paraspinus muscles. Contributions from all three layers form the carotid sheath, which contains the vagus nerve, carotid artery, and internal jugular vein.

Fascial spaces are potential avenues for the spread of infection. The spaces formed by fascia-enveloping muscle offer some resistance of the spread of infection, whereas the spaces formed by fascia-enveloping viscera and vessels are associated with the most serious infections because they offer little resistance to this spread.[25] Clinically, the most

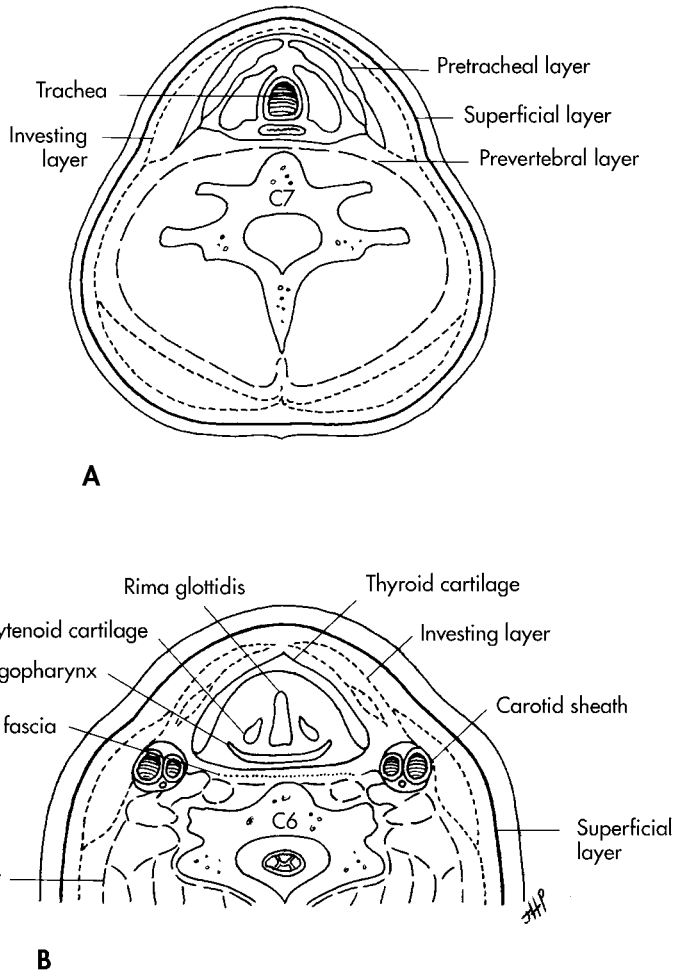


FIGURE 44–11. Fascial planes of the neck. *A*, Schematic representation of the fascial planes of the neck at the level of cervical vertebrae 7. *B*, Schematic representation of the fascial planes of the neck at the level of the thyroid cartilage.

important spaces are the prevertebral, retropharyngeal, lateral pharyngeal, and submandibular (Figure 44–12). The masticator, parotid, and peritonsillar spaces also have significance.

The prevertebral space lies between prevertebral fascia and the vertebral bodies and extends from the base of the skull to the coccyx.[26] Infections of this space may spread from the base of the skull to the psoas muscles.

The retropharyngeal space lies between the prevertebral fascia and the fascia covering the posterior pharyngeal wall and esophagus. It extends from the skull base to the tracheal bifurcation. It contains lymph nodes that receive afferent lymphatic vessels from the nose, nasopharynx, paranasal sinuses, oropharynx, and middle ear. These lymph nodes are prominent in children under 5 years of age and begin to atrophy after this age. This space is a major route for spread of infection into the mediastinum.

The lateral pharyngeal (parapharyngeal) space is cone shaped, with its base at the petrous portion of the temporal bone and its apex at the hyoid bone. Its contents include lymph nodes receiving afferents from the nasal cavity, paranasal sinuses, nasopharynx, oropharynx, and mastoid tip; the carotid artery; internal jugular vein; cranial nerves IX, X, XI, and XII; and the cervical sympathetic chain. The carotid sheath pierces this space at its apex. The lateral pharyngeal space communicates with the submandibular, retropharyngeal, and parotid spaces and is in close proximity to the masticator and peritonsillar spaces. The lateral pharyngeal space is the most frequently involved with serious infections and most frequent route of spread of infection from one region to another.

The submandibular space is divided by the mylohyoid muscle into two compartments, the sublingual space superiorly and submandibular space

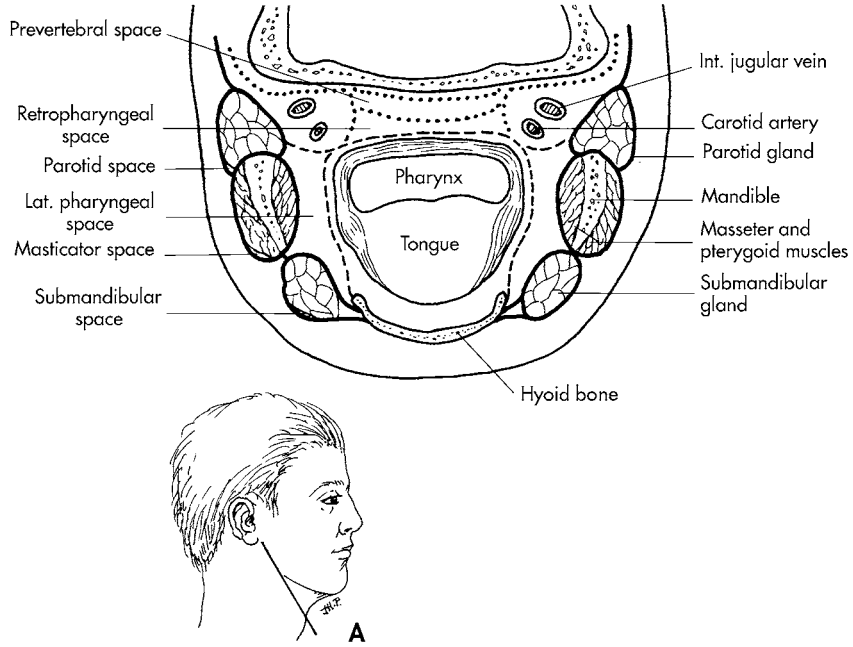


FIGURE 44–12. Fascial spaces of the neck at the plane indicated in *A*.

proper inferiorly. The two compartments are connected through an opening between the mylohyoid and geniohyoid muscles and communicate with their counterparts on the opposite side. The submandibular space is bound by the floor of the mouth, tongue, mandible, anterior layer of deep cervical fascia, and hyoid bone. It contains the sublingual and submandibular glands, lingual and hypoglossal nerves, facial artery, and lymph nodes.

The parotid space is formed when the anterior layer of deep cervical fascia splits and invests the parotid gland. The medial portion of the fascia is incomplete and communicates with the lateral pharyngeal space. Infections of the parotid may extend medially into the lateral pharyngeal space.

The masticator space surrounds the masseter, temporalis, and internal pterygoid muscles and ramus of the mandible and is adjacent to the lateral pharyngeal space. Since its fascia surrounds muscles, it offers some resistance to the spread of infections.

The peritonsillar space is bound by the capsule of the palatine tonsil, the anterior and posterior tonsillar pillars, and superior constrictor muscle. Peritonsillar abscesses may progress to involve the lateral pharyngeal space and carotid sheath.

Many microorganisms are responsible for head and neck space infections, reflecting the numerous possible sources. A mixed bacterial flora is frequently isolated from abscesses. The gram-positive aerobes, beta-hemolytic streptococcus and *Staphyloccocus aureus*, are the predominant pathogens. Gram-negative aerobes and anaerobes are also common. Anaerobes are especially prominent when an odontogenic infection is the primary source. Beta-lactamase-producing microorganisms have been found with increasing frequency in head and neck space infections.[27]

With the significant presence of gram-negative and β-lactamase-producing bacteria, penicillin is no longer the initial antibiotic of choice in the treatment of these infections. Before culture and sensitivity results have been obtained, the empirically chosen antibiotic should have activity against gram-positive and negative aerobes, anaerobes, and β-lactamase-producing microorganisms. Antibiotic combinations, such as clindamycin and cefuroxime and ampicillin and sulbactam, are excellent initial choices.

The diagnosis of a head and neck space infection can often be made after a history and physical examination. Most patients should have a complete blood count with differential, electrolytes, and coagulation studies. If a primary source is suspected, Gram stains and cultures of the infection should be taken, and if there is evidence of sepsis, blood cultures should be taken.

Radiologic studies help in determining the extent and severity of the infection, the primary source, and the presence of complications. The specific studies are determined in the individual case. Both the suspected source and affected space should be evaluated. Anteroposterior and lateral neck films may demonstrate airway compromise, tracheal deviation, and widening of the retropharyngeal space, indicating cellulitis or an abscess (Figure 44–13). An air-fluid level in the retropharyngeal space confirms an abscess. Chest radiographs are required to evaluate the airway and the mediastinum.

Contrast-enhanced computed tomography (CT) is the most valuable radiologic study. It is excellent for evaluation of all of the fascial spaces and may demonstrate findings of infection such as soft tissue edema, obliteration of fat planes, deviation of structures, and airway compromise. It is excellent for differentiating between an abscess and cellulitis. The finding of a rim-enhancing lesion or fluid within the lesion is diagnostic for an abscess. Contrast-enhanced CT will also demonstrate the size and location of an abscess and its position relative to the carotid artery and internal jugular vein (Figure 44–14). Contrast-enhanced magnetic resonance imaging (MRI) may be used, but it does not appear to offer any advantages over CT. Ultrasonography may help localize an abscess.

Although the cause and clinical presentation of these infections are dependent on the space involved, there are several important principles that apply to most of these infections. In children, infections of the upper aerodigestive tract are the most common cause of head and neck space infections, whereas odontogenic infections are the most common cause in adults. It is not uncommon to fail to identify a primary source of the infection. Oropharyngeal pain, fever, and limitation of jaw and/or neck movement are present in most of these infections. Initial treatment consists of intravenous antibiotics, hydration, and analgesics.

Peritonsillar abscess is the most frequent head and neck space infection in children and the most common complication of acute tonsillitis. Most infections occur in the superior pole, but some involve the midtonsillar area and inferior pole. The infection begins as a cellulitis and progresses to an abscess.

The submandibular space is the second most frequent location for head and neck space infections in children. Most infections are odontogenic in origin, usually from the second and third lower molars, which penetrate the mylohyoid muscle into the submandibular space.[28] Infections of the submandibular gland and nodes are other sources of infections. In children, it is not uncommon to find no primary

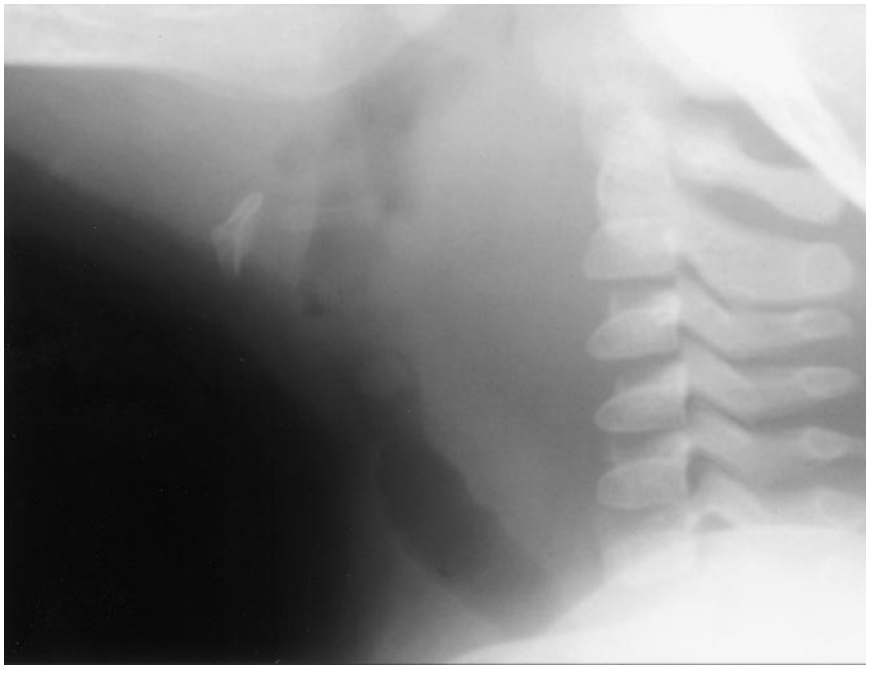


FIGURE 44–13. Lateral neck radiograph demonstrating widening of the retropharyngeal space.

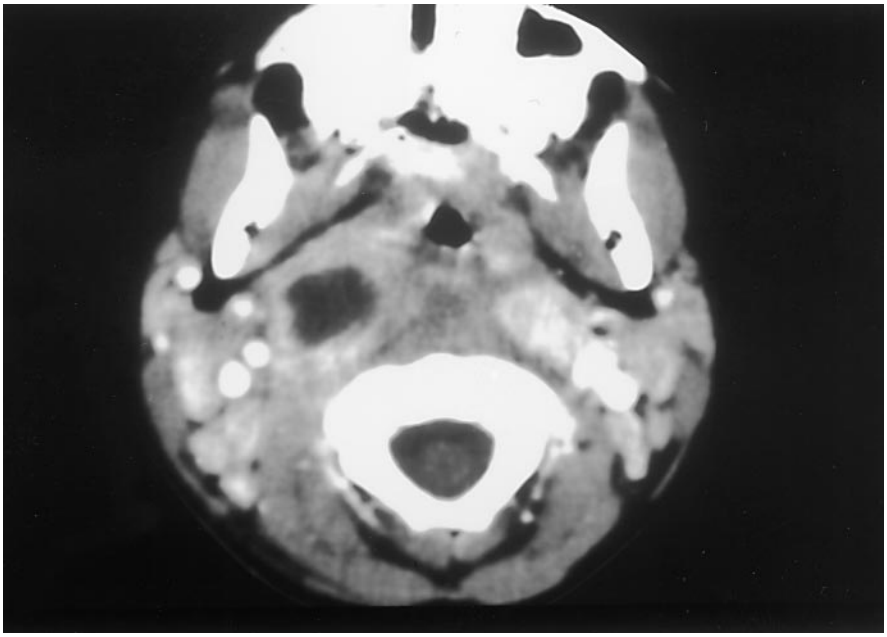


FIGURE 44–14. Contrast-enhanced computed tomographic scan of the neck showing a lateral pharyngeal space abscess medial to the great vessels.

source.[29] Ludwig's angina is a rapidly spreading cellulitis in the superior part of this space.

The initial presentation is dependent on the site of the infection. If the infection begins in the sublingual space superior to the mylohyoid muscle, pain, dysphagia, drooling, and muffled voice are prominent symptoms. Examination may reveal swelling and tenderness of the floor of the mouth, and the tongue may be displaced posteriorly, causing airway compromise. Infections beginning in the submandibular space inferior to the mylohyoid muscle present with cervical pain associated with tense, brawny suprahyoid and submental edema. An infection beginning in either space may spread both to its adjacent space and to the other side, producing a true Ludwig's angina. When the entire submandibular space is involved, the presentation is a combination of the sublingual space and submandibular space proper infections. Dysphagia and airway compromise are more severe because of the increased degree of tissue tension. Trismus is prominent because of involvement of the suprahyoid muscles.

The diagnosis of submandibular space infections can usually be made clinically. Plain lateral and anteroposterior neck radiographs may reveal airway compromise, whereas CT with contrast may demonstrate an abscess. Panoramic tomography may help to determine the primary source of infection.

Treatment must be aggressive. Intravenous antibiotics, management of the primary source, and supportive therapy may be sufficient to treat early cases. With Ludwig's angina, airway obstruction is a main concern. The airway must be secured with an endotracheal tube, often using fiber-optic intubation or a tracheostomy. After the airway is secured, the submandibular space is drained. In Ludwig's angina, colored, weeping exudate, not frank purulence, is usually found. Drains are placed deep to the mylohyoid muscle and into the sublingual space.

Retropharyngeal space infections usually occur primarily as a sequela of upper respiratory tract infections that drain into the retropharyngeal lymph nodes. The disease occurs primarily in young children. Retropharyngeal infections do occur in adults, and the incidence in adults may be greater than has been previously recognized.[30] Foreign bodies and trauma are other causes of these infections.

There is often a history of a recent upper respiratory tract infection. Symptoms include fever, sore throat, odynophagia, dysphagia, drooling, muffled voice, and difficulty breathing. Examination may reveal torticollis; limitation of neck movement; cervical tenderness; boggy, inflamed orpharyngeal mucosa; and bulging of the posterior pharyngeal wall to one side. The classic presentation occurs in less than 10% of children.[31]

With a retropharyngeal space infection, lateral neck radiograph findings include air in the soft tis-

sues, loss of the normal curvature of the cervical spine, and abnormal widening of the prevertebral soft tissue. Radiologic interpretation of this area in children is difficult because of difficulty in positioning, redundancy of normal soft tissue, and variation with inspiration and expiration. Computed tomography with contrast will confirm the presence of an abscess and determine its extent.

In young children, the airway may be obstructed and must be secured. If a retropharyngeal abscess is present, surgical drainage is indicated. Most abscesses can be drained intraorally. To avoid aspiration, the patient is placed in the Rose position (head down), and the neck is extended.

Lateral pharyngeal space infections occur later in childhood than retropharyngeal infections. In children, pharyngotonsillitis is the most common source, whereas odontogenic infections are the most common cause in adults. Infections may easily spread to and from this space.

Patients present with fever, sore throat, odynophagia, dysphagia, trismus, cervical pain, limitation of neck movement, and torticollis. Prominent signs include bulging of the lateral pharyngeal wall, anteromedial displacement of the tonsil, and cervical edema or lymphadenopathy. Respiratory compromise is a concern, especially in younger children. Deficits of cranial nerves IX, X, and XII and Horner's syndrome may occur.

Computed tomography with contrast should be performed in any patient with a suspected infection of the lateral pharyngeal space. It will demonstrate signs of infection and differentiate between cellulitis and abscess. The position, size, and extent of the abscess will be established, as well as its relationship to the great vessels.

If an abscess is identified, drainage is required. Traditionally, the external, cervical approach had been used exclusively to drain these abscesses because it provides excellent exposure, but with the development of CT, many abscesses can be drained transorally. If the CT demonstrates that the abscess is medial to the great vessels, drainage of the abscess can be performed safely and effectively through a transoral approach.[32] The transoral approach is technically easier to perform and is associated with less morbidity and shorter hospitalization. The external, cervical approach remains the approach of choice for abscesses that dissect along or are lateral to the great vessels.

DIFFERENTIAL DIAGNOSIS OF A NECK MASS

The differential diagnosis of a neck mass in adults and children is significantly different. In adults, the incidence of malignancy may be as high as 50%, and the protocol for the evaluation of a neck mass is well established.[33] Although malignancy is always a concern in a child, the incidence is low, and there are no established guidelines for the diagnosis of a pediatric neck mass.[34]

The age of the patient often suggests the type of lesion (Table 44–1). In children, the most common causes of neck masses are inflammatory lymph nodes from viral and bacterial infections. Congenital lesions are also common. Primary malignancies such as Hodgkin's and non-Hodgkin's lymphoma and leukemia occur, but their incidence is low.[35] In adolescents and adults to 40 years, congenital and developmental lesions and sialadenopathies are more common, whereas inflammatory lesions and malignancies are less frequent. In adults over 40 years, the first consideration must be a metastatic neoplasm.

A complete history often helps to narrow the diagnostic possibilities. A rapidly enlarging mass is suggestive of an inflammatory lesion or malignancy. The appearance of a mass after an upper respiratory tract infection or dental treatment indicates an inflammatory process. Constitutional symptoms of fever, weight loss, night sweats, and fatigue suggest a systemic condition such as a mycobacterial infection or lymphoma. A history of foreign travel or exposure to animals and insects makes infection likely.

A thorough head and neck examination is essential. Cervical lymphadenitis is usually the result of an infection elsewhere in the head and neck. Anterior cervical lymphadenitis is usually associated with oropharyngeal infections, whereas nasopharyngeal and paranasal sinus infections frequently cause posterior cervical lymphadenitis. In adults, all of the mucosal surfaces of the upper aerodigestive tract should be examined. Flexible fiber-optic endoscopes have greatly enhanced the ability to examine these mucosal surfaces in the office setting. If there are any questions regarding any lesions, the Storz-Hopkins telescopes can be used. Although flexible endoscopy is not necessary in all children, it should be attempted in children in whom the diagnosis is not apparent and in all adolescents presenting with pos-

TABLE 44–1. Differential Diagnosis of a Neck Mass*

Age (y)		
0–15	*16–40*	*41+*
Congenital/developmental		
Thyroglossal duct cyst	Branchial cleft cyst	Lymphangioma
Dermoid	Thymic cyst	
Laryngocele	Sialadenopathy	
	Pharyngeal diverticulum	
Inflammatory		
Lymphadenitis, viral, bacterial	Lymphadenitis, viral, bacterial	Lymphadenitis, viral, bacterial
Granuloma	Granuloma	Granuloma
	Sialadenitis	
Neoplastic		
Thyroid	Lymphoma	Lymphoma
Lymphoma	Metastatic cancer	Metastatic cancer

*Does not include neurogenic or vascular masses.

terior cervical masses to evaluate for the presence of nasopharyngeal carcinoma (see Chapter 60).

All neck masses should be examined for size, multiplicity, laterality, tenderness, color, mobility, consistency, surrounding tissue changes, and the presence of bruits and thrills. Soft, compressible masses are characteristic of congenital cysts, hemangiomas, and lymphangioma. Lymphomas often present as firm, rubbery, nontender masses. Palpable cervical lymph nodes are unusual in adults but common in children. Localized unilateral cervical lymphadenopathy is usually associated with gram-positive bacterial infections, whereas viral infections usually cause bilateral or generalized lymphadenopathy. In a child, any lymph node larger than 2 cm, especially if it is enlarging, should be evaluated. The location of the mass is especially important. Specific lesions have a predilection for specific sites, and the location may point to a benign or malignant condition. Congenital lesions are usually found in the midline or anterior cervical triangle. Metastatic malignancies are distributed along the pathways of the cervical lymphatic channels. In children, a supraclavicular mass is suggestive of a lymphoma.[36]

After completing the history and physical examination, the diagnosis may be apparent, and no additional studies may be needed. If it is not established, diagnostic studies can be performed. These may include blood and urine studies, skin tests, plain radiographs, ultrasonography, CT, MRI, biopsy, and endoscopy.

In a child with a suspected malignancy, chest radiography should always be performed because many children with lymphoma of the neck have abnormal chest radiographs.[36] Both CT and MRI can identify the consistency of a neck mass and delineate its size and extent. Magnetic resonance imaging has several advantages and is preferred over CT for the evaluation of cervical masses.[37] Magnetic resonance imaging is more sensitive to changes in contrast between the mass and normal tissue, providing superior resolution. It demonstrates better anatomic detail, has multiplanar capability, and involves no radiation exposure. Computed tomography and MRI can be combined with contrast to demonstrate vascular lesions.

If no diagnosis can be established or a malignancy is suspected in a child, an excisional biopsy is preferred. With localized lesions, an excisional biopsy may be curative and provides for a sufficient amount of tissue for additional studies, such as permanent and frozen section, electron microscopy, and tumor

markers. Although needle aspiration biopsies are safe and accurate in adults, they often do not provide enough tissue to diagnose pediatric malignancies. These malignancies are rare and frequently consist of embryonal tissue. Most pathologists do not have sufficient experience with these tumors. In many cases, diagnosis requires special stains, electron microscopy, and tumor markers. Many biopsies are performed because of a suspicion of malignancy, but in most cases, no malignancy is found. If there is a high suspicion of malignancy even after a negative biopsy, a repeat biopsy should be performed.

A needle biopsy is preferred in adults if a biopsy is required for diagnosis. Open biopsy is not recommended because this interrupts both the tumor and lymphatic channels. A primary malignancy and metastatic neck lesion should be treated in a coordinated fashion. Coordinated treatment decreases the risk of wound necrosis, local recurrence, and distant metastasis.[38] If the diagnosis is a metastatic malignancy, endoscopy of the entire upper aerodigestive tract and biopsies of any suspicious mucosal areas and any likely primary sites should be performed.

NASOPHARYNX

Anatomy

The nasopharynx extends from the skull base to the soft palate. The nasopharynx is surrounded by bone. Its superior and posterior walls are continuous and formed by the body of the sphenoid bone, basilar process of the occipital bone, and first two cervical vertebrae. The adenoid tissue (pharyngeal tonsil) lies on the posterosuperior surface. The choanae form its anterior limit.

The eustachian tube passes laterally through the sinus of Morgagni, a defect just above the superior edge of the superior constrictor muscles, and its orifice is located along the lateral wall, just posterior to the choanae. The torus tubarius, which is a prominent cartilaginous elevation, marks the posterior boundary of the orifice. The salpingopharyngeal fold extends inferiorly from the torus tubarius (Figure 44–15). The fossa of Rosenmüller is a deep recess just posterior to the torus tubarius. Gerlach's tonsil, lymphoid tissue, is found around the orifice.

The anterior aspect of its inferior border is the soft palate. The isthmus, the narrowest portion of the nasopharynx, opens into the oropharynx (Figure 44–16).

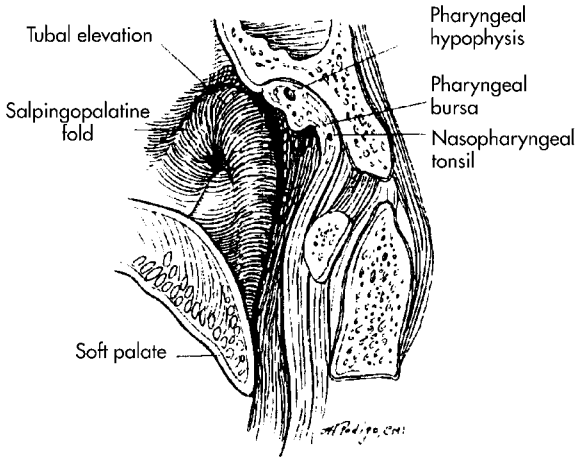


FIGURE 44–15. Sagittal view of the nasopharynx.

Physiology

The nasopharynx functions as a conduit for respired air and secretions from the eustachian tube and nose to pass to the oropharynx.

During swallowing and articulation of speech sounds represented by consonants, except m, n, and ng, the nasopharynx is separated from the oropharynx. During deglutition, there is circular closure of the isthmus by the coordinated actions of several muscles. The palatopharyngeus and salpingopha-

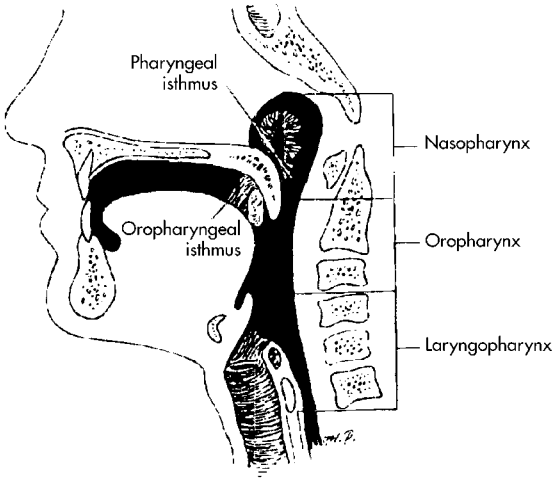


FIGURE 44–16. The regions of the pharynx are diagrammatically depicted.

ryngeus muscles contract to elevate the soft palate. The levator veli palatini lifts the muscular portion of the soft palate upward and backward. Contraction of the superior constrictor muscle narrows the nasopharynx laterally and forms Passavant's pad, a prominence of the posterior pharyngeal wall. The tensor veli palatini contributes to the closure of the isthmus during swallowing. Failure to separate the nasopharynx from the oropharynx results in velopharyngeal insufficiency. There may be reflux of fluid into the nasal airway or difficulty producing speech sounds such as "k."

Along with the nose, the nasopharynx is a resonating chamber that contributes to the quality of the voice. Loss of the resonating capacity produces a hyponasal voice or rhinolalia clausa. For certain vowels, proper velopharyngeal opening is necessary to avoid non-nasal speech. The nasopharynx acts as a resonating chamber for these phonemes.

The eustachian tube, which opens into the nasopharynx, is responsible for protection, ventilation, and clearance of fluid from the middle ear cleft. Proper function of the levator veli palatini and tensor veli palatini is essential for normal eustachian tube function. Children with anatomic abnormalities of the palate such as cleft palate and dysfunction of these muscles are likely to have problems with the middle ear cleft. Enlarged, nasopharyngeal lymphoid tissue may also affect eustachian tube function.

Examination

Evaluation of the nasopharynx can be done by indirect and direct examination. Indirect mirror examination is limited by patient cooperation and mirror distortion of the image. Nasopharyngoscopy, either with a flexible or rigid telescope, provides excellent visualization of most of the important structures, including the adenoid, eustachian tubes, and choanae. Direct examination in an operating room affords the best method of examination. The patient is in the same position as for an adenoidectomy. The nasopharynx may be visualized with mirrors and telescopes. Any suspicious areas may be palpated and even biopsied.

Benign Lesions of the Nasopharynx

These lesions are discussed in Chapter 46 and will be mentioned briefly.

Rathke's pouch develops when remnants of the buccal-mucosal invagination that forms the anterior lobe of the pituitary gland persist. A craniopharyngoma is a tumor that arises from these remnants. Symptoms take the form of pituitary-hypothalamic derangements.

The pharyngeal bursa of Thornwaldt is a midline nasopharyngeal structure located just inferior to Rathke's pouch and represents the proximal end of the notochord. It may form a cyst known as Thornwaldt's cyst. The cyst may become infected.

Teratomas, dermoid cysts, and gliomas are rare but may cause nasal obstruction and respiratory distress in newborns. Teratomas are often associated with other congenital abnormalities. Mucous retention cysts and hamartomas occur occasionally.

Choanal Atresia

Choanal atresia results from the failure of the buccopharyngeal membrane to regress so that the posterior part of the nasal cavity does not open into the nasopharynx. This developmental anomaly is discussed in Chapter 46.

Juvenile Angiofibroma

Juvenile angiofibroma of the nasopharynx is rare but is the most common benign neoplasm of the nasopharynx. It is discussed in Chapter 60.

Nasopharyngeal Stenosis

Nasopharyngeal stenosis occurs when scar tissue obstructs the nasopharynx.[39] The posterior tonsillar pillars and soft palate adhere to the posterior pharyngeal wall. The degree of scarring varies from a thin band to a complete cicatrix, and the obstruction may be partial or complete. Most cases result from tonsillectomy and adenoidectomy, but some cases are secondary to infection.

Nasopharyngeal stenosis may develop when inferolateral adenoid tissue is removed and there is simultaneous denuding or excision of the posterior tonsillar pillar.[24] Patients who undergo secondary adenoidectomy and have large lateral bands of lymphoid tissue are more likely to develop this stenosis. Treatment is dependent on the location and severity of the stenosis and includes corticosteroid injection, lysis of adhesions, and rotation and advancement flaps.

SLEEP APNEA

Sleep apnea is an increasingly recognized and important condition in adults. It has been estimated that 20 million Americans experience significant daytime somnolence because of sleep apnea.[40] Although it is more common in males, a significant number of women also suffer from this condition. A healthy adult may experience apneic periods of 10 seconds or less several times an hour during sleep, without significant clinical signs or symptoms. People with significant manifestations may have 15 or more events per hour, each lasting 20 to 30 seconds or longer.

Obstructive Sleep Apnea

It is estimated that 12 million Americans have OSA.[41] In OSA, the airway obstruction is caused by pharyngeal collapse, which is the end result of complex interactions between structural and neuromuscular factors. In healthy individuals, the patency of the pharynx is maintained by tonic activity of the pharyngeal dilatory muscles. Sleep results in a decrease in the behavioral responses, which maintain this tone, and relaxation narrows the airway.

The etiology of OSA is multifactorial and incompletely understood. Reduced neuromuscular control of the pharyngeal dilator muscles accentuates the physiologic pharyngeal narrowing, further reducing the lumen. Alcohol ingestion, which exaggerates the relaxation of the pharyngeal dilators, and neuromuscular diseases such as muscular dystrophy, myasthenia gravis, cerebrovascular accidents, and olivopontocerebellar degeneration are well known causes of OSA. Anatomic abnormalities cause obstruction, and with obstruction, greater inspiratory pressure is needed to generate airflow. The resulting increased negative pressure also contributes to pharyngeal collapse. Structural abnormalities often associated with OSA include nasal septal deviation, turbinate hypertrophy, elongated soft palate, adenoid and tonsil hypertrophy, macroglossia, retrognathia, and micrognathia. Cardiac disease, chronic obstructive pulmonary disease, obesity, hypothyroidism, acromegaly, gastroesophageal reflux, androgens, caffeine, and tobacco are other factors that may predispose or exacerbate OSA.

Common characteristics of individuals with OSA include male sex, advancing age, hypertension, truncal obesity, and neck circumference greater than 17 inches. The most common symptoms of OSA are snoring and excessive daytime somnolence. Other symptoms include frequent nocturnal arousal periods, restless sleep, memory loss, fatigue, irritability, morning headache, sexual dysfunction, and nocturia.

Although assessment of the upper airway is critical, examination of the patient with suspected OSA is not limited to this region. Predisposing and associated factors must be sought. Assessment of the patient's general medical condition, height, weight, blood pressure, neck size, and cardiopulmonary status is mandatory. The entire upper airway is thoroughly examined, and it is not uncommon to find more than one site of obstruction. Flexible fiberoptic endoscopy is essential. It provides for better inspection of the airway. Patients can be examined either erect or supine, and evaluation of the patient at rest with the mouth closed more closely simulates sleep. The Müller maneuver, fiber-optic examination during inspiration with the oral and nasal cavities closed, can be performed to help assess the pattern and degree of pharyngeal collapse.

Several imaging techniques are available to evaluate the structure and function of the upper airway. They include cephalometry, somnofluoroscopy, CT, and MRI. These imaging modalities are primarily considered investigational tools, and their clinical applications are evolving. Magnetic resonance imaging with three-dimensional reconstruction appears useful in the preoperative assessment of candidates for uvulopalatopharyngoplasty, whereas CT with three-dimensional reconstruction may be considered for candidates of maxillomandibular advancement.[42]

Polysomnography in a sleep laboratory is the definitive diagnostic study for OSA. Other studies have been developed either as screening devices or less expensive alternatives. Portable polysomnography, a convenient, inexpensive study, can be performed at home, but because it presently lacks universally accepted standards, it is not considered definitive. Nocturnal oximetry measures the oxygen levels during sleep. In children, sleep sonography provides objective data regarding the presence and degree of upper airway obstruction, but its ability to predict apnea is not established. The multiple sleep latency test does not diagnose OSA but measures daytime somnolence.

Treatment of OSA begins with management of any predisposing or associated factors. Although a

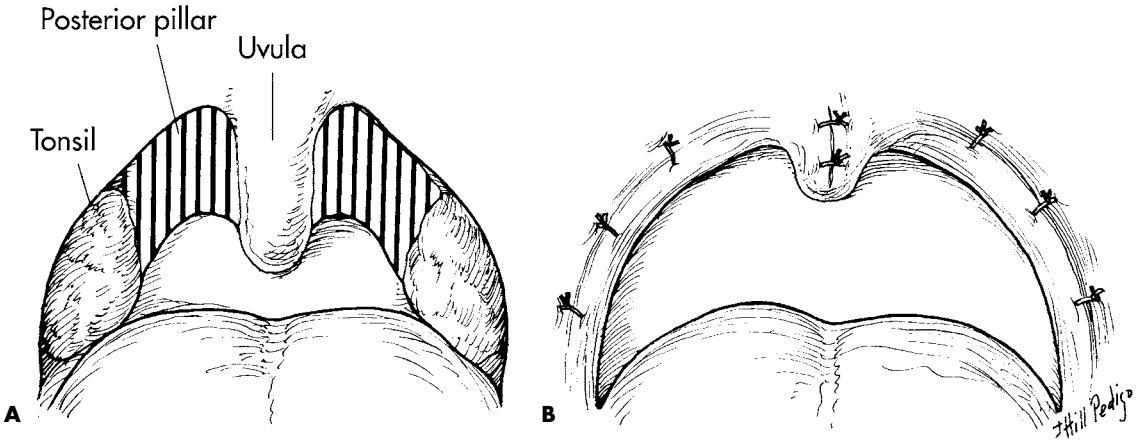


FIGURE 44–17. Diagram of uvulopalatopharyngoplasty. *A*, Obstruction to the oropharyngeal airway is caused by abnormally prominent posterior tonsillar pillars and the large, bulky uvula. *B*, An improved airway has been created by reduction in the size of these structures.

patient may not be obese, moderate weight reduction may improve the severity of the apnea. Alcohol, tobacco, and caffeine should be eliminated.

Nonsurgical treatment of OSA includes positioning, artificial airways, appliances, and positive airway pressure. The supine position often exacerbates OSA, and there are maneuvers and devices that can help one avoid this position. Nasopharyngeal and oral airways, tongue-retaining devices, and mandible advancement appliances may help improve snoring and OSA but are usually not well tolerated. Continuous positive airway pressure (CPAP) and biphasic positive airway pressure (BiPAP) provide pneumatic stenting of the pharyngeal airway. Their face mask delivery systems are cumbersome and discomforting.

The goal of surgery for OSA is relief of airway obstruction by increasing pharyngeal size, decreasing pharyngeal compliance, or both to maintain adequate airflow. The specific procedure depends on the site of obstruction, and consideration includes nasal surgery, tonsillectomy, adenoidectomy, palatal surgery, tongue base reduction, genioglossal advancement with hyoid myotomy, and maxillomandibular osteotomy and advancement. In severe, acute cases, tracheostomy may be considered.

The most common surgery performed for OSA is palatal surgery. Redundant tissue of the soft palate, uvula, tonsil pillars, and lateral pharyngeal walls causes obstruction, and reduction in the size of these structures enlarges the airway. Uvulopalatoplasty and uvulopalatopharyngoplasty (Figure 44–17) are the most widely performed procedures. Uvulopalatoplasty reduces the soft palatal and uvular tissue. It is an outpatient procedure that is often performed with a laser, and recently, radiofrequency energy ablation of these tissues has been used. Uvulopalatopharynoplasty, a more complex procedure, eliminates the palatal, tonsil pillar and pharyngeal redundancy by removing excess mucosal and submucosal tissue. Submucosal resection of the palatoglossus and palatopharyngeus muscles and removal of the posterior portion of the hard palate with advancement of the soft palate are other procedures that have been proposed.[43]

Central Sleep Apnea

Central sleep apnea is much less common than OSA and must be distinguished from it. Central sleep apnea occurs when the neural drive to the respiratory muscles is temporarily abolished, resulting in an absence of respiratory effort. There is little effort to breathe. During sleep, respiration is controlled by an automatic feedback system. Sensory signals from multiple receptors are relayed to the brainstem, where they are integrated, and the brainstem sends signals, which stimulate the muscles of respiration and lungs.

Any condition affecting this feedback system may cause CSA. Congenital central alveolar hypoventilation, which results from a very low sensitivity to increased PCO_2, is a cause of CSA. Chronic neuropathies may also impair the sensory component. Bilateral brainstem lesions from vascular con-

ditions, infections, and degenerative and metabolic diseases affect integrative function and cause CSA. Neuromuscular diseases, such as polio, amyotrophic lateral sclerosis, and muscular dystrophies, impair motor function and may cause CSA. In addition, chronic obstructive pulmonary disease and congestive heart failure are causes of CSA.

With polysomnography, CSA can be diagnosed, and its degree of severity and any associated physiologic changes can be determined. Central sleep apnea can also be confirmed by measuring esophageal pressure, an index of respiratory effort, during sleep.[44]

The primary treatment for CSA is nasal mask ventilation using CPAP, BIPAP, or intermittent positive pressure ventilation. Medications to stimulate respiration and to increase the tone of the genioglossus and geniohyoid muscles have had limited success. In the most severe cases, tracheostomy with ventilation may be required.

REFERENCES

1. Sharp JT. Reiter's syndrome. Curr Prob Dermatol 1973;5:157–79.
2. Handlers JP, Waterman J, Abrams AM, et al. Oral features of Wegner's granulomatosis. Arch Otolaryngol Head Neck Surg 1985;111:267–70.
3. Hurwitz S. Erythema multiforme: a review of its characteristics, diagnostic criteria, and management. Pediatr Rev 1990;11:217–22.
4. Lasarkis G, Stoufi E. Oral pemphigus vulgaris in a 6-year-old girl. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1990;69:609–13.
5. Yoskovitch A, Twefik TL, Duffy CM, et al. Head and neck manifestations of Kawasaki disease. Int J Pediatr Otorhinolaryngol 2000;52:123–9.
6. Campbell SM, Montanaro A, Bardana EJ. Head and neck manifestations of autoimmune diseases. Am J Otolaryngol 1983;4:187–216.
7. Richardson HA, Leibovitz A. Hand, foot, and mouth disease in children. J Pediatr 1965;67:6–12.
8. Terezhalmy GT. A checklist of common oral manifestations of HIV infection. Cleve Clin J Med 1992; 59:562–5.
9. Lasarkis G. Color atlas of oral diseases. New York: Thieme Medical Publishers; 1988.
10. Hadfield PJ, Birchall MA, Novelli V, et al. The ENT manifestations of HIV infection in children. Clin Otolaryngol 1996;21:30–6.
11. Rizer FM, Schechter GL, Richardson MA. Macroglossia: etiologic considerations and management techniques. Int J Pediatr Otorhinolaryngol 1985;8: 225–36.
12. Bartley JR, Morton RP. Simple sublingual ranulas: pathogenesis and management. J Otolaryngol 1995; 24:253–4.
13. Lawson W, Reino AJ. Odontogenic infections. In: Bailey BJ, editor. Head and neck surgery-otolaryngology. 2nd ed. Philadelphia: Lippincott-Raven Press; 1998. p. 671–81.
14. Gonzalez C. Tumors of the mouth and pharynx. In: Bluestone CD, Stool SE, Kenna MA, editors. Pediatric otolaryngology. Philadelphia: WB Saunders; 1996. p. 1108–20.
15. Mitchelmore IJ, Reily PG, Hay AJ, et al. Surface and core cultures in recurrent tonsillitis: prevalence and anaerobes and beta-lactamase producing organisms. Eur J Clin Microbiol Infect Dis 1994;13:542–8.
16. Brook I. Treatment of patients with acute recurrent tonsillitis due to group A beta-haemolytic streptococci: a prospective randomized study comparing penicillin and amoxicillin/clavulante potassium. J Antimicrobial Chemother 1989;24:227–33.
17. Dedio RM, Tom LWC, McGowan KL, et al. Microbiology of the tonsils and adenoids in a pediatric population. Arch Otolaryngol Head Neck Surg 1988;114: 763–5.
18. Clinical indicators 2000. Am Acad Otolaryngol Head Neck Surg Bull 2000;19:19.
19. Herzon FS. Pertitonsillar abscess: incidence, current management practices, and a proposal for treatment guidelines. Laryngoscope 1995;105 Suppl 74:1–17.
20. Clinical indicators 2000. Am Acad Otolaryngol Head Neck Surg Bull 2000;19:34.
21. Marsh RR, Potsic WP, Pasquariello PS. Reliability of sleep sonography in detecting upper airway obstruction in children. Int J Pediatr Otorhinolaryngol 1989; 18:1–8.
22. Gates GA, Aver CS, Prihoda TJ, et al. Effectiveness of adenoidectomy and tympanostomy tubes in the treatment of chronic otitis media with mucoid effusion. N Engl J Med 1987;317:1444–51.
23. Paradise JL, Bluestone CD, Rogers KD, et al. Efficacy of adenoidectomy for recurrent otitis media in children previously treated with tympanostomy tube placement. JAMA 1990;262:2066–73.
24. McLaughlin KE, Jacobs IN, Todd NW, et al. Management of nasopharyngeal and oropharyngeal stenosis in children. Laryngoscope 1997;107:1322–31.

25. Collier FA, Yglesias L. The relation of the spread of infection to fascial planes in the neck and thorax. Surgery 1937;1:323–37.
26. Hollinshead W, editor. Anatomy for surgeons. Vol 1, 3rd ed. Philadelphia: Harper & Row; 1982. p. 269–89.
27. Ungkanont K, Yellon RF, Weissman JL, et al. Head and neck space infections in infants and children. Otolaryngol Head Neck Surg 1995;112:375–82.
28. Tschiassny K. Ludwig's angina. Arch Otolaryngol 1943;38:485–96.
29. Blinder D. Idiopathic submandibular abscesses in children. Int J Oral Maxillofac Surg 1986;15:292–5.
30. Barratt GE, Koopman C, Couthard S. Retropharyngeal abscess—a ten-year experience. Laryngoscope 1984;94:455–63.
31. Thompson JW, Cohen SR, Reddix P. Retropharyngeal abscess in children: a retrospective and historical analysis. Laryngoscope 1988;98:589–92.
32. Nagy M, Pizzuto M, Blackstrom J, et al. Deep neck infections in children: a new approach to diagnosis and treatment. Laryngoscope 1997;107:1627–34.
33. Lee JG, Helmus C. Cervical lymph node biopsy. Mich Med 1970;69:581–3.
34. Wetmore RF, Tom LWC. An approach to the pediatric neck mass. In: Johnson JT, Derkey C, Mendall-Brown MK, Newman RK, editors. Instructional courses. St. Louis: Mosby Year Book; 1991. p. 160–8.
35. Jaffe BF. Pediatric head and neck tumors: a study of 178 cases. Laryngoscope 1973;83:1644–51.
36. Torsiglieri AJ, Tom LWC, Ross AJ, et al. Pediatric neck masses: guidelines for evaluation. Int J Pediatr Otorhinolaryngol 1988;16:199–211.
37. Mafee MF, Campos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.
38. McGuirt WF, McCabe BF. Significance of node biopsy before definitive treatment of cervical metastatic carcinoma. Laryngoscope 1978;88:594–7.
39. Stevenson EW. Cicatrical stenosis of the nasopharynx. Laryngoscope 1968;79:2035–67.
40. Wake up America: a national sleep alert. Vol 1. Report of the National Commission on Sleep Disorders Research. January 1993.
41. Young T, Palta M, Demsey J, et al. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 1993;328:1230–5.
42. Schwab RJ, Goldberg AN. Upper airway assessment: radiographic and other imaging techniques. Otolaryngol Clin North Am 1998;31:931–68.
43. Woodson BT. Transpalatal advancement pharyngoplasty for obstructive sleep apnea. Oper Tech Otolaryngol Head Neck Surg 2000;11:36–40.
44. Guilleminault C, Robinson A. Central sleep apnea. Otolaryngol Clin North Am 1998;31:1049–65.

CHAPTER 45

Congenital Anomalies of the Larynx

Rodney P. Lusk, MD

Congenital lesions of the larynx are relatively rare; however, in their laryngeal laboratory, Chen and Holinger found congenital laryngeal lesions in 33 of 115 specimens.[1] Of all of the anomalies noted, the most common deformity is of the cricoid cartilage. The lesions develop during respiratory differentiation, which occurs during the fourth to tenth weeks of gestation. The clinical manifestations of laryngeal lesions fall into three broad categories:

1. *Respiratory distress*, which may range from complete obstruction with no air movement to minor types of stridor. The characteristics of the stridor are variable and depend on the site and degree of the obstruction. Most laryngeal lesions produce stridor during the inspiratory phase. The reason for this is found in Bernoulli's principle, which states that when a fluid or a gas is in motion, the pressure exerted on the wall of the airway decreases as the velocity of the gas increases. A well-known example is the wing of an airplane, shown in Figure 45–1. When the site of the obstruction is in an area not firmly fixed or supported, such as the supraglottis or larynx, as the air rushes into the airway, it produces a relatively negative pressure, which tends to close the airway (Figure 45–2). The harder the child breathes in, the faster the flow and the greater the created negative pressure, with the net effect of further decreasing the airway lumen. The negative pressure generated during inspiration may also cause supraclavicular and sternal retractions and be associated with nasal flare.
2. *Dysphonia*, which may be caused by laryngeal lesions that interfere with vocalization, with voice quality ranging from complete aphonia to hoarseness.
3. *Failure of the larynx to close the airway* during swallowing, which may cause feeding difficulties, with aspiration, cyanosis, or respiratory compromise.

Although this chapter deals only with laryngeal lesions and airway distress, any airway lesion may result in stridor, tachycardia, tachypnea, cyanosis, or restlessness. When evaluating patients with these signs, a history and examination of the entire airway should be performed. The history and physical examination are the most important diagnostic tools. Radiographic imaging (plain films, computed tomography [CT], and magnetic resonance imaging [MRI]) and laryngoscopy (direct and indirect) are important adjuncts and are discussed in each section.

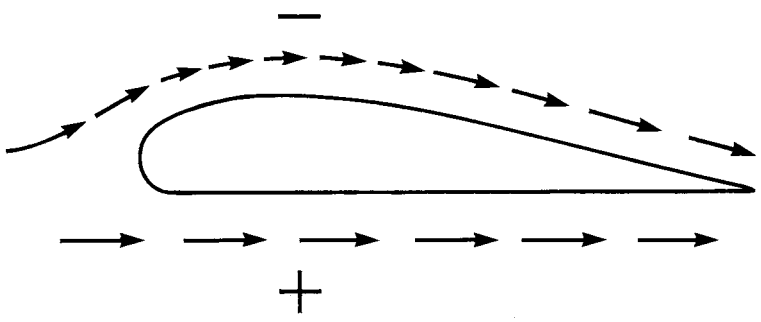


FIGURE 45–1. For a wing to produce lift, the air pressure on the top of the wing must be less than the pressure on the bottom. According to Bernoulli's principle, this can occur only when the velocity of the air is greater over the top of the wing, which is accomplished by increasing its upper curvature.

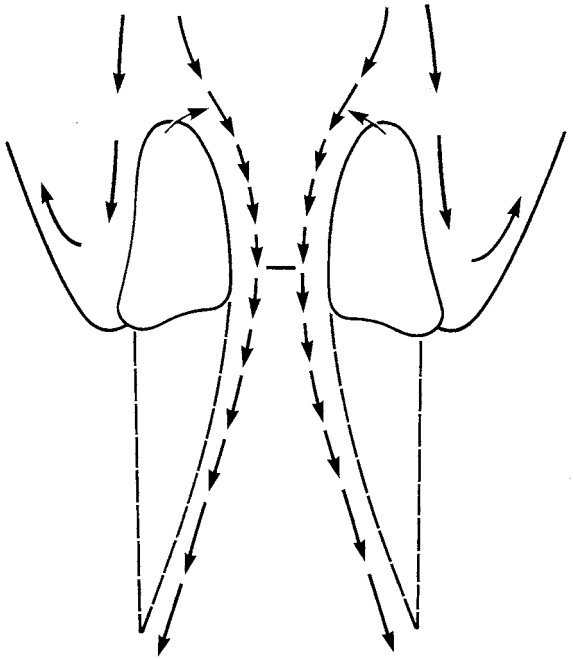


FIGURE 45–2. Demonstration of the arytenoids and the subglottis as compared to an air foil. Note that as the velocity of the air increases in the glottis, it creates negative pressure in the airway, pulling the cuneiform cartilages and aryepiglottic folds into the airway.

LARYNGOMALACIA

Laryngomalacia is a condition in which the laryngeal inlet collapses on inspiration, causing stridor. The terminology is confusing, with congenital laryngeal stridor,[2] congenital laryngeal obstruction,[3] congenital stridor of infants,[4] and inspiratory laryngeal collapse[5] all used synonymously. The term laryngomalacia, introduced by Jackson and Jackson in 1942,[6] is most frequently used. Some physicians confuse this entity with tracheomalacia, which is another disease altogether.

Incidence and Etiology

Laryngomalacia is the most common cause of stridor in the newborn. In some children, it may be an autosomal dominant lesion.[7] Hawkins and Clark evaluated 453 patients with the flexible fiberscope and found 84 with primary and 29 with secondary laryngomalacia.[8] Nussbaum and Maggi found that 68% of 297 children with laryngomalacia had other respiratory disorders, as noted with flexible bronchoscopy.[9]

Signs and Symptoms

The most common symptoms produced by laryngomalacia are inspiratory stridor, feeding problems, and gastric reflux. The inspiratory stridor worsens with increased respiratory effort, such as crying or feeding, and in the supine position. The symptoms begin shortly after birth in most patients,[10] increase in severity until 6 to 8 months of age, plateau at 9 months, and steadily improve thereafter.[5,11,12] No correlation exists between the duration of stridor and the severity or time of onset.[10]

Two theories have been proposed to account for the narrowing of the laryngeal inlet. The first is that a neuromuscular abnormality, by not allowing proper support of the supraglottis, causes increased flaccidity.[13] Kelemen suggested that the cause is ineffective dilators of the supraglottis,[14] but dissections by Belmont and Grundfast showed the dual insertion of the palatopharyngeus, lateral glossoepiglottic, and stylopharyngeal muscles, which could dilate the laryngeal inlet.[13] The incidence of neuromuscular disorders appears to be higher in children with laryngomalacia.[11] Reports of higher incidence of mental retardation in these patients[11] have not been substantiated.[10,13]

The other proposed mechanism is anatomic, with the narrowing resulting from (1) a flaccid epiglottis that folds against the posterior laryngeal wall or into the airway, (2) a long, tubular epiglottis that curls on itself, or (3) short and redundant aryepiglottic folds with varying sizes of cuneiform cartilage that rotate medially into the airway[15] (Figures 45–3 and 45–4). The W-shaped epiglottis is not thought to be a significant factor because it occurs in 30 to 50% of patients,[3,16–18] most of whom do not have stridor.

Feeding problems are frequent with laryngomalacia. Radiographic evidence of gastric reflux occurs in 80% and regurgitation in 40%[13] past 3 months of age. In the study by Zalzal and associates, 5 of 21 patients presented with feeding problems.[19] Some investigators believe that feeding problems are secondary to the high negative pressures generated during inspiration. Aspiration pneumonitis has been reported in 7% of children with laryngomalacia.[13] The mechanism for this is unclear but may be associated with the negative pressure and feeding problems.

Obstructive sleep apnea (23%) and central sleep apnea (10%) have also been noted.[13] Cor pulmonale can develop secondary to laryngomalacia.[19]

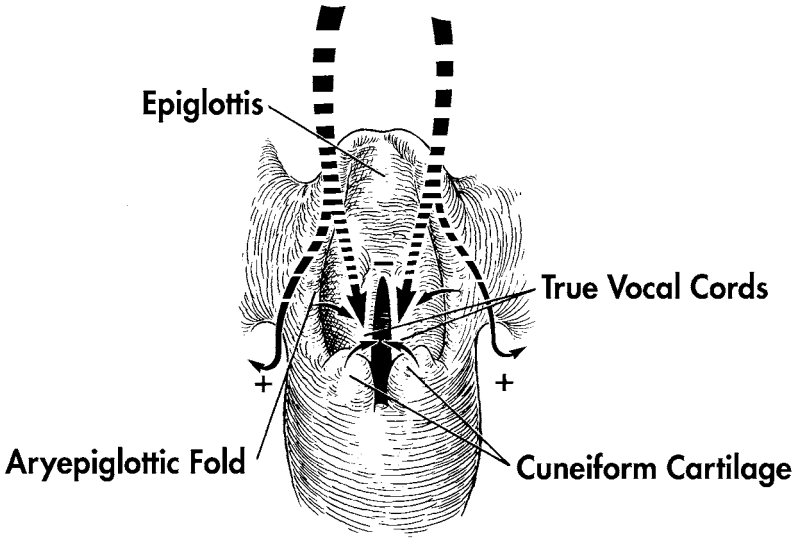


FIGURE 45–3. The airflow through the larynx creates a relative negative pressure and pulls the aryepiglottic folds and cuneiform cartilages into the airway.

Diagnosis

The diagnosis is most frequently based on the symptoms and signs, flexible laryngoscopy, and lateral radiographs of the neck and chest and rigid endoscopy to identify possible synchronous airway lesions such as subglottic stenosis. Hawkins and Clark have shown flexible laryngoscopy to be effective, even in neonates.[8] The advantage of using the flexible scope in an awake patient is that the dynamics of the supraglottis can readily be appreciated. The disadvantage is that the milder forms of laryngomalacia may be missed if the child is crying. The flexible scope can also be used to diagnose other synchronous glottic lesions such as vocal cord paralysis and glottic cysts. Mancuso et al retrospectively evaluated 233 patients with laryngomalacia to assess the necessity of performing a complete diagnostic workup in *all* patients.[20] Their conclusion was that rigid endoscopy in evaluation of an infant with laryngomalacia is rarely necessary because clinically significant synchronous airway lesions that require surgical intervention are rare. Historically, this has not been the case. Hawkins and Clark[8] and Nussbaum and Maggi[9] used flexible endoscopy to evaluate laryngomalacia, but their recommendation was that "all symptomatic" children be evaluated by bronchoscopy. Bluestone et al went "on record as recommending a complete evaluation of the tracheobronchial tree in all symptomatic infants."[21] The limitation of using a flexible scope is that it does not allow accurate assessment of the subglottis and trachea. The incidence of other lesions noted by additional authors has ranged from 18% in patients with congenital laryngeal lesions to 59.8% of all patients presenting with stridor.[22,23] The controversy lies in the likelihood of missing a clinically significant synchronous airway lesion that would otherwise have been treated if rigid endoscopy had been performed. Mancuso et al concluded that rigid endoscopy was not necessary in otherwise asymptomatic children.[20] Bluestone et al reached the opposite conclusion; based on "retrospective careers, [they] have encountered countless numbers of infants who have had clinically significant airway problems in whom the diagnosis of laryngomalacia was previously made by flexible fiberoptic laryngoscopy, but rigid endoscopy was omitted."[21] Deaths have been described in patients with severe disease.[24,25] This controversy is going to require additional data that are prospectively acquired. Currently, it would appear that if symptoms are not compatible with laryngomalacia alone or if the patient fails to thrive, direct laryngoscopy and rigid bronchoscopy should be performed to rule out secondary lesions. It is the surgeon's obligation to follow his or her patients closely until the symptoms resolve.

Treatment

Expectant observation is sufficient in most cases. The small percentage of patients (10 to 15%) with failure to thrive or more than one lesion may require surgical intervention. Iglauer, in 1922, was the first to

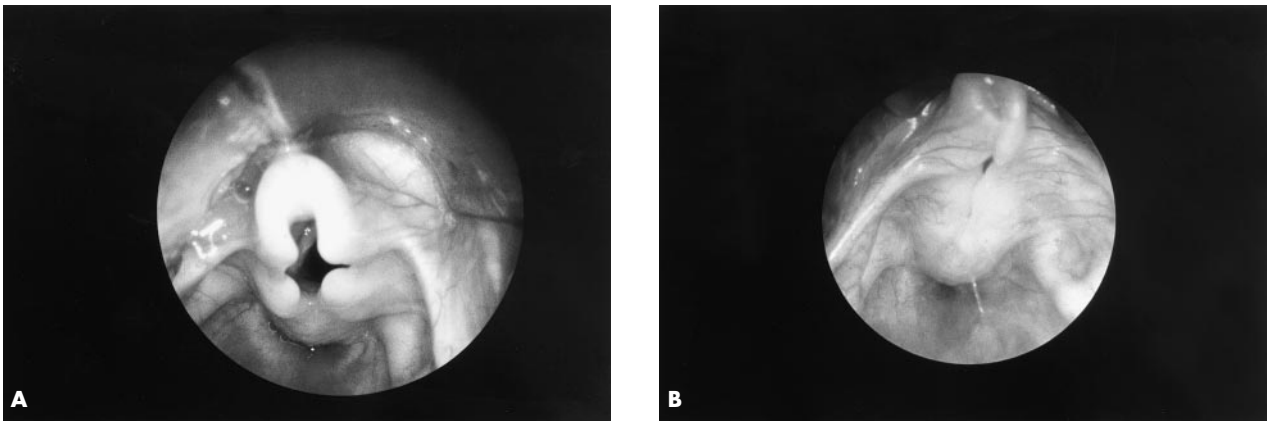


FIGURE 45–4. Example of laryngomalacia. *A*, Supraglottis at beginning of inspiration. *B*, Complete collapse of the supraglottis toward the end of inspiration. Courtesy of Dr. G. B. Healy, Boston's Children's Hospital, Boston, Massachusetts.

alter the supraglottis surgically in laryngomalacia.[26] Schwartz, in 1944, removed a V-shaped wedge from an epiglottis with good results.[5] Fearon and Ellis, in 1971, successfully treated a patient by suturing the epiglottis to the base of the tongue.[27] In that same year, hyomandibulopexy was advocated in France.[28,29] Templer and colleagues, in 1981, performed a supraglottectomy in an adult for longstanding stridor and obstructive sleep apnea secondary to laryngomalacia.[30] Lane and associates, in 1984, excised the tips of the arytenoids, edematous mucosa, a portion of the corniculate cartilages, and a portion of the aryepiglottic fold with microcupped forceps and Bellucci scissors.[31] Seid and colleagues successfully treated three patients with the laser in 1985.[32] Since then, numerous case reports have appeared in the literature.[33] Zalzal and associates trimmed the lateral edges of the epiglottis, aryepiglottic folds, and the corniculate cartilages with microlaryngeal scissors.[19] Solomons and Prescott recommended trimming the supraglottis and performing an anterior epiglottopexy.[16] The trimming of the aryepiglottic folds may involve both sides or only one side. Kelly and Gray recommended unilateral removal of redundant supraglottic tissue (supraglottoplasty) and to revise the other side if the patient is not asymptomatic.[34] We have recommended doing both sides at one setting. Either appears to give satisfactory results.

Holinger and Konior prefer the term supraglottoplasty to describe the surgical procedures for removing the flaccid supraglottic tissues.[15] These investigators used the laser to remove a portion of the cuneiform cartilage and the aryepiglottic folds (Figure 45–5). Care must be taken not to excise or traumatize the strip of mucosa between the arytenoids in the posterior part of the glottis. Holinger and Konior believe that prophylactic antibiotics should be used and that most patients may be extubated in the immediate postoperative period but should be observed overnight in the intensive care unit.[15]

ABSENT OR RUDIMENTARY EPIGLOTTIS

Incidence

The rare bifid epiglottis can be associated with stridor and aspiration. In the only thorough review of the world literature, Montreuil found only five clearly documented cases.[35] Since then, six additional cases have been reported.[36–38]

Signs and Symptoms

Patients have presented with respiratory distress secondary to rotation of the two halves of the epiglottis into the airway. The incidence of multiple congenital anomalies is high. A 44% incidence of polydactyly has been reported.[39] Graham and colleagues reported three cases of hypopituitarism.[37]

Diagnosis

The only means of diagnosing a bifid epiglottis is with flexible or direct laryngoscopy (Figure 45–6).

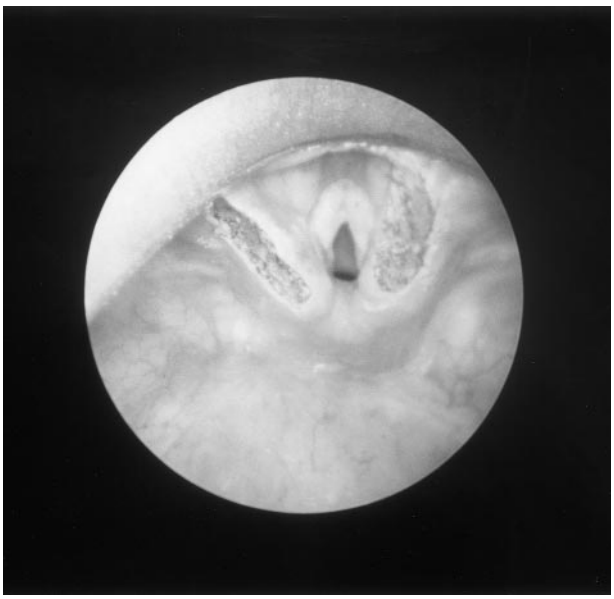


FIGURE 45–5. Example of supraglottoplasty with the laser. Note the sparing of the posterior commissure.

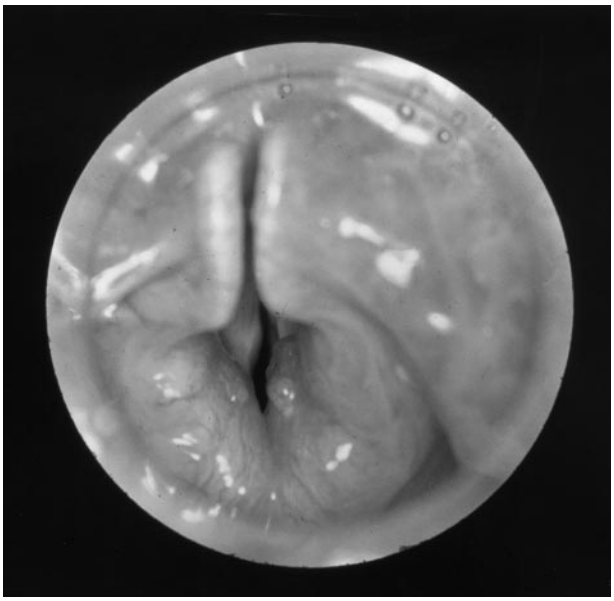


FIGURE 45–6. Bifid epiglottis. Courtesy of Dr. L. D. Holinger, Children's Memorial Hospital, Chicago, Illinois.

Because of the high incidence of associated airway lesions, direct laryngoscopy is suggested for all patients.

Treatment

If the airway distress is significant, a tracheostomy is warranted. Montreuil amputated the epiglottis with good results.[35] Healy and associates performed a tracheostomy in two patients, and in time the epiglottis matured, and both patients were decannulated without surgical intervention on the epiglottis.[39]

ATRESIA OF THE LARYNX

Atresia occurs when the laryngeal opening fails to develop and an obstruction is created at or near the glottis. Tracheal agenesis may also occur but is thought to be a rarer lesion.[40]

Incidence and Etiology

Congenital laryngeal atresia is a rare lesion, with only 51 cases reported in the world literature in 1987.[41] It is thought to be the rarest laryngeal lesion, accounting for only 1 of 846 congenital lesions evaluated by Holinger and colleagues[23,42] and none of 433 congenital lesions evaluated by Fearon and Ellis.[27] The lesion is thought to arise from the premature arrest of normal vigorous epithelial ingrowth into the larynx during the fourteenth to seventeenth stages of embryonic development.[43] In the embryonic larynx, a pharyngoglottic duct divides the larynx into an anterior (membranous) portion and a posterior (cartilaginous) portion.[44]

Signs and Symptoms

At delivery, the child makes strong respiratory efforts but does not move air, cry, or manifest any stridor. The child becomes markedly cyanotic when the umbilical cord is clamped. Most of these children die at birth unless an emergency tracheostomy is performed. Some children survive if they have a large tracheoesophageal fistula and if the esophagus is intubated. Not infrequently, a pharyngotracheal duct is present in the larynx. This should not be confused with a tracheoesophageal fistula, which is located much lower.

Diagnosis

The diagnosis is most frequently made at autopsy.[45] With the increased use of ultrasonography, the diagnosis of laryngeal atresia can be made before birth by noting enlarged edematous lungs, compressed fetal heart, severe ascites, and fetal hydrops.[46–50] Smith and Bain outlined three types of laryngeal

atresia deformities[51]; these investigators indicated that these types are not absolute but are gradations of a continuous spectrum. Type 1 involves the supraglottic and infraglottic larynx with fused arytenoids, an absent vestibule, and a deformed cricoid. Type 2 involves only the cricoid (subglottic area) with normal arytenoids, vestibule, and vocal cords. Type 3 involves a fused glottis with a normal vestibule and cricoid (Figure 45–7). The extent of the obstruction can be diagnosed with ultrasonography.[52] The subglottic lesions have been found to be cartilaginous or membranous.[44] When the lesion is diagnosed prenatally, it is called congenital high airway obstruction syndrome (CHAOS).[53] CHAOS results in a predictable constellation of findings: large echogenic lungs, flattened or inverted diaphragms, dilated airways distal to the obstruction, and fetal ascites and/or hydrops. Ultrasonography should be used to investigate the airway.[53]

Treatment

The patient will not survive unless an emergency tracheostomy is performed or the patient is ventilated through a tracheoesophageal fistula. Tracheoesophageal fistulae are frequently associated with this condition and usually arise at the tracheal bifurcation (Figure 45–8).[51,54,55] Most of the survivors have the membranous type 3 lesion or a type 2 lesion and a tracheoesophageal fistula, which allowed ventilation until a tracheostomy could be performed.

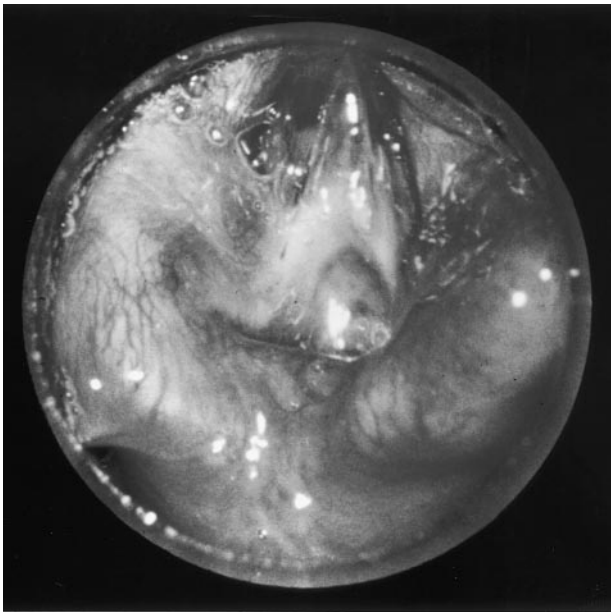


FIGURE 45–7. Type 3 laryngeal atresia. Courtesy of Dr. L. D. Holinger, Children's Memorial Hospital, Chicago, Illinois.

WEBS

Laryngeal webs occur in the glottis. Rarely, webs extend from the epiglottis to the lateral or posterior aspects of the hypopharynx and can result in significant airway compromise.[56] Holinger and associates described a patient with simultaneous supraglottic and laryngeal webs.[57] Most webs are located anteriorly and extend a varying length toward the arytenoids. The webs vary in thickness from a thin structure to one that is thicker and more difficult to eradicate. Benjamin and Mair described a rare type known as a congenital interarytenoid web, which limits the abduction of the arytenoids.[58]

Incidence

Congenital laryngeal webs are uncommon, constituting 5% of all congenital laryngeal lesions.[42,59] Acquired lesions are more common than congenital lesions, in a 60 to 40 ratio.[60] Associated anomalies are frequently seen, especially higher in the airway.[59,61,62] Posterior glottic webs occur in only 1 to 4% of patients.[60]

Signs and Symptoms

Symptoms of laryngeal webs are present at birth in 75% of patients and within 1 year in all patients.[62] The types of symptoms depend on the severity of

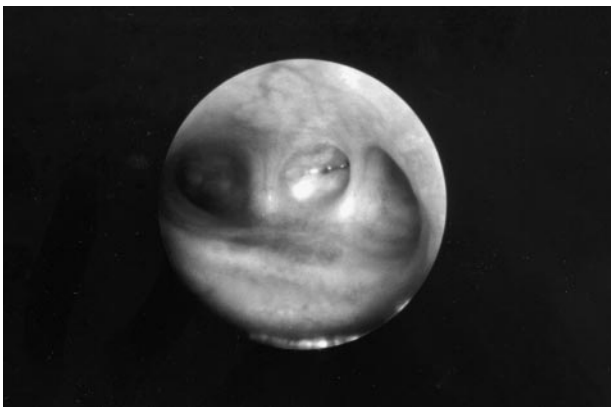


FIGURE 45–8. Tracheoesophageal fistula at the carina.

the web. Many children are asymptomatic until they are stressed, have an infection, or are intubated for a surgical procedure.

Vocal dysfunction is the most frequent symptom and was noted in 47 of 51 patients evaluated by Cohen.[61] The severity of the dysphonia is not necessarily indicative of the severity of the web. Cohen reported on 11 patients who complained of aphonia, 16 with weak or whispery voices, 7 with husky voices, and 4 with some hoarseness.

The second most common symptom is airway obstruction, and the severity is directly proportional to the degree of obstruction. The compromise may be so severe that stridor cannot be produced because of limited air movement through the larynx. If stridor is present, it will occur in both inspiratory and expiratory phases, and if the stridor is severe, the airway must be secured with intubation or a tracheostomy. Forty percent of Cohen's patients required a tracheostomy.[61]

Croup rarely occurs in children younger than 6 months. Cohen reported 15 of 51 children with a significant history of recurrent croup,[61] of whom 7 required tracheostomy, 9 presented with pneumonia, and 6 had tracheobronchitis.

Diagnosis

The only way to diagnose the extent of the web correctly is by direct laryngoscopy under general anesthesia.[63] The flexible scope may also have a role in the initial diagnosis, but experience in using it in patients with laryngeal webs is limited. Lateral radiographs of the neck may allow assessment of the width of the web.

Cohen divided laryngeal webs into four types, based on their appearance and an estimation of the degree of airway obstruction.[61] Type 1 is uniform in thickness, with no subglottic extension, with the true vocal cords clearly visible in the web, and compromises less than 35% of the airway (Figure 45–9). Hoarseness is usually only slight. Type 2 is slightly thicker, with a significantly thicker anterior component that may extend into the subglottis. The web restricts the airway by 35 to 50% and usually causes little airway distress, unless the patient has an acute infection or is traumatized during intubation (Figure 45–10). The voice is usually husky. Type 3 is a thick web that is solid anteriorly, with the true vocal cords not well delineated. The web restricts the airway by 50 to 75%, and one notes marked vocal dysfunction, with a weak and whispery voice. Type 4 is a uniformly thick, solid web occluding 75 to 90% of the airway (Figure 45–11). Respiratory obstruction is severe, and the patient is almost always aphonic.

Treatment

Approximately 60% of patients require surgical intervention.[62] Some webs are not repaired because

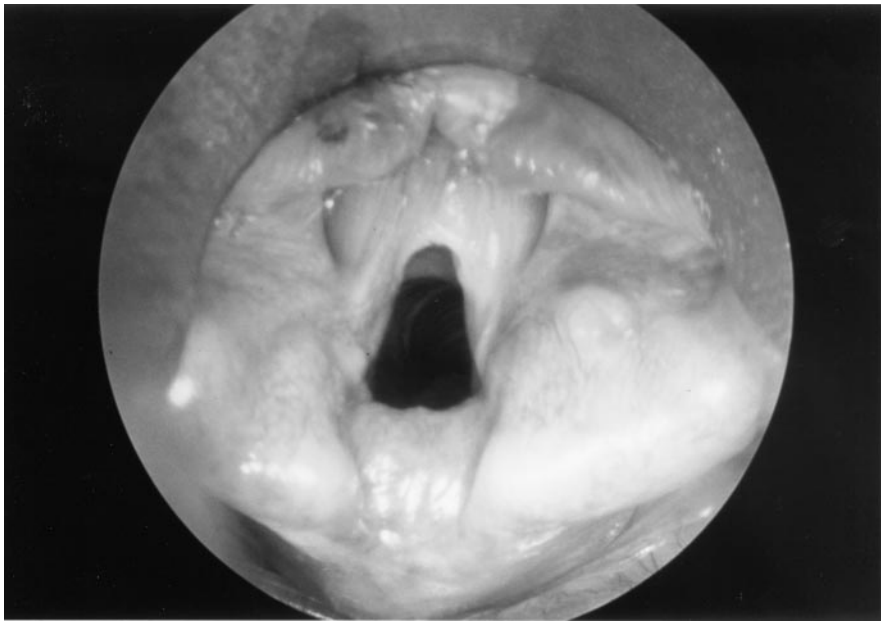


FIGURE 45–9. Type 1 laryngeal web.

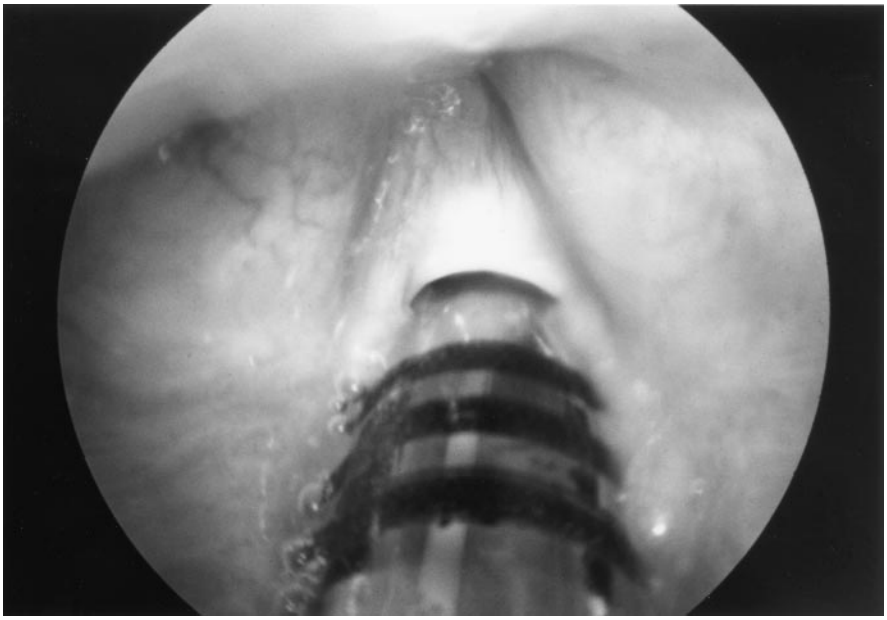


FIGURE 45–10. Type 2 laryngeal web that is intubated. This patient will likely be symptomatic after extubation.

the patients succumb to other systemic complications. Of all patients with laryngeal webs, 30 to 40% require a tracheostomy.[61,62] The type of lesion dictates the surgical approach. In general, the thinner webs are easier to treat; the more severe webs are resistant to surgical management. It is difficult to obtain a crisp anterior commissure, and even if one is obtained, this does not ensure a good voice.

Type I lesions are not life threatening. Many do not require surgery, but if surgery is performed, dilations or excisions with a knife, scissors, or laser are effective. Type 2 lesions require treatment, but not in childhood. Multiple procedures are necessary for excising small portions of the web in multiple steps. Corticosteroids and antibiotics decrease the amount of scarring, and speech therapy may be necessary to maximize phonation. Type 3 lesions frequently require tracheostomy to establish the airway. The child may develop progressive airway distress because of increased airway demands, trauma from intubation, or infection. These lesions may require multiple excisions and frequently necessitate the placement of a keel.[64,65] McGuirt and associates described a method using the laser to develop flaps of the web and reported near-normal results in all patients.[60] Type 4 lesions all require a tracheostomy and excision of the web with placement of a keel. Because they are tracheostomy dependent, these patients should be monitored with apnea monitors. Most surgeons would resect the web and place the keel through a laryngofissure. The airway results are good, but the voice is poor.[61]

SACCULAR CYSTS AND LARYNGOCELES

Laryngoceles and laryngeal saccular cysts are thought to arise from the laryngeal or saccular appendage[66–70] or from retention cysts resulting from obstruction of mucous gland ducts.[71] Morgagni incompletely described the saccule, and Galen also had mentioned it previously.[72] The appendage arises from the anterior part of the ventricle, extends superiorly, and curves slightly posteriorly deep to the false vocal cords and aryepiglottic folds. The orifice opens into the ventricle and measures only 0.5 to 1 mm. The ventriculosaccular fold probably serves to store mucus and direct it posteromedially to lubricate the vocal cords.[73] In the adult, the appendage extends as high as the superior border of the thyroid cartilage. Broyles found that 75% of the saccules measured 6 to 8 mm in length, 25% measured more than 10 mm, and 7% measured more than 15 mm.[74] In the fetus, 25% extend as high as the thyrohyoid membrane.[72] The type of lesion that develops is based on the size of the saccule, whether there is free communication with the laryngeal

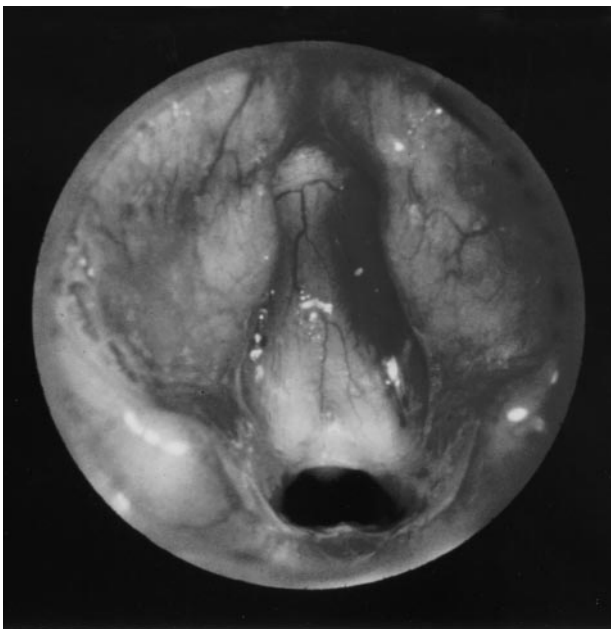


FIGURE 45–11. Type 4 laryngeal web. Courtesy of Dr. L. D. Holinger, Children's Memorial Hospital, Chicago, Illinois.

lumen, and whether there is inflammation within the sac. DeSanto and colleagues believe that the differentiating factor between the development of a cyst or a laryngocele is the patency of the ventricle.[72,75]

Saccular Cysts

Incidence Laryngeal saccular cysts are most likely to become manifest in infancy.[72] Congenital laryngeal cysts are rare, but the awareness of their possibility is important because almost 50% of these cysts are diagnosed at autopsy after the infant has asphyxiated.[73] In 1967, Suehs and Powell found 27 reported cases of congenital laryngeal cysts.[76]

Signs and Symptoms Forty percent of congenital laryngeal cysts are discovered within a few hours of birth, and 95% of the children have symptoms before 6 months of age.[77] The most frequent symptom is stridor (90%), which is primarily inspiratory, although it may be biphasic.[73] The stridor improves in some infants during extension of the head. The cry has been reported as feeble, muffled, shrill, hoarse, or normal.[76,78] Dyspnea, apnea, and cyanosis have been noted in 55% of patients.[78] The associated feeding problems are similar to those in esophageal atresia or tracheoesophageal fistula without atresia and lead to failure to thrive in a large number of patients.[76,77]

Diagnosis Chest and lateral neck radiographs, barium swallow, and CT scan are useful preoperative evaluations of the stridor, but they are not diagnostic.[73,79] Hemangiomas, for example, can cause identical findings, although these lesions appear most frequently in the subglottis. The only way to make the diagnosis definitively is with direct laryngoscopy.[77] Cotton and Richardson suggested that at the time of endoscopy, one should have a large-bore needle, a tracheostomy tray, and an appropriate-size rigid bronchoscope available.[80] The cysts are typically divided into lateral wall saccular cysts and anterior saccular cysts. Lateral saccular cysts are most frequently located in the aryepiglottic fold, epiglottis, or lateral wall of the larynx[76,77] (Figure 45–12). Anterior saccular cysts extend medially and posteriorly between the true and false vocal cords and directly into the laryngeal lumen[77,79] (Figure 45–13). They are most frequently sessile but may be pedunculated. Mitchell and associates reported four subglottic cysts in patients who had previously been intubated and considered the cysts to be probably secondary to intubation.[77]

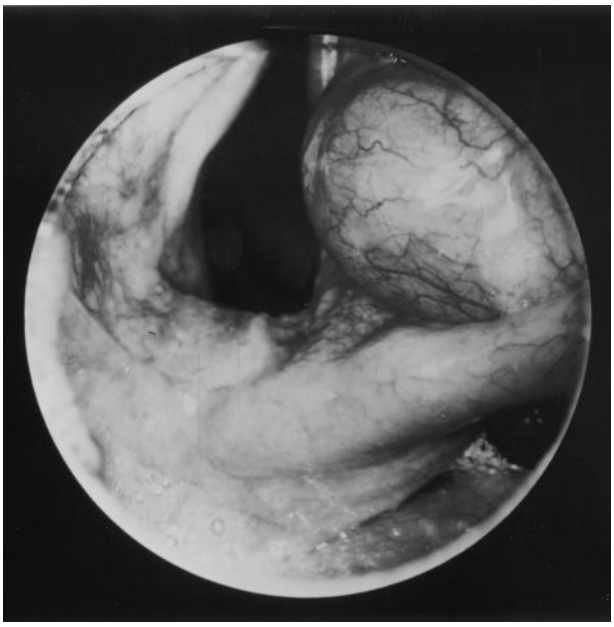


FIGURE 45–12. Lateral saccular cyst. Courtesy of Dr. L. D. Holinger, Children's Memorial Hospital, Chicago, Illinois.

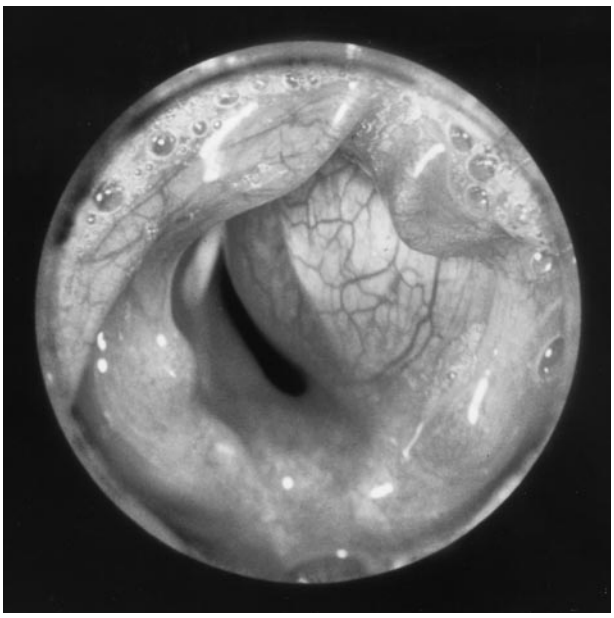


FIGURE 45–13. Anterior saccular cyst. Courtesy of Dr. L. D. Holinger, Children's Memorial Hospital, Chicago, Illinois.

Treatment Treatment of laryngeal cysts may require emergency tracheostomy.[73,81,82] Mitchell and associates noted that 20% of their patients required emergency intervention.[77] In the infant, the cysts should be treated primarily with endoscopic deroofing or aspiration. Suehs and Powell recommended endoscopic incision and drainage as needed.[76] Holinger and colleagues thought that smaller anterior saccular cysts could be effectively treated with cup forceps removal (excision biopsy) at the time of direct laryngoscopy.[73] These investigators would not attempt endoscopic dissection of the entire sac. Aspiration has also been recommended, but the incidence of recurrence is high.[67,83] Mitchell and associates reported successful treatment of 7 of 17 patients with deroofing or marsupialization.[77] A second deroofing was required in six of the seven recurrences. The remaining failure was treated with total excision through a laryngofissure.

Laryngoceles

Incidence As of 1977, 300 laryngoceles had been recorded in the world literature. They may occur in infants and children,[84] and they cause airway obstruction.[85] The presence of laryngoceles in this age group appears to be rare, however. Laryngoceles are much more common in adults and appear most commonly in the fifth decade.[84] Holinger and Brown reported on 12 patients with congenital laryngoceles in infancy.[78] Baker and associates found an increased incidence of laryngoceles in laryngeal cancer and suggested that obstruction of the saccule by the carcinoma was a factor in developing laryngoceles.[86]

Signs and Symptoms Laryngoceles are symptomatic only when they are filled with air or fluid, so the symptoms may be intermittent. Because the laryngocele may rapidly inflate and deflate, several radiographs may be necessary to document the lesion. A fluid-filled laryngocele may be difficult to differentiate from a cyst. If it becomes infected, it fills with mucopus and is called a laryngopyocele.

Diagnosis The definitive diagnosis is made with direct laryngoscopy, but because the symptoms are intermittent, more than one examination may be necessary. Laryngoceles originate from the ventricle and bulge out between the true and false vocal cords or dissect posteriorly into the arytenoid and aryepiglottic fold. In general, three types have been described: internal, external, and mixed[84]; however, this classification is artificial. The internal type is limited to the larynx, whereas the external type extends into the neck through the isthmus of the thyrohyoid membrane. The mixed type involves both internal and external portions and is the most common. Laryngoceles should be diagnosed with direct laryngoscopy when they are symptomatic, in addition to radiographic findings. Trapnell has emphasized that the radiographic appearance of a large saccule does not warrant the diagnosis of a laryngocele.[87]

Treatment Only symptomatic lesions require treatment. DeSanto preferred the external approach[72]; however, most of his experience was in adults. The procedure described by Yarington and Frazer provides adequate exposure for resection.[88] Holinger and associates emphasized that true laryngoceles are rare in children and should be treated endoscopically by deroofing with cup forceps.[73] On occasion, the laryngocele can be aspirated, and a more orderly resection can be performed at a later date.[89] The laser is also an alternative.

LARYNGEAL CLEFTS

Congenital laryngotracheoesophageal cleft is characterized by a deficiency in the separation between the esophagus and the larynx or trachea. Type I clefts are recognized more frequently[90,91]; however, the true incidence is not yet known. Larger clefts are associated with a high mortality.[92]

Incidence

Most clefts occur through the posterior part of the glottis; however, rare ventral or anterior clefts have been reported.[93–95] Clefts may be classified into three broad categories. Type I is called the laryngeal cleft and is found only in the posterior part of the glottis. Types II and III involve the trachea in addition to the larynx (Figure 45–14). Type II extends down to but not beyond the sixth tracheal ring, and type III can extend to the carina. As of 1983, 85 well-documented cases of clefts have been reported.[96] Finlay observed a family in which two of five children developed a cleft larynx.[97] Other investigators have reported similar examples.[85–87] The inheritance pattern is almost always autosomal dominant.[87] Posterior glottic clefts have been associated with tracheal agenesis, as well as many other congenital anomalies.[96]

Signs and Symptoms

In the world literature, of 85 patients, 35 had type I clefts, 36 had type Il clefts, and 14 had type III clefts.

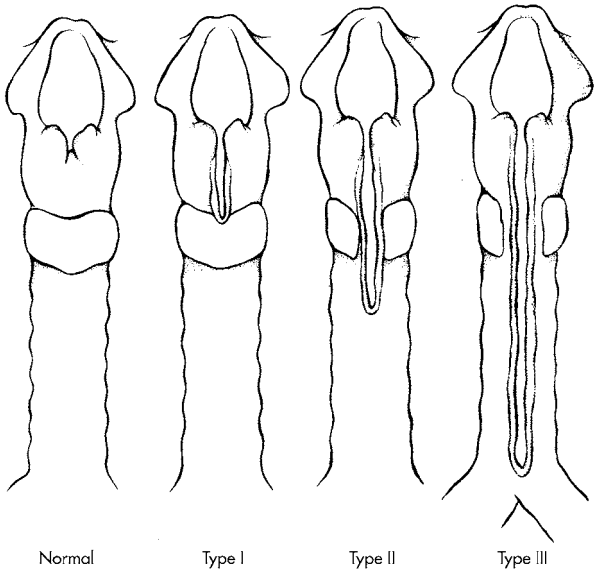


FIGURE 45–14. Classification of laryngeal and tracheoesophageal clefts as proposed by Evans.[98]

The following manifestations, in order of frequency, are associated with laryngotracheoesophageal clefts: aspiration and cyanosis (53%), postpartum asphyxia (33%), increased mucus production (23%), recurrent pneumonia (16%), voiceless crying (16%), stridor (10%), and impaired swallowing (5%).[96] The simultaneous occurrence of increased saliva production; low, soundless, or hurried crying; and stridor should lead to the suspicion of a cleft. The inspiratory stridor is produced by the collapse of redundant mucosa around the cleft and from the intrusion of the arytenoid into the airway.[98,99] Cohen reported that the stridor may be expiratory because of aspirated secretions.[100] One sees a frequent combination of posterior laryngeal clefts and tracheoesophageal fistulae.

Diagnosis

Occult posterior laryngeal clefts have been diagnosed with MRI.[101] A laryngeal cleft is frequently demonstrated by an esophagram with water-soluble contrast medium[102] or is suggested by aspiration pneumonia on a chest radiograph. The most important diagnostic test, however, is direct laryngoscopy. If one does not specifically look for the entity, it frequently escapes detection.[103] When the larynx is examined, redundant mucosa in the posterior cleft is usually the first clue to the defect (Figure 45–15). On inspiration, this redundant mucosa may rotate into the airway. The cleft can best be demonstrated by placing a laryngoscope, such as a Dedo or Jako, into the supraglottis and examining the posterior part of the glottis. If a posterior laryngeal cleft is present, this maneuver will clearly demonstrate the lesion and its extent (Figure 45–16). The posterior part of the glottis can also be palpated with a spatula or a similar thin instrument to demonstrate the defect. Some authors find microlaryngoscopy helpful.[98] Some clefts are submucous, and diagnosis is made only on laryngeal sections.[104,105] The clinical significance of submucous clefts is unknown. There is a high incidence of associated esophageal lesions and tracheoesophageal fistulae, for which careful examination must be made.[98,100,104,106,107]

Treatment

The prognosis for a patient with this lesion is not favorable. In the review presented by Roth and colleagues,[96] 24 patients died before surgical interven-

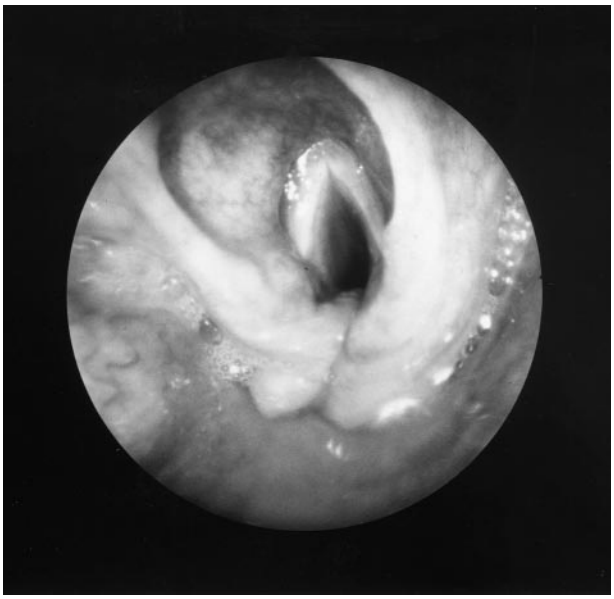


FIGURE 45–15. Laryngeal cleft demonstrating redundant mucosa prolapsing into the larynx on inspiration.

tion; of patients with type III clefts, 13 of 14 died; of those with type II clefts, 13 of 31 died; and of those with type I clefts, 13 of 30 died. Almost two-thirds of the deceased patients had other severe malformations. Clefts were repaired in 48 patients, 13 of whom died. A tracheostomy was performed in 50% of all patients. The first successful operative correction of the defect in the larynx was achieved by Petterson.[108]

Three surgical approaches to the posterior aspect of the larynx have been described: (1) minor type I clefts can be repaired endoscopically[109,110]; (2) lateral pharyngotomy has been used frequently, especially for the smaller laryngeal clefts,[96,98] the major disadvantage of which lies in the danger of injuring the recurrent nerve; the exposure is limited, and it is difficult to obtain a good layered closure; and (3) anterior laryngofissure does not risk the recurrent laryngeal nerve and has been used in neonates.[96,99,100,111–113] Evans reported better exposure and a better two-layer closure in type II and III clefts with a laryngofissure[98] (Figure 45–17). Prescott successfully repaired a cleft with a posterior rib graft through a laryngofissure.[114] A disturbance of the laryngeal growth appears unlikely. Intraoperative airway management has been a problem in repairing these lesions. Ahmad and associates solved the problem by using a bifurcated endotracheal tube.[115] Geiduschek and colleagues reported the use of extracorporeal membrane oxygenation for extensive cleft repair.[116] We have used bypass to repair long-segment tracheal stenosis.

Cotton and Schreiber have noted that, after repair of posterior clefts, the patient may have continued esophageal reflux and aspiration of gastric contents.[117] There appears to be a significant decrease in successful repair if a tracheoesophageal fistula is present.[107] When this occurs, these investigators recommend a high gastric section with a dou-

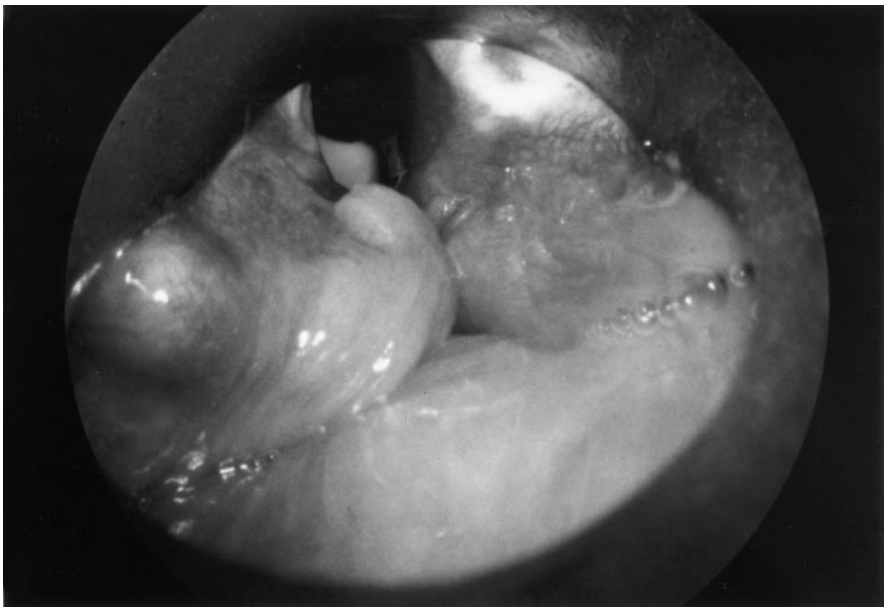


FIGURE 45–16. Jako laryngoscope inserted into the posterior commissure, demonstrating the laryngeal cleft.

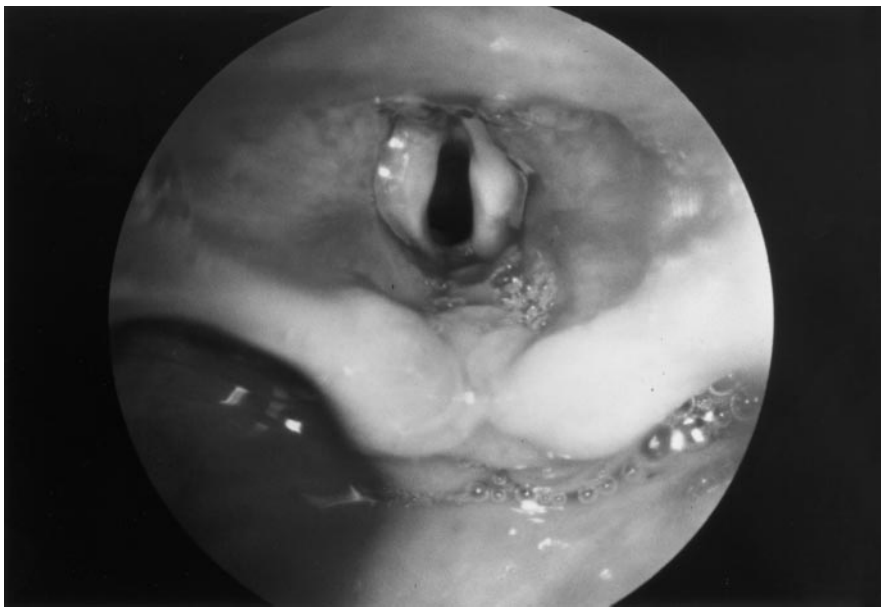


FIGURE 45–17. Type I cleft repaired through a laryngofissure.

ble gastrostomy to minimize the chance of aspiration of gastric contents. Robie and associates reported that of 170 clefts repaired, 19 required revision surgery.[113]

VOCAL CORD PARALYSIS

Vocal cord paralysis can be categorized into congenital and acquired lesions. One or both of the true vocal cords may be involved, and bilateral vocal cord paralysis is more frequent.[118–120] Apnea is frequently seen in bilateral vocal cord paralysis.[121–123] Congenital lesions are frequently associated with central nervous system lesions, including hydrocephalus, meningomyelocele, Arnold-Chiari malformation, meningocele, encephalocele, cerebral agenesis, nucleus ambiguus dysgenesis, neuromuscular disorders, and myasthenia gravis.[124,125] Arnold-Chiari is the most frequent malformation and is likely a contributing factor in most cases.[126]

Incidence and Etiology

Estimates of the frequency of vocal cord paralysis range from 1.5 to 23%[120,127,128]; according to some authors, it ranks second in frequency among all congenital laryngeal lesions.[124] Holinger and associates found that congenital lesions were more frequent than acquired.[125] The acquired group can be further categorized into traumatic, infectious, or neoplastic. Traumatic lesions are most frequent secondary to stretching of the recurrent laryngeal nerve during vaginal delivery or surgical trauma in management of bronchogenic cysts, tracheoesophageal fistulae, or patent ductus arteriosus.[124] Infectious diseases such as whooping cough, encephalitis, poliomyelitis, diphtheria, rabies, tetanus, syphilis, and botulism are now rarely seen but can cause vocal cord paralysis.[123] Tumors of the brain and spinal column are also rare but can cause unilateral or bilateral vocal cord paralysis.

The pathophysiology of bilateral vocal cord paralysis is unclear, but the condition may result from (1) compression of the vagus nerves in their course through the foramen magnum, (2) traction of the cervical rootlets of the vagus nerves by the caudal displacement of the brainstem, or (3) brainstem dysgenesis.[129,130] Most authors favor the compression theory because, with timely decompression of hydrocephalus or the Arnold-Chiari malformation, the vocal cords regain function. Familial bilateral vocal cord paralysis[131–133] and persistent apnea after tracheostomy[121] appear to be most appropriately explained by the dysgenesis theory. Probably, more than one lesion can cause vocal cord paralysis.

Laryngeal lesions such as subglottis stenosis, laryngomalacia, and posterior laryngeal clefts have also been associated with, but do not cause, vocal cord paralysis.[134]

Signs and Symptoms

Any of the three laryngeal functions of respiration, voice production, and deglutition can be affected by vocal cord paralysis. With unilateral vocal cord paralysis, the voice is breathy and weak, but the patient has an adequate airway unless stressed. Stridor, weak cry, and some degree of respiratory distress can be seen in all patients with bilateral vocal cord paralysis.[130] Dedo noted that if *both* the recurrent and superior laryngeal nerves are paralyzed, the vocal cords will be in the intermediate position, and the airway then will frequently be sufficient to allow adequate ventilation.[124,135] If the recurrent nerves only are paralyzed, the vocal cords will be in the paramedian position, resulting in an inadequate airway.

Stridor is the most frequent presenting symptom of bilateral vocal cord paralysis,[125] and its onset may be sudden.[130] Older children suppress laughing and coughing because of the increased respiratory demand. The airway becomes narrower, and there will be an increase in stridor and the development of nasal flaring, restlessness, and the use of accessory respiratory muscles, with an indrawing of the sternum and epigastrium. The stridor may progress to cyanosis, apnea, and respiratory and cardiac arrest if not recognized and treated.

Aspiration and dysphagia are frequently noted in patients with bilateral vocal cord paralysis.[129,136] Pneumonia may first be apparent when signs of increasing cranial pressure appear.

Diagnosis

The diagnosis is made by flexible laryngoscopy or direct laryngoscopy. Vocal cord mobility may be difficult to examine in an infant, and the intermediate versus paramedian position may be impossible to assess.

Treatment

Once the diagnosis of vocal cord paralysis is made, the airway should be secured if the patient has significant airway distress. The airway is best established with intubation, followed by a full workup to ascertain the cause of the vocal cord paralysis. One must look specifically for associated findings of meningomyelocele, Arnold-Chiari malformation, and hydrocephalus.[125] If the compression of the nerve is relieved within 24 hours, the vocal cords will regain function within 2 weeks2[125,130]; otherwise, vocal cord function may not return for 1½ years, if at all. If intervention has been timely, the larynx should be examined periodically to assess vocal cord function. If the patient shows no evidence of function within 1 to 2 weeks, a tracheostomy should be performed to relieve the airway distress,[123] but central sleep apnea may continue even with appropriate early decompression.[121] Once the tracheostomy is in place, periodic examinations will be necessary to assess vocal cord function.

Approximately 50% of children with bilateral vocal cord paralysis require tracheostomy.[118,125,130] Bluestone and colleagues reported a 50% mortality rate secondary to shunt failure or infection in patients with Arnold-Chiari malformation.[130] Of those who survive, 25 to 48% can be decannulated.[119,129]

SUBGLOTTIC LESIONS

The principal subglottic lesions are stenoses and hemangiomas.

Stenosis

Stenoses of the subglottic area can be divided into cartilaginous stenoses, which are usually congenital, and acquired membranous or soft tissue stenoses.

Incidence and Etiology A stenosis is considered to be congenital when the patient has no history of endotracheal intubation or trauma. The normal subglottic lumen is 4.5 to 5.5 mm in a full-term neonate and 3.5 mm in a premature neonate. We do not know in how many infants' extubations fail because of congenitally small cricoid cartilages.

Congenital subglottic stenosis is the third most common congenital abnormality. Cricoid cartilage deformities are usually congenital and consist of abnormal shapes and sizes. The cricoid cartilage may have a normal shape but may be small or hypoplastic. It may also be elliptical, flattened, or otherwise distorted. The first tracheal ring can also be trapped under the cricoid cartilage, resulting in a narrowed subglottis. The primary causes of acquired or membranous subglottic stenosis (Figure 45–18) in children are (1) external injury from blunt trauma or a high tracheostomy and (2) internal injury from prolonged intubation and chemical or thermal burns. External injuries are rare in infants. Internal trauma, secondary to prolonged intubation, is thought to

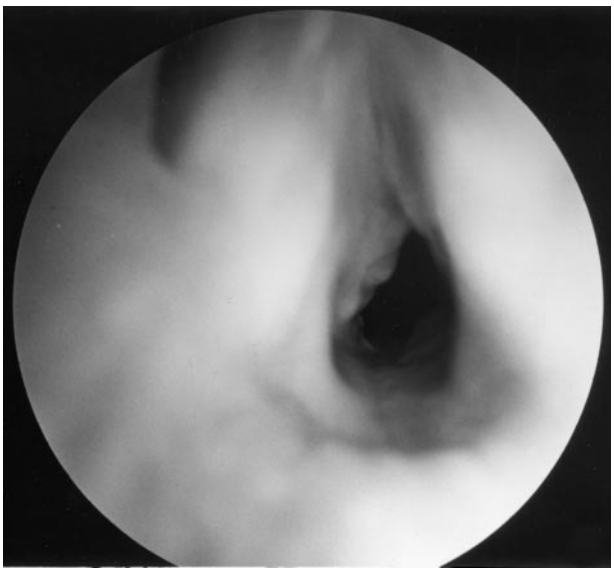


FIGURE 45–18. Membranous subglottic stenosis.

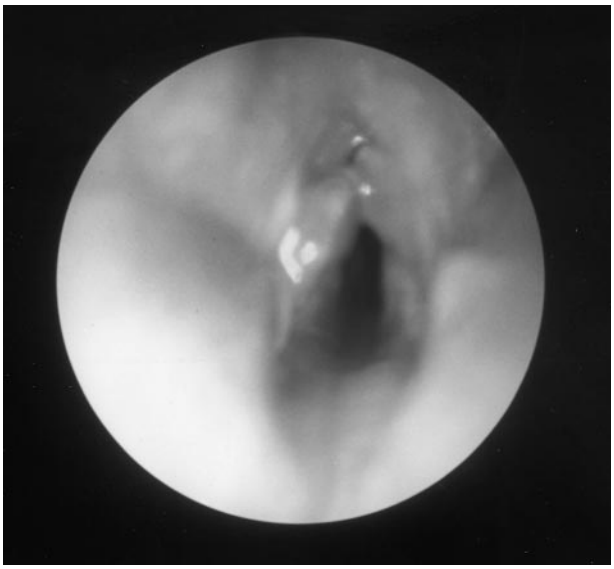


FIGURE 45–19. Ulcerated right vocal process.

account for approximately 90% of acquired subglottic stenosis.[137,138] The incidence of stenoses after intubation ranges from 0.9 to 8.3%.[139–142] The stenoses most frequently occur in the subglottis because (1) the cricoid cartilage is a complete circular ring, without any "give" for edema; (2) the edema can accumulate rapidly in the loose areolar tissue of the subglottis; (3) the pseudostratified, ciliated columnar respiratory epithelium is delicate and easily traumatized; and (4) the narrowest portion of the upper airway is in the subglottis and is therefore the most likely to be traumatized.[138]

The pathophysiology of acquired subglottic stenoses is well described. The endotracheal tube causes pressure necrosis of the respiratory epithelium. Edema and superficial ulceration begin, and the normal ciliary flow is interrupted. As the ulcer deepens, secondary infection of the areolar tissue and perichondrium begin. Chondritis may eventually occur, with necrosis and collapse of the cricoid cartilage. Benjamin characterized the traumatic lesions formed in the larynx with prolonged intubation.[143] He noted that, in the acute phase, one sees formation of posterolateral ulcerations at the vocal processes ("ulcerated troughs") (Figure 45–19), and usually one sees an "intact median strip" of mucosa and an "annular ulceration" of mucosa in the subglottis. "Tongues of granulation tissue" (Figure 45–20) form anterior and posterior in the ulceration, and frequently one sees generalized inflammation and edema of the ventricle that results in ventricular protrusions. The long-term complications of these lesions are characterized as posterior glottic synechiae, "healed furrows," posterior subglottic and glottic stenoses (Figure 45–21), "healed fibrous nodules," submucosal mucous gland hyperplasia with ductal cysts, and submucosal fibrosis and stenosis. The scar tissue forms in the subglottis and limits the airway.[144–148] Healing is inhibited in part by poor blood supply in the subglottis and constant motion of the larynx.

Neonates tolerate prolonged intubation better than adults. The reasons for this are unclear, but more pliable cartilage[144] and the higher position of the larynx in the neck[146] have been suggested as considerations.

Certain factors can increase the chances of developing subglottic stenosis. An oversized endotracheal tube or a tube of appropriate size in a patient with a small cricoid cartilage can increase the mucosal pressure and result in a deep ulceration. Primary intubation can traumatize the subglottis. In children, an endotracheal tube that allows a leak at pressure less than 20 cm H_2O should be chosen. Reintubation,[144,149] shearing motion of the tube on movement of the head,[150] and superimposed local or systemic bacterial infections[145] increase the risk. Gastroesophageal reflux can increase the inflammation

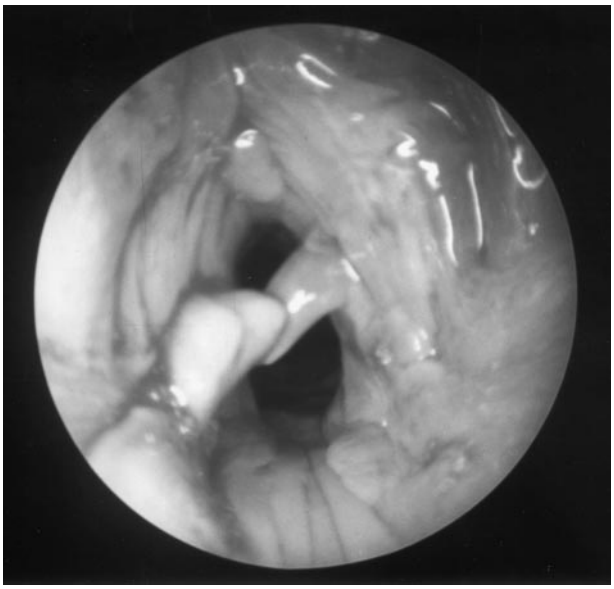


FIGURE 45–20. Tongues of granulation tissue.

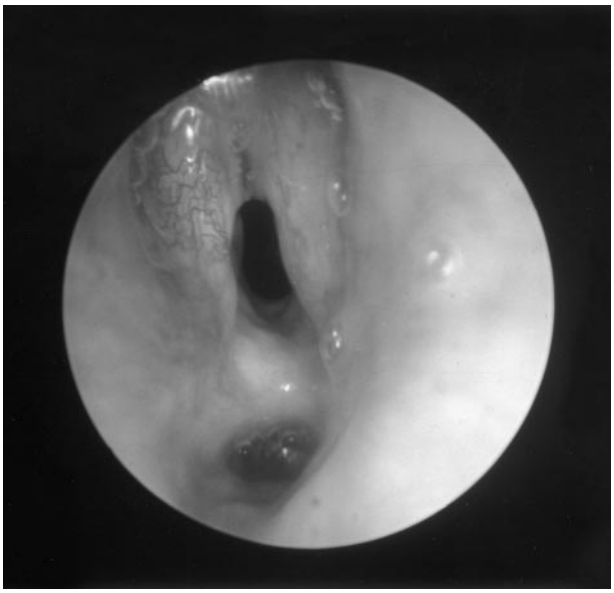


FIGURE 45–21. Posterior glottic web.

and tissue trauma. Nasogastric tubes and endotracheal tubes have been noted to cause pressure necrosis of the cricoid[145] and increase the risk of reflux.

Systemic factors such as immunodeficiency, anemia, neutropenia, toxicity, hypoxia, dehydration, and poor perfusion increase the risk of developing mucosal ulceration and subsequent scar formation.

Signs and Symptoms In the intubated neonate, evidence of subglottic stenosis may not manifest until the patient is ready for extubation. If a subglottic ulcer is present, the airway may be compromised immediately or edema may accumulate over a few hours. In some patients, symptoms do not develop until 2 to 4 weeks after intubation. If the patient has mild to moderate congenital subglottic stenosis, symptoms may not appear until an infection of the upper respiratory tract causes additional narrowing and respiratory distress.

If the stenosis is congenital, the only manifestation may be prolonged or recurrent croup. A common dictum is that "there is no such thing as croup under 1 year of age." As respiratory demands increase, the infant may become symptomatic.[80]

The main symptoms and signs relate to airway, voice, and feeding. Stridor is the primary sign and is biphasic, with the inspiratory phase always louder. With progressive narrowing of the airway, respiratory distress ensues. If the vocal cords are affected, hoarseness, abnormal cry, and aphonia will indicate that an anterior web is present. Dysphagia and aspiration pneumonia can occur.

Diagnosis Soft tissue radiographs of the lateral neck may demonstrate subglottic narrowing. Xeroradiography demonstrates the tissue–air interface better and is believed by many to be the best method for evaluating chronic airway problems.[151–153] Computed tomographic scans do not give adequate additional information.

Direct laryngoscopy is the most important diagnostic step in assessing the thickness and length of the stenosis and involvement of the larynx. Flexible fiber-optic laryngoscopy is most useful in assessing vocal cord function.[151] Because the flexible scope provides only a limited view of the posterior part of the glottis and subglottis, rigid endoscopy is necessary to assess the size and patency of the lumen. The airway may be sized with the bronchoscope or endotracheal tubes. Storz Hopkins rod lens telescopes are especially important in visualizing the extent of the stenosis in the subglottis. Because of the wide-angle view, estimating the actual dimensions of the lumen is difficult. Often the narrowing is so great that only the telescope can be used.

Treatment Congenital subglottic stenosis that is mild and causing mild symptoms and signs can be

treated expectantly. Some patients outgrow their lesions and are unlikely to need surgical correction.[154] Treatment must be individualized.

Tracheostomy is required in fewer than half of patients with congenital subglottic stenosis.[144] Normal growth and development may allow decannulation within 2 to 5 years.

The *anterior cricoid split operation* was initially devised to treat acquired subglottic stenosis,[155] and it was later applied to patients with the congenital form.[156,157] This procedure breaks the cartilaginous cricoid ring anteriorly to allow expansion of the subglottis. It is used in patients who have confirmed stenosis and failed extubation but do not require airway support and have mature lungs with oxygen requirements less than 35%. The procedure involves making a vertical incision through the lower third of the thyroid cartilage, the cricoid cartilage, and the first two tracheal rings. Many of the cricoid rings spring open, and the endotracheal tubes are readily seen through the incision. Others do not open to any significant degree. Some surgeons have recommended placement of auricular or rib grafts at the time of the decompression.[158–160] The patient is left intubated with an endotracheal tube one size bigger for 5 to 7 days and is re-examined at the time of extubation. Corticosteroids are usually administered 3 to 5 days before planned extubation.

In weighing the advantages of endoscopic and open procedures, the surgeon must take into account the extent of the scarring and the expertise required. The goal of any surgical procedure is to extubate or decannulate the patient by repairing the stenosis with minimal effect on the voice.

Dilatation is useful if the ulceration is still present and granulation tissue is forming. It is most appropriately used with immature scar or submucosal fibrosis. Gentle dilation is performed with a round, smooth instrument, usually in conjunction with corticosteroid injection.[161] Aggressive dilatation with corticosteroid injection can induce additional trauma and cause significant necrosis of the cartilage.

The use of *corticosteroids* is controversial. They tend to decrease scar formation through their anti-inflammatory action and by delay of collagen synthesis. They may be used systemically or injected locally. Corticosteroids have not been successfully used to treat mature scar in the subglottis.

A variety of methods for *endoscopic correction* of subglottic stenosis have been suggested, including microcauterization,[162] cryosurgery,[163,164] serial electrosurgical resection,[165] and carbon dioxide laser.[166] The carbon dioxide laser appears to be the current modality of choice.[146,166–171] The laser can only be used to resect membranous stenosis, and the procedure has to be performed in stages. The more aggressively the laser is used, the less the likelihood of a successful outcome and the greater the risk of inducing additional scarring. Cotton and Manoukian identified several factors associated with poor results: (1) circumferential scarring, (2) scar tissue greater than 1 cm in length, (3) scar in the posterior commissure, (4) severe bacterial infection of the trachea after a tracheostomy, (5) exposure of the perichondrium or cartilage with the laser, (6) combined laryngotracheal stenosis, (7) failure of previous endoscopic procedures, and (8) previous loss of cartilaginous framework.[172] Prophylactic systemic antibiotic therapy is recommended for endoscopic procedures. Adequate exposure of the subglottis is necessary. The subglottiscope designed by Healy is useful if the patient does not have a tracheostomy in place. Supraglottic jet ventilation provides the best exposure. Thin webs are most appropriately treated with the laser.

The *open procedures* have been traditionally used for the more severe lesions. Surgeons are increasingly using single-stage reconstruction.[159,173] This allows for more rapid correction of the problem and is perhaps more cost effective. The *hyoid interposition* was first reported by Looper.[174] Bennett also reported the use of a hyoid bone graft for the treatment of subglottic stenosis.[175] Bone resorption was a problem, which led to the development of pedicle hyoid-sternohyoid myo-osseous flaps.[176–178]

The anterior cricoid split with endotracheal tube stenting has been recommended for treatment of milder forms of stenosis. The standard reconstruction has been with autogenous costal cartilage.[179–181] Because of its abundance, costal cartilage is a better material to use than thyroid cartilage, hyoid bone, or muscle pedicle.

The *costal cartilage reconstruction* has been the standard method for subglottic reconstruction for the past several years. The fifth or sixth cartilaginous rib or costal margin cartilage is harvested. One incision can be used to harvest two ribs, but this is seldom necessary. The perichondrium is left on the lateral surface of the rib, with incisions through it along the superior and inferior borders and stripped

from the medial surface. Some surgeons use the rib stripper on small double-prong hooks and a freer elevator. The incision in the neck is usually a U-shaped flap through the tracheostoma or a horizontal incision over the cricoid if a tracheostomy is not in place. The larynx and trachea are exposed, and a midline incision is made through the length of the stenosis. If the stenosis is severe, the lumen can first be identified with a 25-gauge needle under endoscopic visualization. When only an anterior graft is used, the incision usually extends from the tracheal rings through the cricoid and the lower third of the larynx. If a posterior graft and stent are required, a full laryngofissure will be necessary. The intraluminal scar and mucosa are incised strictly in the midline. This is best performed by first marking the lumen transtracheally with a 25-gauge needle. Once the airway is open, the scar is not removed. The length of the trachea and larynx to be reconstructed is measured, and the rib graft is designed to fit the defect. The graft is elliptical, with the perichondrium toward the lumen. The rib is not thinned, and care is taken to shape the cartilage to be as wide as possible, with a lip or shelf to prevent the graft from falling into the airway (Figure 45–22).

In patients with greater subglottic stenosis, a posterior graft is needed, and the posterior cricoid lamina is divided in the midline to, but not through, the hypopharyngeal mucosa. In complete stenosis, a four-quadrant split has been recommended. The cricoid is separated to a distance of approximately 1 mm per year of age up to 10 mm between the two halves.[172] An appropriate wedge of rib with the perichondrium facing the lumen is sutured into the posterior cricoid cleft. If a posterior graft has been performed, a stent will be required to maintain the lumen postoperatively. The current stent of choice is the Aboulker-styled Teflon stent, which is round. Unfortunately, it can cause significant blunting, trauma, and granulation tissue formation at the anterior commissure.

Once the posterior cricoid graft is positioned, the stent is placed and the anterior tracheal wall is closed. Usually, posterior and anterior grafts are used concomitantly. When a stent is used, it must first be positioned in the airway with the tracheostomy in place. The best method of securing the tracheostomy to the stent is with wire, as described by Cotton and Manoukian.[172] Placement must be checked endoscopically. The stent should come to the level of the false vocal cords. If necessary, the stent length should be revised to hold the proper position. The rib is sutured into place with 4-0 Vicryl or polydioxanone suture, taking care to go through the cartilage and not allowing the suture to be exposed intraluminally. If suture is located in the lumen, granulation tissue and chronic infection will result. The neck wound is closed in layers. The stent is left in place for a variable length of time based on the surgeon's judgment and the severity of the stenosis.

Recently, auricular cartilage has been used as a graft material.[158,159] It is useful because it is malleable;

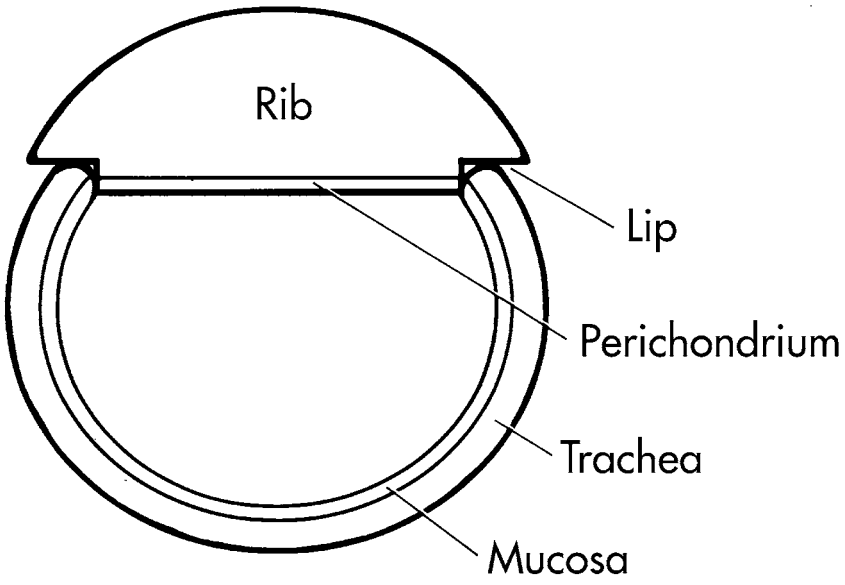


FIGURE 45–22. Drawing depicting the rib graft, which has not been thinned, and a lip to prevent collapse into the lumen.

its curved shape allows for greater anteroposterior dimensions than available with the rib graft (Figure 45–23). The auricular cartilage graft has not been used extensively with staged reconstructions.

The results of successful decannulation depend in part on the severity of the stenosis. Viral infections with respiratory syncytial virus and bacterial infections with *Pseudomonas* have been associated with an increased failure rate.[182] The ultimate outcome of grade 1 is 92% successful, that of grade 2 is 85% successful, that of grade 3 is 70% successful, and that of grade 4 is 36% successful.[183]

Subglottic Hemangioma

Incidence and Etiology Congenital subglottic hemangiomas are relatively rare lesions and develop primarily in the submucosa. Fifty percent are associated with cutaneous hemangiomas. No correlation exists between the size of the cutaneous lesion and the size of the laryngeal lesion.[184] A 2 to 1 female preponderance is recognized.

Signs and Symptoms The subglottic hemangioma develops in a typical growth pattern with increasing size, usually causing symptoms in the first 8 weeks of life. Almost all of these lesions will become manifest before 6 months of age. Some lesions extend into the perichondrium and tracheal rings and beyond the trachea. Variable and fluctuant respiratory distress usually progresses to persistent distress. The stridor is more prominent on inspiration but is present during expiration. The voice is altered to a varying degree, depending on the involvement of the larynx. Altered cry, hoarseness, barking cough, and failure to thrive are the other frequently noted manifestations. Recurrent croup is the most frequent erroneous diagnosis.

Diagnosis Lateral radiographic studies suggest a subglottic abnormality consistent with hemangioma. This is not diagnostic, however, and the diagnosis can only be made endoscopically (Figure 45–24). The appearance of these lesions is characteristic and can be made by the experienced endoscopist without a biopsy. The lesion is sessile, fairly firm, but compressible; pink, red, or bluish; and poorly defined. The vessels are rarely, if ever, cavernous. If a biopsy is performed, the bleeding is usually not profuse, but the airway must be maintained, most frequently with a tracheostomy. The lesion is usually unilateral or asymmetric. Multiple hemangiomas of the airway have been reported.[185] A thorough workup of the possible extent of the hemangioma in the mediastinum is recommended.

Treatment If the airway distress is significant at the time of diagnosis, immediate airway control will be necessary. Endoscopy may compromise the airway. If necessary, the airway may be secured with intubation or a tracheostomy. If the patient is intubated, tracheostomy is more likely to be required.

A wide range of treatments have been advocated. They include the following: tracheostomy with observation and awaiting spontaneous resolution; prolonged endotracheal intubation, with and without corticosteroid injection; injection of sclerosing agents; the use of systemic corticosteroids;

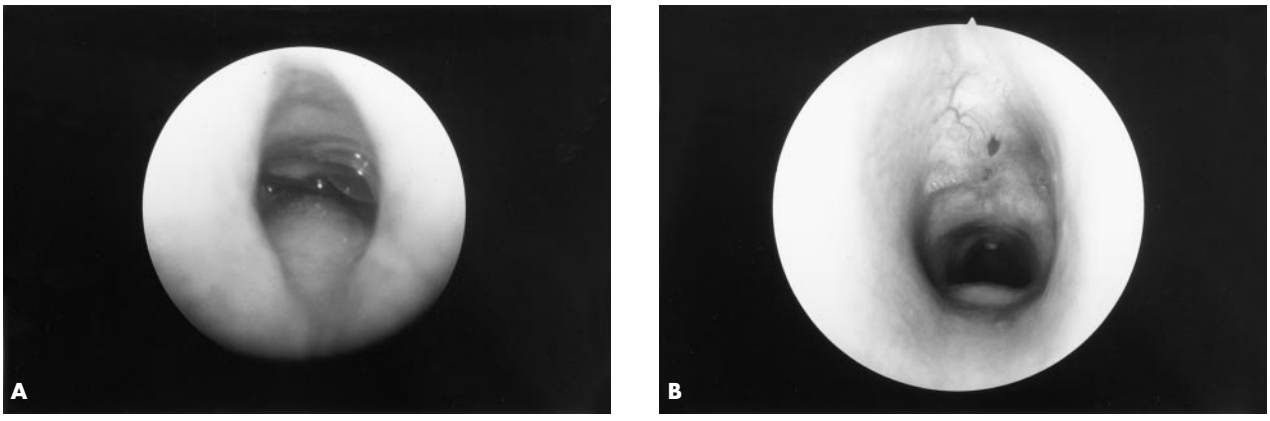


FIGURE 45–23. Preoperative (*A*) and postoperative (*B*) photographs after a laryngotracheoplasty with auricular cartilage.

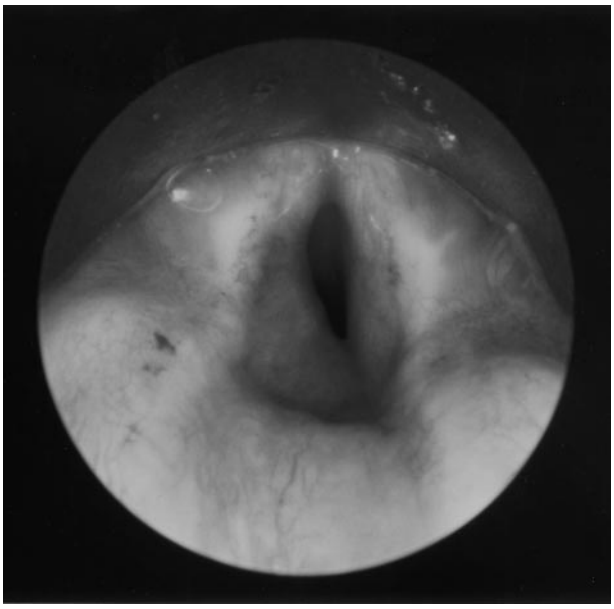


FIGURE 45–24. Subglottic hemangioma.

cryosurgery; external beam irradiation; gold seed irradiation; and laser excision. Prospective studies are not available. The judicious use of a laser to excise the lesion in stages, with the patient's airway protected by a tracheostomy when necessary, appears to be the treatment favored by most surgeons. External beam irradiation has been reported to have a cure rate of 93%, but the threat of radiation-induced malignant tumors of the thyroid gland strictly limits its use. Gold seeds are the treatment favored by Benjamin.[184] Unrecognized or mistreated subglottic hemangiomas have a high mortality rate.

REFERENCES

1. Chen JC, Holinger LD. Congenital laryngeal lesions: pathology study using serial macrosections and review of the literature. Pediatr Pathol 1994;14: 301–25.
2. Thomson JM. On infantile respiratory spasm (congenital laryngeal stridor). Edinb Med J 1892;38: 205–16.
3. Sutherland GA, Lack HM. Congenital laryngeal obstruction. Lancet 1897;2:653–5.
4. Thomson JM, Turner AL. On the causation of congenital stridor of infants. BMJ 1900;2:1561–3.
5. Schwartz L. Congenital laryngeal stridor (inspiratory laryngeal collapse). New theory as to its underlying cause and the desirability of a change in terminology. Arch Otolaryngol 1944;39:403–12.
6. Jackson C, Jackson CL. Diseases and injuries of the larynx. New York: Macmillian; 1942.
7. Ford GR, Irving RM, Jones NS, Bailey CM. ENT manifestations of Fraser syndrome. J Laryngol Otol 1992;106:1–4.
8. Hawkins DB, Clark RW. Flexible laryngoscopy in neonates, infants, and young children. Ann Otol Rhinol Laryngol 1987;96:81–5.
9. Nussbaum E, Maggi JC. Laryngomalacia in children. Chest 1990;98:942–4.
10. McSwiney PF, Cavanagh NP, Languth P. Outcome in congenital stridor (laryngomalacia). Arch Dis Child 1977;52:215–8.
11. Apley J. The infant with stridor: a follow-up survey of 80 cases. Arch Dis Child 1953;28:423–35.
12. Holinger PH, Johnston KC. The infant with respiratory stridor. Pediatr Clin North Am 1955;2:403–11.
13. Belmont JR, Grundfast K. Congenital laryngeal stridor (laryngomalacia): etiologic factors and associated disorders. Ann Otol Rhinol Laryngol 1984; 93:430–7.
14. Kelemen G. Congenital laryngeal stridor. Arch Otolaryngol 1953;58:245–68.
15. Holinger LD, Konior RJ. Surgical management of severe laryngomalacia. Laryngoscope 1989;99: 136–42.
16. Solomons NB, Prescott CA. Laryngomalacia. A review and the surgical management for severe cases. Int J Pediatr Otorhinolaryngol 1987;13:31–9.
17. Cotton RT, Richardson MA. Pediatric otolaryngology. Philadelphia: WB Saunders; 1983.
18. Ferguson CF. Congenital abnormalities of the infant larynx. Otolaryngol Clin North Am 1970;3:185–200.
19. Zalzal GH, Anon JB, Cotton RT. Epiglottoplasty for the treatment of laryngomalacia. Ann Otol Rhinol Laryngol 1987;96:72–6.
20. Mancuso RF, Choi SS, Zalzal GH, Grundfast KM. Laryngomalacia. The search for the second lesion. Arch Otolaryngol Head Neck Surg 1996;122:302–6.
21. Bluestone CD, Healy GB, Cotton RT. Diagnosis of laryngomalacia is not enough [letter, comment]. Arch Otolaryngol Head Neck Surg 1996;122:1417–8.
22. Holinger LD. Etiology of stridor in the neonate, infant and child. Ann Otol Rhinol Laryngol 1980; 89:397–400.
23. Holinger PH, Johnson KC, Schiller F. Congenital anomalies of the larynx. Ann Otol Rhinol Laryngol 1954;63:581–606.

24. Reardon TJ. Congenital laryngeal stridor. Am J Med Sci 1907;134:242–52.
25. Blackader AD, Muckleston HS. Inspiratory stridor and dyspnea in infants. Arch Pediatr 1909;26: 401–16.
26. Iglauer S. Epiglottidectomy for the relief of congenital laryngeal stridor, with report of a case. Laryngoscope 1922;32:56–9.
27. Fearon B, Ellis D. The management of long term airway problems in infants and children. Ann Otol Rhinol Laryngol 1971;80:669–77.
28. Narcy P, Bobin S, Contencin P. Anomalies laryngies du nouveau-ne apropos de 687 observations. Ann Otolaryngol Chir Cervicofac 1984;101:363–73.
29. Cagnol C, Garcin M, Unal D. Un cas de stridor larynge congenital grave laoite pan hyomandibulopexie. J Fr Otorhinolaryngol 1971;20:625–6.
30. Templer J, Hast M, Thomas JR, Davis WE. Congenital laryngeal stridor secondary to flaccid epiglottis, anomalous accessory cartilages and redundant aryepiglottic folds. Laryngoscope 1981;91: 394–7.
31. Lane RW, Weider DJ, Steinem C, Marin-Padilla M. Laryngomalacia. A review and case report of surgical treatment with resolution of pectus excavatum. Arch Otolaryngol 1984;110:546–51.
32. Seid AB, Park SM, Kearns MJ, Gugenheim S. Laser division of the aryepiglottic folds for severe laryngomalacia. Int J Pediatr Otorhinolaryngol 1985;10: 153–8.
33. Katin LI, Tucker JA. Laser supraarytenoidectomy for laryngomalacia with apnea. Trans Pa Acad Ophthalmol Otolaryngol 1990;42:985–8.
34. Kelly SM, Gray SD. Unilateral endoscopic supraglottoplasty for severe laryngomalacia. Arch Otolaryngol Head Neck Surg 1995;121:1351–4.
35. Montreuil F. Bifid epiglottis, report of a case. Laryngoscope 1949;59:194–9.
36. DelMonico ML, Haar JG. Bifid epiglottis. Report of a case. Arch Otolaryngol 1972;96:178–81.
37. Graham JM, Brown FE, Saunders RL, et al. Bifid epiglottis, hand anomalies, and congenital hypopituitarism [letter]. Lancet 1985;2:443.
38. Tobeck A. Halbseitige Aplasie der Epiglottis. Z Laryng Rhinol Otol 1949;28:502–3.
39. Healy GB, Holt GP, Tucker JA. Bifid epiglottis: a rare laryngeal anomaly. Laryngoscope 1976;86:1459–68.
40. Mar S, Essex C. Type 2 tracheal agenesis without tracheoesophageal fistula. Int J Pediatr Otorhinolaryngol 1997;39:159–61.
41. Gatti WM, MacDonald E, Orfei E. Congenital laryngeal atresia. Laryngoscope 1987;97:966–9.
42. Holinger PH, Brown WT. Congenital webs, cysts, laryngoceles and other anomalies of the larynx. Ann Otol Rhinol Laryngol 1967;76:744–52.
43. Woo P, Karmody CS. Congenital laryngeal atresia. Histopathologic study of two cases. Ann Otol Rhinol Laryngol 1983;92:391–5.
44. McIlwain JC. The posterior glottis. J Otolaryngol 1991;20 Suppl 2:1–24.
45. Asano S, Yamagiwa K, Konnai I. Congenital laryngeal atresia: two autopsy cases, one describing the use of computed tomography. APMIS 2000;108: 313–7.
46. Dolkart LA, Reimers FT, Wertheimer IS, Wilson BO. Prenatal diagnosis of laryngeal atresia. J Ultrasound Med 1992;11:496–8.
47. Watson WJ, Thorp JM, Miller RC, et al. Prenatal diagnosis of laryngeal atresia. Am J Obstet Gynecol 1990;163:1456–7.
48. Garel C, Legrand I, Elmaleh M, et al. Laryngeal ultrasonography in infants and children: anatomical correlation with fetal preparations. Pediatr Radiol 1990; 20:241–4.
49. Balci S, Altinok G, Ozaltin F, et al. Laryngeal atresia presenting as fetal ascites, olygohydramnios and lung appearance mimicking cystic adenomatoid malformation in a 25-week-old fetus with Fraser syndrome. Prenat Diagn 1999;19:856–8.
50. Morrison PJ, Macphail S, Williams D, et al. Laryngeal atresia or stenosis presenting as second-trimester fetal ascites—diagnosis and pathology in three independent cases. Prenat Diagn 1998;18:963–7.
51. Smith II, Bain AD. Congenital atresia of the larynx. Ann Otol Rhinol Laryngol 1965;74:338–49.
52. Kalache KD, Chaoui R, Tennstedt C, Bollmann R. Prenatal diagnosis of laryngeal atresia in two cases of congenital high airway obstruction syndrome (CHAOS). Prenat Diagn 1997;17:577–81.
53. Hedrick MH, Ferro MM, Filly RA, et al. Congenital high airway obstruction syndrome (CHAOS): a potential for perinatal intervention. J Pediatr Surg 1994;29:271–4.
54. Peison B, Levitzky E, Sprowls JJ. Tracheoesophageal fistula associated with tracheal atresia and malformation of the larynx. J Pediatr Surg 1970;5:464–7.
55. Cohen MS, Rothschild MA, Moscoso J, Shlasko E. Perinatal management of unanticipated congenital laryngeal atresia. Arch Otolaryngol Head Neck Surg 1998;124:1368–71.

56. Gerson CR, Tansek K, Tucker GF Jr. Pharyngolaryngeal web. Report of a new anomaly. Ann Otol Rhinol Laryngol 1983;92:331–2.
57. Holinger LD, Wong HW, Hemenway WG. Simultaneous glottic and supraglottic laryngeal webs: report of a case. Arch Otolaryngol 1975;101:496–7.
58. Benjamin B, Mair EA. Congenital interarytenoid web. Arch Otolaryngol Head Neck Surg 1991;117: 1118–22.
59. McHugh HE, Loch WE. Congenital webs of the larynx. Laryngoscope 1942;52:43–65.
60. McGuirt WF, Salmon J, Blalock D. Normal speech for patients with laryngeal webs: an achievable goal. Laryngoscope 1984;94:1176–9.
61. Cohen SR. Congenital glottic webs in children. A retrospective review of 51 patients. Ann Otol Rhinol Laryngol Suppl 1985;121:2–16.
62. Benjamin B. Chevalier Jackson Lecture. Congenital laryngeal webs. Ann Otol Rhinol Laryngol 1983; 92:317–26.
63. Milczuk HA, Smith JD, Everts EC. Congenital laryngeal webs: surgical management and clinical embryology. Int J Pediatr Otorhinolaryngol 2000;52:1–9.
64. Parker DA, Das Gupta AR. An endoscopic Silastic keel for anterior glottic webs. J Laryngol Otol 1987; 101:1055–61.
65. Mouney DF, Lyons GD. Fixation of laryngeal stents. Laryngoscope 1985;95:905–7.
66. Laff HI. Cysts in the ventricular area of the larynx. Laryngoscope 1953;63:227–40.
67. Davidson JI. Congenital cyst of the larynx. Lancet 1943;2:508–9.
68. Ahlen G, Ranstrom S. Congenital laryngeal cysts in early infancy. Acta Otolaryngol (Stockh) 1944;32: 483–95.
69. Beautyman W, Haidak GL, Taylor M. Laryngopyocele: report of a fatal case. N Engl J Med 1959; 260:1025–7.
70. Horowitz S. Laryngoceles. J Laryngol Otol 1951; 65:724–34.
71. Hockmuth LN, Martin SJ. An obstructing laryngeal cyst in a newborn. Conn Med 1962;26:691.
72. DeSanto LW. Laryngocele, laryngeal mucocele, large saccules, and laryngeal saccular cysts: a developmental spectrum. Laryngoscope 1974;84:1291–6.
73. Holinger LD, Barnes DR, Smid LJ, Holinger PH. Laryngocele and saccular cysts. Ann Otol Rhinol Laryngol 1978;87:675–85.
74. Broyles EN. Anatomical observations concerning the laryngeal appendix. Ann Otol Rhinol Laryngol 1959;68:461–70.
75. DeSanto LW, Devine KD, Weiland LH. Cysts of larynx: classification. Laryngoscope 1970;80:145–76.
76. Suehs OW, Powell DB. Congenital cyst of the larynx in infants. Laryngoscope 1967;77:654–62.
77. Mitchell DB, Irwin BC, Bailey CM, Evans JN. Cysts of the infant larynx. J Laryngol Otol 1987; 101:833–7.
78. Holinger PH, Brown WT. Congenital webs, cysts, laryngoceles and other anomalies of the larynx. Ann Otol Rhinol Laryngol 1967;76:744–52.
79. Shackelford GD, McAlister WH. Congenital laryngeal cyst. AJR Am J Roentgenol 1972;114:289–92.
80. Cotton RT, Richardson MA. Congenital laryngeal anomalies. Otolaryngol Clin North Am 1981; 14:203–18.
81. English GM, DeBlanc GB. Laryngocele: a case presenting with acute airway obstruction. Laryngoscope 1968;78:386–95.
82. Ferguson GB. Laryngocele. Laryngoscope 1967; 77:1368–75.
83. Donegan JO, Strife JL, Seid AB, et al. Internal laryngocele and saccular cysts in children. Ann Otol Rhinol Laryngol 1980;89:409–13.
84. Canalis RF, Maxwell DS, Hemenway WG. Laryngocele—an updated review. J Otolaryngol 1977;6: 191–9.
85. Chu L, Gussack GS, Orr JB, Hood D. Neonatal laryngoceles. A cause for airway obstruction. Arch Otolaryngol Head Neck Surg 1994;120:454–8.
86. Baker HL, Baker SR, McClatchey KD. Manifestations and management of laryngoceles. Head Neck Surg 1982;4:450–6.
87. Trapnell DH. The radiological diagnosis of laryngoceles. Clin Radiol 1962;13:68–72.
88. Yarington CT, Frazer JP. An approach to the internal laryngocele and other submucosal lesions of the larynx. Ann Otol Rhinol Laryngol 1966;75:956–60.
89. Thomas DM, Madden GJ. Bilateral laryngoceles. Ear Nose Throat J 1993;72:819–21.
90. Parsons DS, Herr T. Delayed diagnosis of a laryngotracheoesophageal cleft. Int J Pediatr Otorhinolaryngol 1997;39:169–73.
91. Parsons DS, Stivers FE, Giovanetto DR, Phillips SE. Type I posterior laryngeal clefts. Laryngoscope 1998;108:403–10.
92. Samuel M, Burge DM, Griffiths DM. Prenatal diagnosis of laryngotracheoesophageal clefts. Fetal Diagn Ther 1997;12:260–5.
93. Cohen SR, Thompson JW. Ventral cleft of the larynx: a rare congenital laryngeal defect. Ann Otol Rhinol Laryngol 1990;99:281–5.

94. Decreton SJ, Clement PA. Comparative study of standard x-ray of the maxillary sinus and sinuscopy in children. Rhinology 1981;19:155–9.
95. Pattisapu JV, Parent AD. Subdural empyemas in children. Pediatr Neurosci 1987;13:251–4.
96. Roth B, Rose KG, Benz-Bohm G, Gunther H. Laryngo-tracheo-oesophageal cleft. Clinical features, diagnosis and therapy. Eur J Pediatr 1983; 140:41–6.
97. Finlay HV. Familial congenital stridor. Arch Dis Child 1949;24:219–23.
98. Evans JN. Management of the cleft larynx and tracheoesophageal clefts. Ann Otol Rhinol Laryngol 1985;94:627–30.
99. Jahrsdoerfer RA, Kirchner JA, Thaler SU. Cleft larynx. Arch Otolaryngol 1967;86:108–13.
100. Cohen SR. Cleft larynx. A report of seven cases. Ann Otol Rhinol Laryngol 1975;84:747–56.
101. Garel C, Hassan M, Hertz-Pannier L, et al. Contribution of MR in the diagnosis of 'occult' posterior laryngeal cleft. Int J Pediatr Otorhinolaryngol 1992;24:177–81.
102. Delahunty JE, Cherry J. Congenital laryngeal cleft. Ann Otol Rhinol Laryngol 1969;78:96–106.
103. Corbally MT. Laryngo-tracheo-oesophageal cleft. Arch Dis Child 1993;68:532–3.
104. Holinger LD, Tansek KM, Tucker GF Jr. Cleft larynx with airway obstruction. Ann Otol Rhinol Laryngol 1985;94:622–6.
105. Tucker GF Jr, Maddalozzo J. "Occult" posterior laryngeal cleft. Laryngoscope 1987;97:701–4.
106. Tyler DC. Laryngeal cleft: report of eight patients and a review of the literature. Am J Med Genet 1985;21:61–75.
107. Walner DL, Stern Y, Collins M, et al. Does the presence of a tracheoesophageal fistula predict the outcome of laryngeal cleft repair? Arch Otolaryngol Head Neck Surg 1999;125:782–4.
108. Petterson G. Inhibited separation of larynx and the upper part of trachea from oesophagus in a newborn. Acta Chir Scand 1955;110:250–4.
109. Koltai PJ, Morgan D, Evans JN. Endoscopic repair of supraglottic laryngeal clefts. Arch Otolaryngol Head Neck Surg 1991;117:273–8.
110. Nuutinen J, Karja J, Karjalainen P. Measurements of impaired mucociliary activity in children. Eur J Respir Dis 1983; Suppl 128:454–6.
111. Pettit PN, Butcher RB, Bethea MC, Danks DM. Surgical correction of complete tracheoesophageal cleft. Laryngoscope 1979;89:804–11.
112. Froehlich P, Truy E, Stamm D, et al. Cleft larynx: management and one-stage surgical repair by anterior translaryngotracheal approach in two children. Int J Pediatr Otorhinolaryngol 1993;27:73–8.
113. Robie DK, Pearl RH, Gonsales C, et al. Operative strategy for recurrent laryngeal cleft: a case report and review of the literature. J Pediatr Surg 1991; 26:971–3; discussion 973.
114. Prescott CA. Cleft larynx: repair with a posterior cartilage graft. Int J Pediatr Otorhinolaryngol 1995;31:91–4.
115. Ahmad R, Horwitz PE, Sami KA, Rabeeah A. A bifurcated endobronchial tube in the management of laryngotracheo-oesophageal cleft repair [published erratum appears in Br J Anaesth 1994; 72:371]. Br J Anaesth 1993;70:696–8.
116. Geiduschek JM, Inglis AF, O'Rourke PP, et al. Repair of a laryngotracheoesophageal cleft in an infant by means of extracorporeal membrane oxygenation. Ann Otol Rhinol Laryngol 1993;102:827–33.
117. Cotton RT, Schreiber JT. Management of laryngotracheoesophageal cleft. Ann Otol Rhinol Laryngol 1981;90:401–5.
118. Cohen SR, Geller KA, Birns JW, Thompson JW. Laryngeal paralysis in children: a long-term retrospective study. Ann Otol Rhinol Laryngol 1982; 91:417–24.
119. Swift AC, Rogers J. Vocal cord paralysis in children. J Laryngol Otol 1987;101:169–71.
120. Phelan PD. "Oscopy" in children: laryngoscopy and bronchoscopy. Aust Fam Physician 1979;8:853–7.
121. Hoffman HJ, Hendrick EB, Humphreys RP. Manifestations and management of Arnold-Chiari malformation in patients with myelomeningocele. Childs Brain 1975;1:255–9.
122. Krieger AJ, Detwiler JS, Trooskin SZ. Respiratory function in infants with Arnold-Chiari malformation. Laryngoscope 1976;86:718–23.
123. Cavanagh F. Vocal palsies in children. J Laryngol Otol 1955;69:399.
124. Dedo DD. Pediatric vocal cord paralysis. Laryngoscope 1979;89:1378–84.
125. Holinger LD, Holinger PC, Holinger PH. Etiology of bilateral abductor vocal cord paralysis: a review of 389 cases. Ann Otol Rhinol Laryngol 1976; 85:428–36.
126. Gardner E, O'Rahilly R, Prolo D. The Dandy-Walker and Arnold-Chiari malformations. Clinical, developmental, and teratological considerations. Arch Neurol 1975;32:395–407.

127. Mackenzie IJ, Kerr AI, Cowan DL. A review of endoscopies of the respiratory tract and oesophagus in a children's hospital. Health Bull (Edinb) 1984;42:78–80.
128. Richardson MA, Cotton RT. Anatomic abnormalities of the pediatric airway. Ear Nose Throat J 1985;64:47–60.
129. Holinger PC, Holinger LD, Reichert TJ, Holinger PH. Respiratory obstruction and apnea in infants with bilateral abductor vocal cord paralysis, meningomyelocele, hydrocephalus, and Arnold-Chiari malformation. J Pediatr 1978;92:368–73.
130. Bluestone CD, Delerme AN, Samuelson GH. Airway obstrucion due to vocal cord paralysis in infants with hydrocephalus and meningomyelocele. Ann Otol Rhinol Laryngol 1972;81: 778–84.
131. Gacek RR. Hereditary abductor vocal cord paralysis. Ann Otol Rhinol Laryngol 1976;85:90–3.
132. Watters GV, Fitch N. Familial laryngeal abductor paralysis and psychomotor retardation. Clin Genet 1973;4:429–33.
133. Plott D. Congenital laryngeal-abductor paralysis due to nucleus ambiguus dysgenesis in three brothers. N Engl J Med 1964;27:593–7.
134. Cohen SR, Geller KA, Birns JW, Thompson JW. Laryngeal paralysis in children: a long-term retrospective study. Ann Otol Rhinol Laryngol 1982; 91:417–24.
135. Holinger PC, Vuckovich DM, Holinger LD, Holinger PH. Bilateral abductor vocal cord paralysis in Charcot-Marie-Tooth disease. Ann Otol Rhinol Laryngol 1979;88:205–9.
136. Bressler KL, Kaiser PC, Dunham ME, Holinger LD. Primary closure of persistent tracheocutaneous fistula in children. Ann Otol Rhinol Laryngol 1994; 103:835–7.
137. Cotton RT, Evans JN. Laryngotracheal reconstruction in children. Five-year follow-up. Ann Otol Rhinol Laryngol 1981;90:516–20.
138. Holinger LD. Treatment of severe subglottic stenosis without tracheotomy: a preliminary report. Ann Otol Rhinol Laryngol 1982;91:407–12.
139. Hawkins DB. Hyaline membrane disease of the neonate prolonged intubation in management: effects on the larynx. Laryngoscope 1978;88: 201–24.
140. Papsidero MJ, Pashley NR. Acquired stenosis of the upper airway in neonates. An increasing problem. Ann Otol Rhinol Laryngol 1980;89:512–4.
141. Parkin JL, Stevens MH, Jung DL. Acquired and congenital subglottic stenosis in the infant. Ann Otol Rhinol Laryngol 1976;85:573–81.
142. Whited RE. A prospective study of laryngotracheal sequelae in long-term intubation. Laryngoscope 1984;94:367–77.
143. Benjamin B. Prolonged intubation injuries of the larynx: endoscopic diagnosis, classification, and treatment. Ann Otol Rhinol Laryngol Suppl 1993; 160:1–15.
144. Holinger PH, Kutnick SL, Schild JA, Holinger LD. Subglottic stenosis in infants and children. Ann Otol Rhinol Laryngol 1976;85:591–9.
145. Sasaki CT, Horiuchi M, Koss N. Tracheostomy-related subglottic stenosis: bacteriologic pathogenesis. Laryngoscope 1979;89:857–65.
146. Healy GB. An experimental model for the endoscopic correction of subglottic stenosis with clinical applications. Laryngoscope 1982;92: 1103–15.
147. Cotton RT, Silver P, Nuwayhid NS. Chronic laryngeal and tracheal stenosis. In: Paparella MM, Shumrick DA, editors. Otolaryngology. 2nd ed. Philadelphia: WB Saunders; 1980. p. 2931–50.
148. Fee WE, Wilson GG. Tracheoesophageal space abscess. Laryngoscope 1979;89:377–84.
149. Lindholm CE. Prolonged endotracheal intubation, a valuable alternative to tracheostomy. Bronches 1968;18:398–408.
150. Pashley NR. Risk factors and prediction of outcome in acquired subglottic stenosis in children. Int J Pediatr Otorhinolaryngol 1982;4:1–6.
151. McMillan WG, Duvall AJ. Congenital subglottic stenosis. Arch Otolaryngol 1968;87:272–8.
152. Marshak G, Grundfast KM. Subglottic stenosis. Pediatr Clin North Am 1981;28:941–8.
153. Noyek AM. Xeroradiography in the assessment of the pediatric larynx and trachea. J Otolaryngol 1976;5:468–74.
154. Cotton RT. Management of subglottic stenosis in infancy and childhood: review of a consecutive series of cases managed by surgical reconstruction. Ann Otol Rhinol Laryngol 1978;87:649.
155. Cotton RT, Seid AB. Management of the extubation problem in the premature child. Anterior cricoid split as an alternative to tracheotomy. Ann Otol Rhinol Laryngol 1980;89:508–11.
156. Cotton RT. Prevention and management of laryngeal stenosis in infants and children. J Pediatr Surg 1985;20:845–51.

157. Holinger LD, Stankiewicz JA, Livingston GL. Anterior cricoid split: the Chicago experience with an alternative to tracheotomy. Laryngoscope 1987; 97:19–24.
158. Lusk RP, Kang DR, Muntz HR. Auricular cartilage grafts in laryngotracheal reconstruction. Ann Otol Rhinol Laryngol 1993;102:247–54.
159. Lusk RP, Gray S, Muntz HR. Single-stage laryngotracheal reconstruction. Arch Otolaryngol Head Neck Surg 1991;117:171–3.
160. Richardson MA, Inglis AF. A comparison of anterior cricoid split with and without costal cartilage graft for acquired subglottic stenosis. Int J Pediatr Otorhinolaryngol 1991;22:187–93.
161. Otherson HB. Steroid therapy for tracheal stenosis in children. Ann Thorac Surg 1974;17:254.
162. Kirchner FR, Toledo PS. Microcauterization in otolaryngology. Arch Otolaryngol 1974;99:198–202.
163. Rodgers BM, Talbert JL. Clinical applications of endotracheal cryotherapy. J Pediatr Surg 1978;13:662–8.
164. Strome M, Donahoe PK. Advances in management of laryngeal and subglottic stenosis. J Pediatr Surg 1982;17:591–6.
165. Downing TP, Johnson DG. Excision of subglottic stenosis with the urethral resectoscope. J Pediatr Surg 1979;14:252–7.
166. Simpson GT, Strong MS, Healy GB, et al. Predictive factors of success or failure in the endoscopic management of laryngeal and tracheal stenosis. Ann Otol Rhinol Laryngol 1982;91:384–8.
167. Friedman EM, Healy GB, McGill TJ. Carbon dioxide laser management of subglottic and tracheal stenosis. Otolaryngol Clin North Am 1983;16: 871–7.
168. Healy GB, McGill T, Simpson GT, Strong MS. The use of the carbon dioxide laser in the pediatric airway. J Pediatr Surg 1979;14:735–40.
169. Koufman JA, Thompson JN, Kohut RI. Endoscopic management of subglottic stenosis with the CO_2 surgical laser. Otolaryngol Head Neck Surg 1981; 89:215–20.
170. Lyons GD, Owens R, Lousteau RJ, Trail ML. Carbon dioxide laser treatment of laryngeal stenosis. Arch Otolaryngol 1980;106:255–6.
171. Strong MS, Healy GB, Vaughan CW, et al. Endoscopic management of laryngeal stenosis. Otolaryngol Clin North Am 1979;12:797–805.
172. Cotton RT, Manoukian JJ. Glottic and subglottic stenosis. In: Cummings CW, Fredrickson JM, Harker LA, et al, editors. Otolaryngology-head and neck surgery. St. Louis: CV Mosby; 1986. p. 2159–80.
173. Seid AB, Pransky SM, Kearns DB. One-stage laryngotracheoplasty. Arch Otolaryngol Head Neck Surg 1991;117:408–10.
174. Looper EA. The use of hyoid bone as graft in laryngeal stenosis. Arch Otolaryngol 1938;28:106.
175. Bennett T. Laryngeal strictures. South Med J 1960;53:1101.
176. Finnegan DA, Wong ML, Kashima HK. Hyoid autograft repair of chronic subglottic stenosis. Ann Otol Rhinol Laryngol 1975;84:643–9.
177. Thawley SE, Ogura JH. Panel discussion: the management of advanced laryngotracheal stenosis. Use of the hyoid graft for treatment of laryngotracheal stenosis. Laryngoscope 1981;91:226–32.
178. Ward PH, Canalis R, Fee W, Smith G. Composite hyoid sternohyoid muscle grafts in humans: its use in reconstruction of subglottic stenosis and the anterior tracheal wall. Arch Otolaryngol Head Neck Surg 1977;103:531–4.
179. Cotton RT. Pediatric laryngotracheal stenosis. J Pediatr Surg 1984;19:699–704.
180. Fearon B, Cinnamond M. Surgical correction of subglottic stenosis of the larynx: clinical results of the Fearon-Cotton operation. J Otolaryngol 1976;5:475–8.
181. Fearon B, Cotton RT. Surgical correction of subglottic stenosis of the larynx in infants and children. Progress report. Ann Otol Rhinol Laryngol 1974;83:428–31.
182. Ludemann JP, Hughes CA, Noah Z, Holinger LD. Complications of pediatric laryngotracheal reconstruction: prevention strategies. Ann Otol Rhinol Laryngol 1999;108:1019–26.
183. Cotton RT, Gray SD, Miller RP. Update of the Cincinnati experience in pediatric laryngotracheal reconstruction. Laryngoscope 1989;99:1111–6.
184. Benjamin B. Congenital disorders. In: Cummings CW, Fredrickson JM, Harker LA, et al, editors. Otolaryngology-head and neck surgery. St. Louis: CV Mosby; 1986. p. 2329–38.
185. Wooley AL, Lusk RP. Multiple vascular lesions of an infant's airway. Otolaryngol Head Neck Surg 1994;111:305–8.

CHAPTER 46

Congenital Anomalies of the Head and Neck

Lee D. Rowe, MD

Congenital head and neck anomalies represent a diverse group of clinical disorders. Frequently, these present as upper aerodigestive tract neoplasms and neck masses, with thyroglossal duct abnormalities most common, followed by branchial arch defects, lymphangiomas (cystic hygromas), and subcutaneous vascular anomalies (hemangiomas, arteriovenous malformations). Less common are teratomas, heterotopic neural tissue, and nasopharyngeal neoplasms. Additional disorders encountered include congenital disorders of the oral cavity including cleft lip and palate, Pierre Robin sequence (PRS), and other aberrations such as primary ciliary dyskinesia, Kartagener's syndrome, and craniofacial anomalies.

CONGENITAL DISORDERS OF THE NECK

Branchial Cleft Anomalies

Lateral cervical lesions, termed branchial cleft cysts, are congenital developmental defects that arise from the primitive branchial apparatus (branchial arch, cleft, and pouches). First branchial cleft anomalies are discussed elsewhere in this chapter.

Embryogenesis The branchial arches consist of five parallel mesodermal bars, each with its nerve supply and blood vessel (primitive aortic arches that develop during the third and fourth embryonic weeks). The branchial arches are separated externally by branchial clefts consisting of ectoderm and internally by endodermally lined branchial (pharyngeal) pouches. A branchial plate is located between each arch, separating pouch from cleft. The nerves are anterior to their respective arteries, except in the fifth arch, where the nerve is posterior to the artery. Caudal to all of the arches is the twelfth nerve, which arises from the epicardial ridge. The sternocleidomastoid muscle is derived from cervical somites posterior and inferior to the foregoing arches. In each arch, a central artery develops, connecting the two ventral and two dorsal aortas. The two ventral aortas fuse completely, whereas the two dorsal aortas only fuse caudally. With continued development, the ventral aortas become the external and common carotid arteries, whereas the arteries of the first and second arch degenerate. Segments of the dorsal aortas persist as the internal carotid arteries along with the artery of the third arch. The left fourth arch artery becomes the arch of the aorta, and the right fourth arch artery becomes the proximal portion of the subclavian artery. The primitive clefts pass between the corresponding arteries.[1] The most widely accepted theory of the genesis of branchial cleft anomalies is that fistulae, sinuses, and cysts result from incomplete closure of the connection between the cleft and the pouch, with rupture of the branchial plate. Recently, an induced first and second branchial arch syndrome in an animal model was shown to be secondary to disturbed migration of neural crest tissue during early embryogenesis.[2] Additionally, targeted mutations of two related basic helix-loop-helix transcription factors in the developing ventricles of the heart in mice have been associated with branchial arch hypoplasia, suggesting a molecular basis for branchial cleft anomalies.[3] These anomalies are generally lined with stratified squamous epithelium containing subepithelial lymphoid follicles, keratin, hair follicles, sweat glands, cartilage, and sebaceous glands.

Clinical Presentation Typically, branchial clefts present as smooth, round, fluctuant, nontender masses along the anterior border of the sternocleidomastoid muscle, anywhere from the external auditory canal to the clavicle. During upper respiratory tract infections, a painful increase in size is common and occasionally may be associated with external drainage through an unrecognized fistula. Small

cysts may not be recognized until the second decade in life. Male and female incidences are equal, and nearly all of these lesions are recognized by the time the patient reaches 30 years of age.

Second Branchial Cleft Anomalies Second branchial cleft lesions are the most common anomalies and are encountered most frequently in the anterior triangle of the neck along the anterior border of the sternocleidomastoid muscle inferior to the angle of the mandible. The fistula tract, if present, ascends along the carotid sheath, crosses over the hypoglossal and glossopharyngeal nerves, and courses between the internal and external carotid arteries. It ends in the tonsillar fossa, a second branchial pouch derivative.

Second pouch remnants may form blind sinuses in the pharyngeal tonsil and frequently cause recurrent unilateral tonsillitis. Treatment requires tonsillectomy and excision of the sinus tract.

Third Branchial Cleft Anomalies Third branchial cleft anomalies are unusual and constitute less than 1% of all branchial cleft anomalies.[4] The external ostium occurs at the same position as the second branchial cleft anomaly, and the cyst may be located anywhere along the fistula. The fistula tract ascends along the carotid sheath behind the internal carotid artery and over the hypoglossal nerve, and it enters the piriform sinus, piercing the middle constrictor muscle below the glossopharyngeal nerve. Clinically, the anomaly may mimic suppurative thyroiditis or symptoms of an external laryngocele and can cause recurrent infection, discharge, and, rarely, stridor.

Fourth Branchial Cleft Anomalies Fourth branchial cleft anomalies are extremely rare. Only 40 cases have been reported. The fistula tract descends along the carotid sheath, enters the chest, passing under either the aortic arch on the left or the subclavian artery on the right, and ascends in the neck to open at the apex of the piriform sinus. The sinus tract or cyst may become clinically manifested by recurrent episodes of neck abscess or acute suppurative thyroiditis (particularly in infants). A left predominance has been reported in one series.[5]

Treatment Complete surgical excision is the treatment of choice for branchial cleft anomalies and is indicated for recurrent infection, cosmetic deformity, and potential for malignant degeneration. Preoperative assessment with computed tomographic (CT) scanning or magnetic resonance imaging (MRI) is essential and may be combined with a fistulogram or pharyngoesophagram when indicated. A "stepladder" surgical approach is used to avoid long, cosmetically deforming incisions, paralleling the sternocleidomastoid muscle. To avoid recurrence, combined endoscopic examination for a pharyngeal pouch and sinus tract with meticulous dissection of a sinus tract, if present, will facilitate complete resection.

Branchiogenic Carcinoma The hypothesis that squamous cell carcinoma arises in a branchial cleft cyst (branchiogenic carcinoma) is controversial. Cystic squamous cell carcinoma presenting in the neck without an apparent primary is almost universally secondary to metastasis from a neoplasm arising in the faucial or lingual tonsillar crypt epithelium or nasopharyngeal tissue.[6] Malignant transformation in a branchial cleft cyst is therefore a rarity. Management is wide excision of the tumor and ipsilateral radical neck dissection followed by radiation therapy.[7]

Lymphangiomas, Hemangiomas, and Vascular Malformations

Congenital lymphangiomatous malformations of the head and neck represent a wide clinical spectrum. Lymphangiomas result from abnormal development of the lymphatic system at sites of lymphatic-venous connection, with obstruction of lymph drainage from the affected area causing multicystic endothelium-lined spaces. The neck is the most common site (25% of all cases). Over half of these lesions are present at birth, with 90% becoming apparent by 2 years of age. Those lymphangiomas arising above the mylohyoid tend to extend from skin to mucosa and are more infiltrative, whereas those below are more discreet and cystic.

Cystic Hygroma Cystic hygromas are large lymphangiomas most commonly found in the posterior triangle of the neck and axilla in children.[8] Cervical cystic hygromas commonly appear before 30 weeks gestation and are usually associated with karyotypic abnormalities, various malformation syndromes, and several teratogenic agents.[9] The prognosis is poor for these types of hygromas. By contrast, cys-

tic hygroma developing late in pregnancy has a more favorable outcome and is more likely to be encountered by the head and neck surgeon. Cystic hygromas are soft, painless, and compressible masses that may increase when the patient cries. Two-thirds are asymptomatic. After an upper respiratory tract infection, however, sudden enlargement with inflammation, infection, dysphagia, and stridor may develop. This is more commonly seen if the anterior triangle of the neck is involved or in patients with pharyngolaryngeal extension or intraoral involvement.

Vascular Lesions Other congenital vascular lesions of the head and neck include hemangiomas and arteriovenous malformations or lymphovenous lesions. Diagnostic imaging of these anomalies is based on the need for surgical treatment. Only those lesions that cause functional impairment or developmental disturbance are surgically addressed. Angiography, combined with MRI, allows separation into low-flow lesions (hemangiomas, venous, and lymphatic malformations) and high-flow lesions (arteriovenous malformations and invasive, combined lymphovascular malformations).

Treatment Treatment of small cystic hygromas that have not regressed but have enlarged is by surgical excision with staged debulking of larger cystic hygromas. Neural structures such as the facial nerve, vagus nerve, and phrenic nerves should not be sacrificed. Recurrence is uncommon when gross neoplasm is removed. Hemangiomas are frequently multiple with the parotid gland, a common site of occurrence. They grow rapidly for 1 to 2 years and then stabilize, and often regress spontaneously. Their growth may be slowed or inhibited with corticosteroids or interferon. Surgical excision is reserved for severe cosmetic deformity or life-threatening aerodigestive tract impairment.

For low-flow lesions, sclerosant therapy is effective, either alone in small lesions or combined with surgical resection or embolization. Preoperative embolization at the time of selective angiography and surgical excision are the treatment of choice in high-flow malformations.

Aberrant Thyroid Tissue

Ectopic thyroid tissue can occur anywhere from the foramen cecum to the lower neck (Figure 46–1). Most frequently, it occurs as a thyroglossal duct cyst associated with a normal thyroid gland. Less common is total ectopia, appearing as a lingual thyroid. True lateral neck thyroid ectopia lateral to the carotid artery and jugular vein cannot be confirmed based on embryogenesis.[10] Therefore, laterally situated thyroid tissue should direct attention to the ipsilateral thyroid lobe. If papillary carcinoma is found in an aberrant position, then either ipsilateral thyroid lobectomy with isthmusectomy or total thyroidectomy is performed.

Embryology In the 4-week-old embryo, the primitive thyroid gland begins as a ventral diverticulum of endodermal origin, arising in the floor of the pharynx between the tuberculum impar and the copula. Ultimately, the site of origin of the diverticulum becomes the foramen cecum, and the copula becomes the posterior third of the tongue. The thyroid descends caudally through or adjacent to the primitive hyoid bone. The developing thyroglossal duct reaches its final position in the midline of the

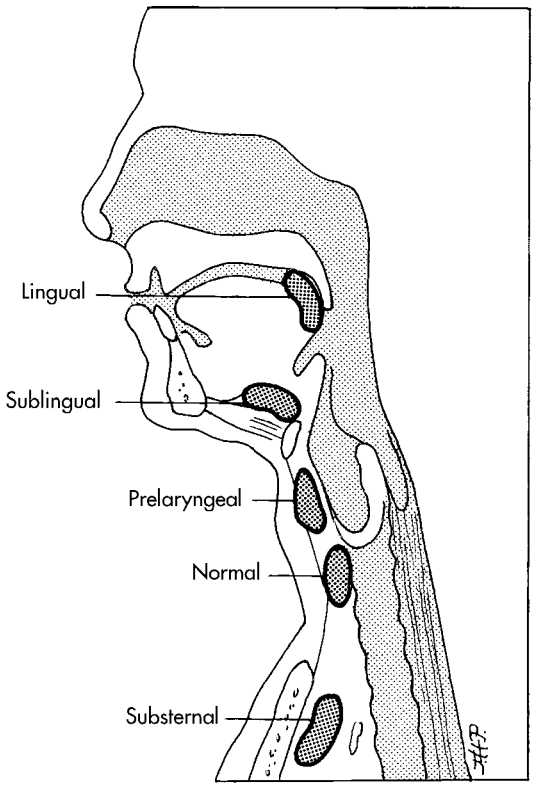


FIGURE 46–1. Most common locations of midline aberrant thyroid tissue schematically illustrated.

neck and develops into the median lobe of the thyroid. The duct normally persists as a hollow stalk for 6 weeks and then atrophies. Thyroglossal duct anomalies result from a failure of complete obliteration of the thyroglossal duct and are located anywhere along the descent of the gland.

Lingual Thyroid Lingual thyroid presents in the midline as a sessile, nontender reddish mass in the base of the tongue, anterior to the valleculae. It is the most frequent benign mass encountered in the oropharynx and, unlike other aberrant thyroid anomalies, has a 7 to 1 female preponderance. Symptoms may include dysphagia, cough, dysphonia, dyspnea, and hemorrhage. Lingual thyroid may become apparent during pregnancy because of increased thyroid function, and approximately 70% of these women have hypothyroidism.

Lingual thyroid is covered with stratified squamous epithelium and often exhibits an abundant aberrant vascular supply. The mass consists of normal or immature thyroid tissue, which may be either functional or dysfunctional. Indications for surgical removal include uncontrollable hyperthyroidism, hemorrhage, symptomatic enlargement, or a question of malignancy. Carcinoma arising in a lingual thyroid is rare, with only 26 cases reported.[11] Preoperatively, a thyroid uptake study and scan should be obtained to determine whether functioning thyroid tissue exists in its normal cervical location. Magnetic resonance imaging may also be of value because it allows multiplanar imaging and provides the best soft tissue definition. Excision is performed through the transhyoid route.

Thyroglossal Duct Anomalies *Signs and Symptons* A thyroglossal duct cyst is the most common congenital neck mass and the second most common of all childhood cervical neck masses. Cysts, sinuses, and fistulae of the thyroglossal duct manifest as anterior midline neck masses from the foramen cecum to the thyroid gland, typically before 20 years of age (most before 10 years). Thyroglossal ducts and fistulae are often asymptomatic. They may also become recurrently infected during upper respiratory tract infections, which can cause cyst enlargement, abscess development, and rupture with external sinus formation. Thyroglossal duct cysts are most commonly found below the hyoid bone and above the thyroid gland, displaying movement with anterior tongue protrusion and swallowing. Usually, the cysts are 1 to 3 cm in diameter and are smooth, round, and fluctuant.

A precise pathogenesis of thyroglossal duct cysts has not been determined. However, a hereditary etiology is suspected. Overall, for non-Americans, a predominantly autosomal dominant pattern of inheritance has been reported in older patients (mean age 13.9 years) versus an autosomal recessive pattern in younger patients (mean age 6.2 years).[12] In the United States, on the other hand, a female preponderance is observed with autosomal dominance and a younger mean age, which may be accounted for by genetic imprinting that is felt to be secondary to variations in DNA methylation.

Histopathologic Features. Thyroglossal duct cysts and fistulae are lined with squamous, ciliated columnar, or transitional cell epithelium. They may be surrounded by fibrous tissue with infiltrating inflammatory cells lacking organized lymphoid tissue. Islands of ectopic thyroid tissue and mucous glands are not uncommonly identified. The cyst or sinus tract is filled with mucoid or mucopurulent material.

Differential Diagnosis. The differential diagnosis includes dermoid cyst, pyramidal lobe hyperplasia or cyst, teratoma, hamartoma, lipoma, sebaceous cyst, cavernous hemangioma, hypertrophic lymph node, and malignant primary or metastatic neoplasm.

Treatment. As with lingual thyroid, a preoperative thyroid scan and uptake study are mandatory, even though concomitant agenesis of the thyroid gland is extremely rare. Malignant degeneration, recurrent infections, undesirable cosmetic appearance, and, rarely, intermittent upper airway obstruction are indications for surgical excision of thyroglossal duct cysts and fistulae. Incision and drainage may be necessary in the interim if an abscess has developed. Because of the high recurrence rate after excision, the Sistrunk procedure is recommended to prevent recurrence.[13] This involves a transverse incision over the cyst or a fusiform incision around an external fistula. All abnormal tissue, including the cyst, fistula, body of the hyoid bone (preferably more than 15 mm), and the fibrous cord extending to the foramen cecum, should be resected. Recurrence devel-

ops in up to 20% of cases as a result of failure to remove all abnormal tissue, including accessory ducts. Malignant degeneration of a thyroglossal duct cyst is rare (0.7%).[14]

Approximately 150 cases of thyroglossal duct carcinoma, predominantly of the papillary type, have been reported. Additional cases of Hürthle cell adenoma have also been described.[15] Most patients with papillary carcinoma are in their forties, although 20% are less than 20 years of age. Typically, the neoplasm is small and clinically unsuspected, and, in rare cases, it may be associated with regional or distant metastasis. Rarely, squamous cell carcinoma arising directly from the lining epithelium of a thyroglossal duct remnant has been reported and, as a rule, has a poor prognosis. Treatment consists of excision of the thyroglossal duct cyst, total or near-total thyroidectomy, ablative radioactive iodine therapy for residual metastatic thyroid neoplasm tissue, and careful interval follow-up with whole-body thyroid scanning.

THYMIC CYSTS

Cervical thymic cysts are extremely rare. Only 84 cases have been reported in the English language, with most occurring asymptomatically in children and adults. Only seven cases have involved patients younger than 1 year of age.[16] Cysts can arise from nests of thymic tissue anywhere along the descent of the thymic primordia from the angle of the mandible to the mediastinum and are primarily located anterior and deep to the middle one-third of the sternocleidomastoid muscle. In up to 50% of cervical thymic cysts, there is a mediastinal connection possibly requiring a sternotomy. Surgery is the definitive treatment (Figure 46–2).[16] Three additional cases of solid, ectopic, cervical thymus have been reported requiring complete excision.[17]

NEUROGENIC NEOPLASMS

Congenital neurogenic lesions involving the head and neck include all neoplasms or anomalies originating in the neural tissue or its covering. Two groups have been recognized: (1) heterotopic brain lesions with developmental defects and (2) neoplasms of neurogenic origin including neurinomas, neurofibromas, neuromas, ganglioneuromas, and meningiomas. Neurinomas originating from the connective tissue sheath of the nerve are termed schwannomas and may appear as multiple neurofibromas arising from cutaneous, visceral, and cranial nerves in neurofibromatosis (NF) 1 or 2. The incidence of head and neck manifestations in patients with NF (a genetic disease) varies between 14 and 37%. In NF1 (classic von Recklinghausen's disease), functional deficits include speech and voice abnormalities, airway obstruction, dysphagia, facial paresis, lip incompetence, and impaired mastication. Patients with NF2 (bilateral acoustic schwannomas) present with hearing loss.[18]

Typically, neuromas appear as solitary, encapsulated lesions with elongated, spindle-shaped cells with an oval or flattened nucleus. Ganglioneuromas, rarely seen in the head or neck, are characterized by ganglion and glial cells. Meningiomas are usually benign and arise from embryonic arachnoid rests. They may appear extracranially at the nasal root or in the sinuses. Whorl-like fibroblastic nuclei with hyaline formations producing a sand-like appearance may be present (psammoma body). Pharyngeal neuromas are rare and appear as a smooth, firm, rounded, and yellow mass. Symptoms depend on the size and location of the tumor. If multiple brown discolorations of the skin or café au lait spots are present, neurofibroma may be suggested. Treatment is surgical excision, and resection must be complete because recurrence is common.

CONGENITAL DISORDERS OF THE NOSE AND PARANASAL SINUSES

Heterotopic Neural Tissue

Heterotopic neural tissue or gliomas may manifest as isolated ectopic brain tissue with only a fibrous band connecting it to the endocranium. A glioma may be of the external or endonasal type. The external nasal glioma is typically found in the nasion as a red, relatively firm, mobile mass located subcutaneously that does not increase in size when the patient cries. The endonasal glioma is less common and arises from the middle turbinate or the lateral nasal vault, where it may be mistaken for a polyp. By contrast, a meningocele is a hernial protrusion of the meninges, and if it contains brain tissue, it is called an encephalocele.

An encephalocele is caused by a defect of the fetal skull and contains an ependyma-lined cavity

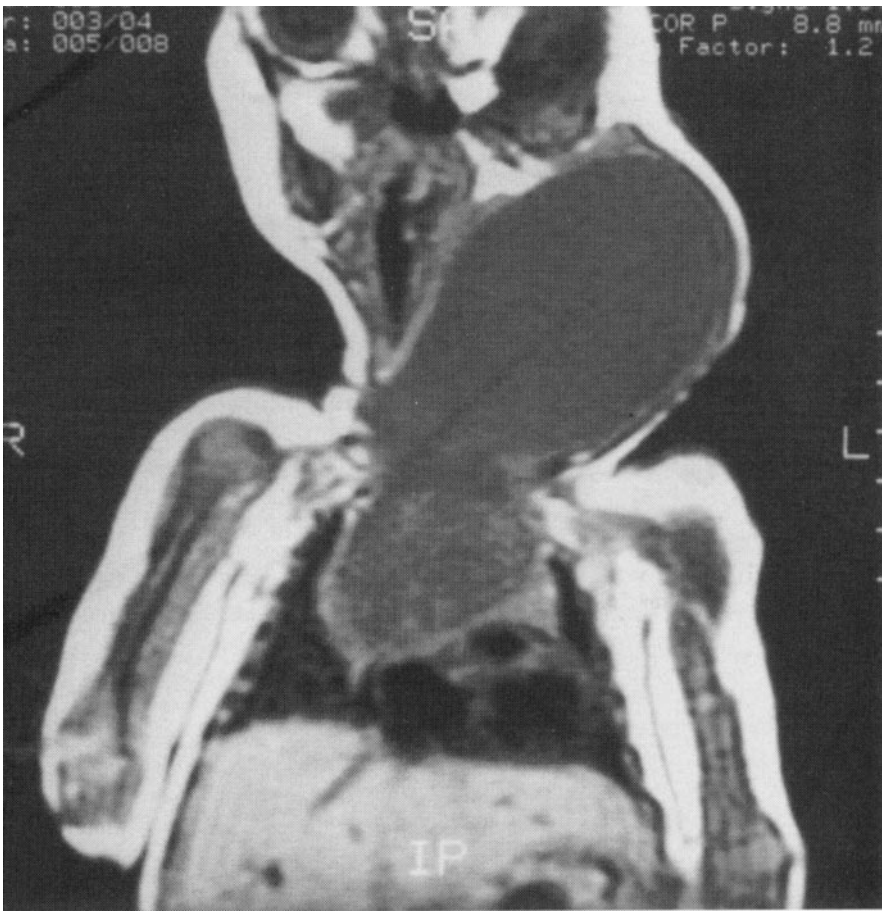


FIGURE 46–2. Coronal T_1-weighted magnetic resonance image of a neonate with a cervical thymic cyst with mediastinal involvement. Reproduced with permission from Nguyen Q et al.[16]

filled with cerebrospinal fluid. Two basic types of hernial protrusions are identified: the sincipital type and the basal type. The sincipital type is uncommon and is associated with termination of the meninges near the base of the nose. This type has three different forms: (1) nasofrontal, in which the encephalocele extends between the nasal and frontal bones, resulting in a midline swelling at the base of the nose; (2) nasoethmoidal, in which a defect among the nasal, frontal, and ethmoid bones allows the encephalocele to appear as a mass beneath the skin of the bony-cartilaginous junction; and (3) naso-orbital protrusion of the encephalocele through the suture lines among the lacrimal, frontal, and ethmoidal bones appearing as a conjunctival mass. The basal type, in which the hernia extends into the naso-orbital and pharyngeal region, is rarer than the sincipital. Three varieties have been identified: (1) sphenoid-pharyngeal, in which the encephalocele extends through the ethmoid or sphenoid bones or their sutures lines into the nasal or nasopharyngeal cavity; (2) sphenoid-orbital, in which the encephalocele extends through the spheno-orbital fissure into the posterior aspect of the orbit, resulting in pulsatile exophthalmos; and (3) sphenomaxillary, in which the encephalocele herniates through the spheno-orbital fissure into the orbit, with extension inferiorly through the inferior orbital fissure into the pterygomaxillary fissure.[19] This results in a mass bulging into the cheek or into the oropharynx, medial to the ramus of the mandible. In all cases, unlike with gliomas, pulsations and an increase in the size of the mass can be observed when the patient coughs or strains. Computed tomography and MRI are necessary to determine the appropriate combined transfacial and intracranial approach for surgical resection.[20]

Dermoid Cysts

Dermoid cysts occasionally occur in the neck, usually in the midline. Overall, fewer than 10% of all dermoid cysts occur in the head and neck. One-

fourth are found in the floor of the mouth, with the remainder in the periorbital region. These lesions are seen primarily in children and adolescents during the second decade of life, but they also may occur in infants. Cyst walls consist of squamous cell epithelium containing epidermal appendages. Periorbital dermoids are thought to originate from displacement of epidermal elements during the intramembranous growth phase of the nasal bones in the embryo. Because of the variability in clinical presentation and contiguous structure involvement, segregation of periorbital lesions into three distinct subgroups is helpful: (1) brow region dermoid cysts, (2) orbital region dermoid cysts, and (3) nasoglabellar dermoid cysts. Midline nasoglabellar cysts may have associated sinus tracts or fistulae (10 to 45%), which, in rare cases, may have intracranial extensions. In those patients with intracranial extension, the sinus tract traverses either the cribriform plate or foramen cecum and is attached to the dura, falx cerebri, or other intracranial structures.[21]

The treatment of choice is surgical excision after CT and MRI to evaluate intracranial extension.[22] All abnormal tissue must be removed to prevent recurrences. In patients with intracranial extension, a bicoronal flap is employed to facilitate removal and to prevent postoperative meningitis and abscess formation.

Other congenital midface cysts are those associated with the facial clefts, including the nasoethmoidal cleft cyst, nasolabial cyst, subalar cleft cyst, globulomaxillary cyst, cysts connected with cleft lip or palate, premaxillary cyst, nasopalatine cyst, foraminal incisor cyst, and Jacobson's organ cysts.

Choanal Atresia

Bilateral atresia of the choanae is the most frequently encountered congenital nasal anomaly (1 in 7,000 to 8,000 live births) and is a common cause of neonatal respiratory distress. Fifty to 60% of cases are unilateral. A recent analysis of 47 CT scans or histopathologic sections demonstrated 29% pure bony, 71% mixed bony and membranous, and no pure membranous atresia.[23] A female predisposition is seen in choanal atresia, and recent evidence points to an autosomal recessive mode of inheritance.

Failure of breakdown of the buccopharyngeal membrane on gestational day 45 is considered to be the cause of choanal atresia. Other theories include abnormalities in the migration of the cephalic neural crest following neural tube closure. Several craniofacial abnormalities, including skull base defects and systemic malformations, have been described in association with choanal atresia, including the CHARGE association (Table 46–1). The CHARGE association is a nonrandom pattern of congenital anomalies that has an estimated prevalence of 1 in 10,000 births. It occurs with choanal atresia (C = coloboma; H = congenital heart disease; A = atresia choanae; R = retarded growth and development; G = genital anomalies in males; E = ear abnormalities and deafness). Almost all patients have malformed pinnae and exhibit hypoplastic incudes, cochlear anomalies, and absent semicircular canals on CT scan (Figure 46–3). The majority demonstrate severe conductive or mixed hearing loss and vestibular dysfunction.[24] In addition, almost half exhibit facial nerve palsies. Laryngotracheal anomalies occur in one-third of cases. The combination of malformations in the CHARGE association suggests that this syndrome is a polytopic developmental field defect involving the neural tube and neural crest cells.[25]

Symptoms Respiratory distress at birth is the sine qua non of bilateral choanal atresia. In spite of vigorous attempts at respiration, effective air exchange does not occur until the neonate begins to cry, bypassing the nasal obstruction. Once the crying ceases, however, the neonate's mouth closes, and a pattern of cyclic obstruction gradually develops, resulting in increasing respiratory failure. Because neonates are obligatory nasal breathers, placement of an oral airway or McGovern nipple is lifesaving. By contrast, children with unilateral choanal atresia

TABLE 46–1. Congenital Anomalies Associated with Choanal Atresia

Branchial anomalies
CHARGE association
Humeroradial synostosis
Mandibular facial synostosis
Microcephaly
Micrognathia
Nasopharyngeal anomalies
Palatal defects

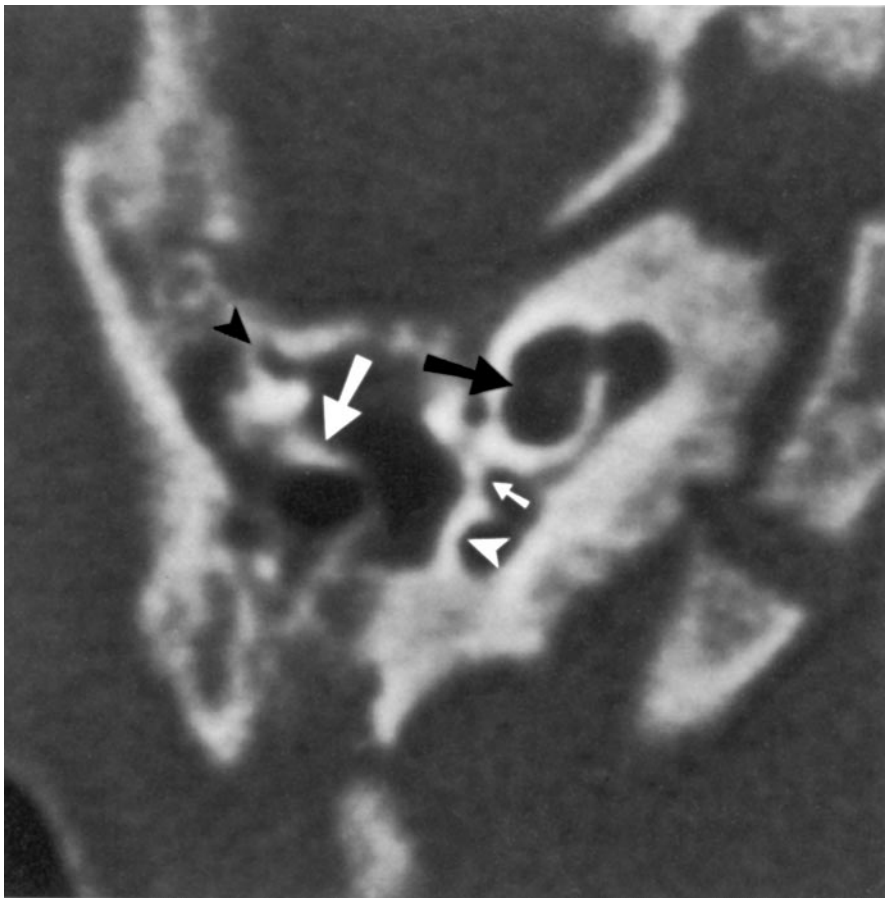


FIGURE 46–3. Axial 1 mm thick computed tomographic section through right mesotympanum in patient with CHARGE association. Note fixation of the malleus head to the anterior tympanic wall (*black arrowhead*), malformation and narrow rotation of the incus (*large white arrow*), and absence of stapes. Also note absence of the oval window (*small white arrow*) and presence of a round window niche with bony obliteration (*white arrowhead*). Tympanic sinus, pyramidal eminence, and stapedius muscle are absent. Interscalar septum between middle and apical turns of the cochlea is missing, as well as a normal cochlear modiolus (*black arrow*). Reproduced with permission from Dhooge I, Lemmerling M, Lagache M, et al. Otological manifestations of CHARGE association. Ann Otol Rhinol Laryngol 1998;107:935–41.

present later in life with unilateral rhinorrhea without respiratory distress.

Diagnosis The diagnosis of bilateral choanal atresia is confirmed by the inability to pass a No. 5 or 6 French feeding catheter at least 3 cm through the nose into the nasopharynx. In addition, direct observation with nasofiber-optic endoscopy and CT scan are essential to determine the type of obstruction. Specifically, the CT scan demonstrates the thickness of the atretic plate as well as its encroachment of the posterior aspect of the medial part of the maxilla.[26]

Treatment Multiple methods are available to repair choanal atresia. The appropriate method is determined by the age of the patient and whether the atresia is bilateral or unilateral. In all methods, the challenge is to provide adequate mucosal lining to the new choana and to prevent granulation tissue formation and subsequent stenosis. Treatment requires perforation of the atresia plate followed by stenting for 6 weeks with preservation of mucosal flaps.

Transnasal Approach. The most direct and simplest route for choanal atresia repair is transnasal, using a No. 2 Lempert or similar curette inserted 3 to 3.5 cm beyond the nares along the nasal floor and exerting firm pressure against the bony atresia plate until it is perforated. The anterior nasal mucous membrane is sacrificed after satisfactory enlargement of the choana to permit placement of a 3.5 mm Silastic endotracheal tube as a stent. Posterior membranous flaps are developed in a stellate fashion to cover the denuded bone (Figures 46–4 and 46–5). Care must be taken to prevent injury to the roof of the choana and the skull base because reported complications of this approach have included meningitis, cerebrospinal fluid leaks, brain injuries, Gradenigo's syndrome, and cervical vertebral subluxation. Studies indicate that

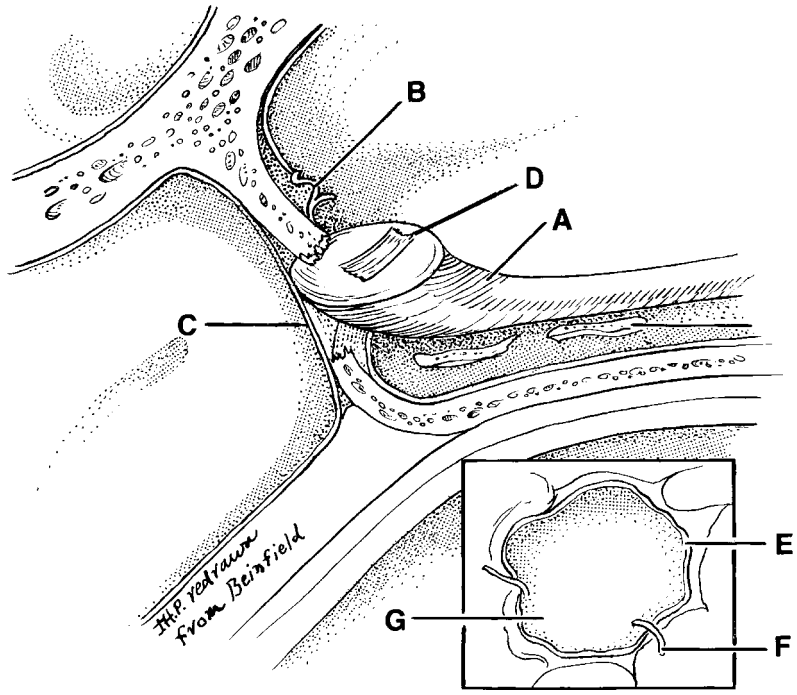


FIGURE 46–4. *A*, curette; *B*, shreds of nasal mucous membrane; *C*, intact nasopharyngeal mucous membrane; *D*, spicule of bone; *E*, bony atresia almost completely removed; *F*, shreds of mucous membrane; and *G*, intact pharyngeal mucous membrane.

transnasal curettage has a higher incidence of restenosis than newer techniques and requires additional revisions or dilatations. Other methods include the use of the carbon dioxide laser, endoscopic endonasal approach, and neodymium: yttrium-aluminum-garnet (Nd:YAG) laser technique. Currently, an endoscopic approach with a 2.5 or 4 mm telescope and powered instrumentation using attachable burs and blades with continuous suction is preferred, especially for neonates (Figure 46–6).[27] After effective removal and stenting and the use of topical corticosteroids to diminish granulation tissue formation, stenosis may still occur, and dilatation may be necessary. If this fails, a transpalatal approach will be required. A transseptal approach, which is reserved for older children with unilateral atresia, offers better exposure of the posterior part of the vomer without risk of palatal injury and impairment of growth. This approach can be modified either via sublabial extension or by an external rhinoplasty approach to improve visualization.

Transpalatal Approach. The transpalatal approach provides superior visualization of membranous or bony atresia and may be useful for both bilateral and unilateral atresia. Although impaired palatal growth is a potential problem in the neonate, the procedure can be safely used in patients older than 5 years who are treatment failures if one takes great care to avoid injury to the posterior palatine canals.

The procedure is carried out with the patient in the supine position and the neck extended. After infiltration of local anesthesia, a horseshoe-type incision is developed several millimeters posterior to the alveolar ridge, and a mucoperiosteal flap is then raised based on the greater palatine vessels. Once the posterior edge of the hard palate is reached, the mucosa is divided, providing access to the atretic plate. An attempt is made to maintain mucosal flaps on both sides of the atretic plate. The atretic plate, posterior part of the vomer, and medial aspect of the posterior part of the maxilla are removed with a diamond bur, otologic curette, or rongeur (Figure 46–7). After removal of bone, mucosal flaps are transposed to cover the superior and inner surfaces of the choana, with stents placed as previously described. The palatal flap is reapproximated anteriorly with absorbable sutures. If stenosis occurs after surgical repair, patients generally slowly develop progressive respiratory difficulties. Revision surgery and/or dilatation may be necessary.

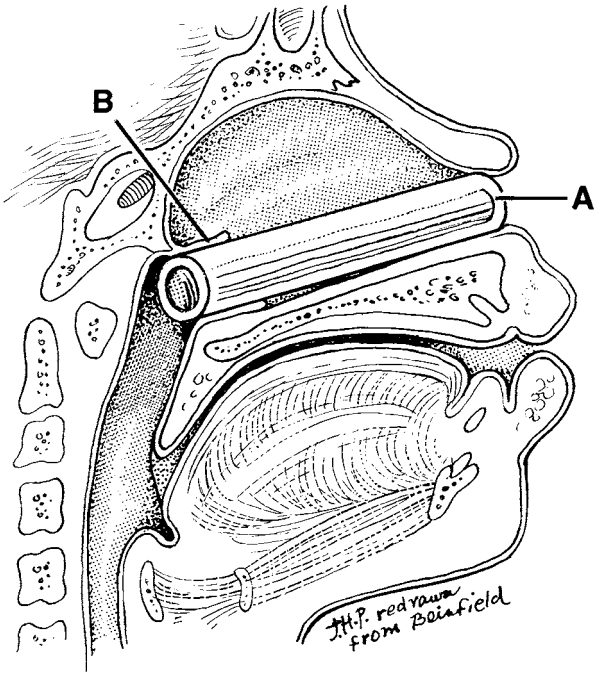


FIGURE 46–5. The tube in place. *A*, the anterior end of the tube lying just within the nares. *B*, Pharyngeal mucous membrane covering the raw bony surface.

CONGENITAL DISORDERS OF THE NASOPHARYNX AND OROPHARYNX

THORNWALDT'S CYST AND RATHKE'S POUCH CYST

If the pharyngeal segments of the primitive notochord remain connected to the endoderm in the nasopharynx, a bursa or embryonic pouch occurs. In approximately 3% of individuals, this invaginated connection persists, and the resulting sac and canal, located in the posterior midline of the nasopharynx, extends posteriorly and cephalad toward the occipital bone. If the bursa is occluded by inflammation, a cyst known as Thornwaldt's cyst will develop, and if the cyst becomes infected, an abscess will result. Anterior to the invagination and above the bursa is a small pharyngeal hypophysis developed from Rathke's pouch, which sometimes persists as the craniopharyngeal canal running from the sella turcica through the body of the sphenoid.

Most Thornwaldt's cysts appear clinically in the second and third decades of life, with males and females affected equally. Symptoms include intermittent or persistent postnasal discharge of tenacious mucus or purulent material, associated with odynophagia, halitosis, unpleasant taste, and, occasionally, a dull occipital headache exacerbated by head movement with associated stiffness of the posterior cervical muscles. Nasopharyngoscopy reveals a smooth, submucosal, 1 to 2 cm midline cystic mass, superior to the adenoidal pad. Frequently, a central dimple or fistula is identified. A CT scan demonstrates a soft tissue mass located high on the posterior nasopharyngeal wall with sharp borders. Magnetic resonance studies of the nasopharynx reveal high signal intensity on T_1-weighted, T_2-weighted, and fluid-attenuated inversion recovery images.[28] Microscopically, Thornwaldt's cyst is lined by respiratory epithelium with minimal amounts of lymphoid tissue in the wall. Treatment requires either excision to the periosteum or wide marsupialization.

Rathke's pouch cysts, on the other hand, are basically cysts lined by columnar or cuboidal, ciliated epithelium that may become secondarily infected or rupture intracranially and are most commonly associated with headache followed by galactorrhea, visual field loss, and hypopituitarism. Treatment is transsphenoidal drainage of the cyst with biopsy of the wall. If no demarcation between normal and involved tissue is evident at surgery, the cyst and residual pituitary are radically removed.[29] Radiation therapy is not indicated.

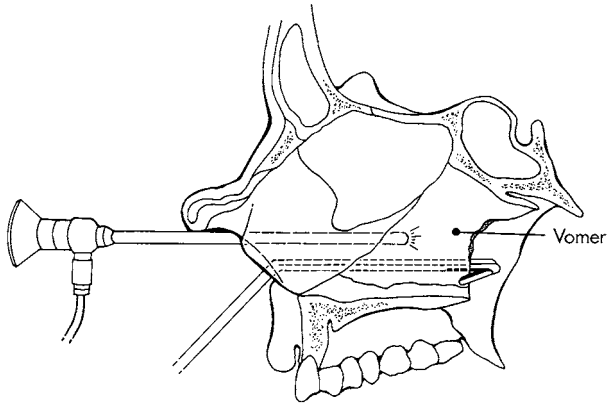


FIGURE 46–6. An endoscopic approach to choanal atresia repair with a backbiting instrument placed in the nasal cavity to resect a portion of the vomer. Reproduced with permission from Park AH et al.[27]

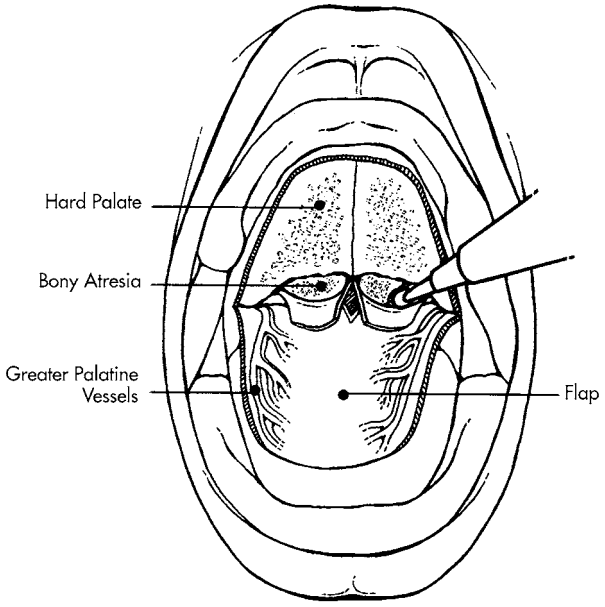


FIGURE 46–7. A transpalatal approach using a drill to remove the bony atretic plate; mucosa over the atretic plate has been preserved for reconstruction of the neochoana. Reproduced with permission from Park AH et al.[27]

Chordomas

Chordomas are rare malignant neoplasms arising from primitive notochordal remnants, primarily in the fifth or sixth decade of life. Fifty percent occur in the spheno-occipital area, with a mean patient age of 40 to 45 years. Males and females are approximately equally affected. Presenting signs and symptoms include an expanding nasopharyngeal mass, frontal headaches, cranial nerve palsies (sixth nerve involvement is seen in 60% of patients followed equally by ninth and tenth nerve involvement), and pituitary abnormalities.

Children under 5 years of age have a wider range of presenting symptoms, a greater prevalence of atypical histologic findings with aggressive behavior, and a higher incidence of metastasis to lung and lymph nodes. These tumors show no chondroid component compared to a 17.1% incidence in older patients.[30] Magnetic resonance imaging with gadolinium is essential to evaluate the extent of the tumor, and it may be helpful in distinguishing classic chordoma from the chondroid variant. Computed tomographic scanning frequently reveals extensive bone destruction at the skull base and is used to plan surgical resectability. Treatment is surgical resection via a skull base approach combined with postoperative radiation therapy or proton-photon therapy.[31] The prognosis in patients with chordomas is unfavorable, with median 5-year survival rates at 50% for patients with surgical resection only. The survival rates for combined surgery and radiation therapy are statistically significantly longer than for surgical treatment alone. The overall prognosis is directly related to the histologic pattern of the tumor: atypical chordomas carry a poor prognosis. There appears to be no significant difference for prognosis in patients older than 5 years of age with a classic or chondroid tumor.[30] For metastatic chordoma, multiagent systemic chemotherapy (ifosfamide and doxorubicin) has been efficacious, whereas systemic chemotherapy with intrathecal or intraventricular hydrocortisone, cytarabine, and methotrexate has been effective in controlling central nervous system involvement with chordoma.[32]

Craniopharyngiomas

Craniopharyngiomas arise from Rathke's pouch and are located in the sellar, parasellar, and third ventricular regions. They are composed of well-differentiated epithelial elements including cysts and ameloblasts as well as bone. Rarely, this lesion is first seen in the nasopharynx but most commonly occurs intracranially above the sella turcica. It accounts for approximately 10 to 15% of all childhood and adolescent intracranial neoplasms. Clinical manifestations include visual field defects, sudden blindness, extraocular motor paralysis, and hypopituitarism. The median age at presentation in children is 7.5 years, with an approximately equal female-to-male ratio. Magnetic resonance imaging with gadolinium or gadodiamide injection is mandatory for pretreatment assessment. Management of craniopharyngiomas has been controversial. Currently, surgical resection followed by radiation therapy achieves long-term control with low morbidity for tumors smaller than 5 cm. Ten-year actuarial overall survival rates are 90%. Despite major microsurgical advances, however, total removal of these tumors is associated with a high risk of death, endocrinologic complications, and behavioral dysfunction. Craniopharyngiomas located within an enlarged sella that are not adherent to parasella structures are removed

through a transsphenoidal approach with a lower incidence of recurrence.[33] Primary linear accelerated stereotactic radiation therapy has also proven valuable in the initial management of this tumor. A long-term (greater than 5 years) follow-up is essential to assess tumor control.[34,35]

Teratomas

Teratomas are true neoplasms that contain tissues foreign to the site in which they arise. The haphazard arrangement of tissue with asynchronous maturation is believed to be attributable to escape from the controlling influence of the primitive streak notochord or adjacent structures. Teratomas grow aggressively, and in the head and neck, they most commonly occur in the cervical area, followed by the nasopharynx. Nasopharyngeal teratomas occur with a female-to-male ratio of 6 to 1. Overall, teratomas of the head and neck comprise approximately 2 to 9% of all teratomas.

Teratomas of the nasopharynx typically arise on the lateral or superior wall. Four basic types are recognized: (1) dermoid cyst, the most common form, composed of ectoderm and mesoderm arising as an epithelium-lined cavity with variable numbers of skin appendages; (2) teratoid cyst, derived from all three germ layers but poorly differentiated; (3) true teratoma, composed of ectoderm, mesoderm, and endoderm, with specific tissue and organ differentiation; and (4) epignathus, in which well-developed fetal parts are recognizable; this type of teratoma arises from the soft or hard palate and is frequently incompatible with life.

Clinically, dermoids are more common than true teratomas, and in the nasopharynx, the dermoid cyst is the most common developmental anomaly found. Known as "hairy polyps," these lesions appear at birth as a pedunculated mass filling the nasopharynx, often with oropharyngeal extension. Patients with teratomas have nasal obstruction, dysphagia, and copious secretions. Infants with nasopharyngeal teratomas have a higher incidence of preterm birth, neonatal airway distress, and associated congenital anomalies along with polyhydramnios than do infants with dermoids.[36] Computed tomography and MRI are critical to define the extent of the neoplasm and to exclude either a nasoencephalomeningocele or intracranial extension of a sphenoid-based teratoma through the craniopharyngeal canal. Most cervical teratomas occurring in the neonate are benign. The tumor can surround or encroach on the airway, thereby causing progressive dysphagia and airway obstruction.

Surgical removal must be carefully planned to ensure a controlled airway throughout the intraoperative and postoperative periods. Tracheostomy is generally not necessary if orotracheal intubation is combined with a transoral removal of pedunculated nasopharyngeal teratomas or with early excision of cervical teratomas. Operative and postoperative bleeding in patients with nasopharyngeal teratomas is usually only slight because these tumors are poorly vascularized, and if removal is complete, recurrences are rare. More sessile tumors arising in the nasopharynx require a transpalatal approach. Malignant metastasizing cervical teratoma is extremely rare.

CONGENITAL DISORDERS OF THE ORAL CAVITY AND LIP

Numerous congenital defects of the oral cavity and related structures may occur. Common major anomalies, such as cleft lip and palate, are described in detail, whereas less common deformities are mentioned for completeness.

Congenital lesions of the tongue include aglossia, microglossia, and macroglossia (secondary to lymphangioma or hemangiomas or as commonly found in Down syndrome). Ankyloglossia owing to varying degrees of underdevelopment of the lingual frenulum is not uncommon. Indications for frenuloplasty with Z-plasty are notching of the protruding tongue tip, inability to contact the maxillary alveolar ridge, and/or restriction of protrusion beyond the mandibular alveolar ridge. Other glossal anomalies include dermoid cysts, hamartomas, lingual thyroid, fissured (scrotal) tongue, median rhomboid "glossitis," and enteric duplication cysts.

Nonodontogenic cysts derived from epithelial remnants trapped in embryonic fusion lines during the developmental stage include midline maxillary cysts (median alveolar, nasal palatal, median palatal) and lateral maxillary cysts (globomaxillary, nasolabial). Treatment varies from simple incision and drainage to removal of the cysts from the palatal side, with or without adjacent teeth. Noncleft lip congenital anomalies encountered include microstomia, congenital pits, and "double lip."

Finally, congenital epulis or gingival granular cell tumor arises exclusively from the alveolar ridge and is a rare lesion of unknown origin found in newborn female infants.

Cleft Lip and Palate

Cleft lip and palate are the most common congenital malformations of the head and neck, occurring approximately once in every 700 births. The risk of additional siblings having a defect increases when an older sibling is affected. Both cleft lip with or without cleft palate and isolated cleft palate may be further classified into those classes associated with (syndromic) or without (nonsyndromic) another recognized malformation. The cause of syndromic clefting may be single-gene transmission, chromosomal aberrations, teratogenicity, or environmental factors. Recent evidence points to multigenic inheritance with allelic variation at different loci. However, some pedigrees are monogenic, with either autosomal dominance or recessive inheritance. Exposure to teratogens such as ethanol, retinoids, or folate antagonists during the first trimester is associated with an increased risk of syndromic cleft lip or cleft palate.[37] In addition, a dose-response relationship between maternal smoking and an increased risk of having a child with cleft lip or cleft palate has been demonstrated.[38] Associated head and neck anomalies such as maxillary or malar hypoplasia, abnormal pinnae or atresia, facial nerve paresis or paralysis, and mandibular dysmorphism or abnormal excursion may be identified. Nonsyndromic clefting is associated with no obvious first or second arch anomalies or systemic organ malformation. Multifactorial inheritance is the cause of these clefts, and calculated risk rates are essential to provide genetic counseling to families.

Anatomy The lips are a movable muscular sphincter composed primarily of orbicularis oris muscle with motor supply from the seventh cranial nerve. The lips are covered by skin on the outer surface and mucosa on the inner. The vermillion or lip edge forms a junction with the skin (white line) that creates a gentle arching in the upper lip or "Cupid's bow." The cephalic extension from the handle of the Cupid's bow encloses a small depressed area at the base of the columella called the philtrum. These anatomic landmarks, plus the protrusion of the vermillion below the Cupid's bow or tubercle, are used for orientation in cleft lip repair.

The palate is composed of a hard anterior portion and a soft muscular posterior portion. The anterior primary palate is formed by the alveolar ridge and four incisor teeth plus the triangular premaxilla. The remaining hard palate is formed by the palatine processes of the maxillary and palatine bones, which, with the soft palate, comprise the secondary palate. The soft palate is a muscular curtain formed by the tensor veli palatini muscle innervated by the fifth cranial nerve and the levator veli palatini muscle, musculae uvulae, and glossopalatine and palatopharyngeus muscles innervated by the pharyngeal plexus.

Embryogenesis Clefts of the lip, alveolar ridge, or palate are midfacial soft tissue and skeletal fusion abnormalities. In the normally developing fetus, the median nasal process fuses with the maxillary processes to form the upper lip, premaxilla, and corresponding segments of the alveolar process. Together, the premaxilla and the alveolar ridge form the primary palate. The palatal processes fuse with each other to form the secondary palate, along with the nasal septum and premaxilla in the vicinity of the incisive foramen between gestational weeks 8 and 10. Palatal growth proceeds posteriorly toward the uvula. Failure of ingrowth of mesodermal tissue at this point results in a lack of cohesion of the palatal segments, causing a cleft palate, which may be seen in conjunction with clefts of the lip and/or alveolar process or alone. Interruption in the migration of mesodermal tissue during the first 2 months of embryonic life results in cleft lip deformities. Ultimately, a lack of fusion of the median nasal processes with the maxillary processes causes a cleft of the upper lip, premaxilla, and alveolar process. The incidence of cleft palate increases in the presence of a cleft lip, which occurs in an earlier stage than cleft palate. Moreover, the tongue is positioned higher in the oral cavity in cleft lip, further interfering with palatal fusion and leading to greater palatal clefting.

Management The head and neck surgeon is part of a cleft palate or craniofacial team that includes pediatrics, ophthalmology, general plastic surgery, neurosurgery, audiology, speech and language pathology, orthodontics, prosthodontics, oral surgery, genetics, dentistry, cytology, psychiatry, social

work, and nursing. The goals of this team are to restore normal anatomy and physiology, with an emphasis on muscular reconstruction of the lip and/or palate to allow normal facial development and to minimize growth disturbance.

The initial priority for infants with clefts is to establish adequate feeding and nutrition. Infants with a unilateral or bilateral cleft lip and alveolar ridge feed generally well by either breast or bottle. Infants with bilateral cleft lip, alveolar ridge, and palate have significant feeding problems and require modified nipples, with feeding in the upright position to minimize nasal regurgitation.

Although the timing of surgical repair is controversial, most experts perform cleft lip repair at 3 months of life and cleft palate repair at 12 months to establish a competent velopharyngeal sphincter. Two schools of thought have evolved: one advocating early closure of the lip and palate, a procedure imparting a high priority to early speech development, and the other recommending delayed closure of the hard palate, thus according a high priority to maxillary growth.[39] Surgical treatment is dictated by the anatomic defect, and a classification scheme is helpful in planning surgical repair: (1) cleft lip, unilateral or bilateral; (2) cleft palate, either with unilateral or bilateral cleft lip; and (3) isolated cleft palate.

Unilateral Cleft Lip. In the basic complete defect, the floor of the nose communicates with the oral cavity, and the alveolar defect passes through the developing dentition. One sees significant nasal deformity, including columellar displacement, nasal dome deformity, and alar flattening and retrodisplacement. The goals of surgical repair are restoration of orbicularis muscle function, alar base, and columellar height and creation of symmetry of philtral columns, tip height, Cupid's bow, and vermillion. This can be accomplished with a Millard rotation advancement procedure in concert with either preoperative orthodontic appliances or lip adhesion.

Bilateral Cleft Lip. The floor of the nose is absent bilaterally, and the nasal and oral cavities communicate freely. The central alveolar arch is displaced forward and superiorly, whereas the nasal tip is widened and the columella is short. The goals of bilateral cleft lip repair are the same as for unilateral defects. Bilateral lip adhesion, if indicated, is performed when the patient is 2 to 4 weeks old, with definitive lip repair followed at 4 to 6 months of age. If no lip adhesion is necessary to align the underlying maxillary segments before definitive lip repair, bilateral lip repair is performed at 3 months.

Cleft Palate. The basic defect, absence of the nasal floor, may be complete or incomplete, with or without associated cleft lip. The goals of cleft palate repair are to close the defect without inhibiting maxillary arch growth, to achieve adequate velopharyngeal function, and to normalize oronasal resonance and speech patterns.[40]

Although the exact timing of the repair is controversial, the procedure is performed when the patient is 10 to 18 months old and may require preoperative orthopedic devices to move the premaxilla posteriorly and to expand the lateral maxillary segments, facilitating surgical closure of the lip. Of the many techniques described for bilateral cleft palate repair, the two-flap palatoplasty by Bardach is commonly employed: (1) complete two-layer closure (oral and nasal sides) and (2) dissection, redirection, and suturing of the soft palatal musculature.

Secondary problems of cleft lip and palate that require subsequent treatment include correction of nasal airway impairment, cosmetic nasal defects, velopharyngeal insufficiency with associated hypernasality, and recurrent acute and chronic otitis media with effusion.

Pierre Robin Sequence. "Pierre Robin" is one of the most recognized diagnostic eponyms. It is a poorly understood, nonspecific grouping of malformations, however, which as yet have no prognostic significance. The classic PRS of glossoptosis, micrognathia, and cleft palate (all three findings in 50% of cases) may occur as an isolated, nonsyndromic congenital disorder or as part of a larger anomaly that may include facial dysmorphism, cardiac defects, mental retardation, and/or musculoskeletal anomalies.[41] Möbius' syndrome or nonprogressive congenital facial diplegia is such an example of PRS, with concomitant brachial and thoracic muscle aplasia along with cranial neuropathies (six and seven). Pierre Robin sequence has no sex predilection, and it may be secondary to an intrauterine insult during the fourth month of gestation or owing to hereditary causes. Syndromic PRS is often associated with ocular anomalies (detached

retina or glaucoma) or aural abnormalities (chronic otitis media and low-set pinna). Heredity appears to be a factor in isolated nonsyndromic PRS, and usually a complete U-shaped cleft palate is the primary determinator of the associated triad.[42]

Infants with PRS are at increased risk of airway obstruction and resulting hypoxemia, cor pulmonale, failure to thrive, and cerebral anoxia. Obstructive sleep apnea and snoring are common. Aspiration is a primary cause of death. Aerophagia associated with vomiting frequently results in aspiration. When compared with normal infants, infants with PRS exhibit a shorter tongue and mandibular length, with a narrow airway and a posteroinferior position of the hyoid bone resulting in airway compromise. Patients are initially given a trial of prone positional management with high-calorie gavage feeding. When continued respiratory distress and failure to thrive develop, a modified glossopexy, attaching the tongue at the mandibular alveolar ridge and the lower lip, is performed. The genioglossus is also released to lengthen the tongue. Bedside polysomnography may help to identify those infants without clinically severe airway obstruction who may benefit from a modified tongue lip adhesion. Tracheostomy is reserved for those patients following tongue lip adhesion who have continued failure to thrive and respiratory embarrassment. Patients with syndromic PRS have a higher rate of tracheostomy and gastrostomy tube placement and exhibit significantly lower Apgar scores and longer hospital stays.[43]

CONGENITAL DISORDERS OF THE SALIVARY GLANDS

Congenital anomalies of the salivary glands are uncommon. Rarely, the parotid gland may be absent in the first and second branchial arch syndromes. Congenital sialoadenitis with cystic dilatation of the salivary ducts can occur, resulting in single or multiple cystic spaces. Sialography in the infant under sedation may help to demonstrate ductal ectasia. If obstruction and stasis develop, recurrent infections may result, necessitating parotidectomy with seventh nerve preservation. Heterotopic salivary gland tissue may also appear in or adjacent to the salivary glands, developing from a defect in the growth of the ductal system or in relation to first branchial cleft anomalies. Usually, these congenital lesions present as a discharging sinus in the base of the neck.[44]

FIRST BRANCHIAL CLEFT ANOMALIES

First branchial cleft anomalies have been classified into two groups, types 1 and 2. Type 1 anomalies are first cleft ectodermal defects or duplication anomalies of the membranous external auditory canal. Type 1 cysts are found anterior and inferior to the pinnae in association with the parotid gland. The fistulous tract extends superiorly along the facial nerve and parallel to the external auditory canal, where it terminates. Type 2 branchial cysts are primarily first cleft and arch defects (ectodermal and mesoderm) with duplication of the external canal and pinnae. Type 2 cysts are found below the angle of the mandible along the sternocleidomastoid muscle. The fistulous tract, like in type 1 anomalies, is variable and may run lateral to, medial to, or between the main branches of the facial nerve, ultimately ending in the external auditory canal.

Clinically, first branchial cleft cysts are associated with repeated infection and may require incision and drainage of an apparent neck abscess, which is typically located above a horizontal plane passing through the hyoid bone. Occasionally, purulent otorrhea may occur in the absence of demonstrable middle ear disease. A third common clinical presentation is preauricular swelling in the parotid area. Microscopically, the cyst is lined with hair follicles, apocrine or sweat glands, and sebaceous glands. Treatment is surgical excision of the cyst, sinus tract, and skin surrounding the sinus tract orifice. Because of the intimate relationship between the seventh nerve and the sinus tract, it is critical to expose the seventh nerve as in a parotidectomy, using the bony landmarks (styloid process, mastoid tip, and tympanomastoid suture line) to help identify the nerve, instead of the tragal pointer.[45] A nerve integrity monitoring system may prove exceedingly beneficial in these difficult cases.

Hemangiomas and lymphangiomas are congenital lesions that may involve the salivary glands and are discussed earlier in this chapter.

PRIMARY CILIARY DYSKINESIA (KARTAGENER'S SYNDROME)

Primary ciliary dyskinesia syndrome is an inherited disorder characterized by recurrent upper and lower respiratory tract infections secondary to abnormal ciliary structure and function. Also termed immotile

cilia syndrome, its most recognizable form is in Kartagener's syndrome (bronchiectasis, chronic sinusitis, situs inversus, and sterility). It has also been recognized in Usher's syndrome type I, congenital heart disease, congenital esophageal dysfunction, and cutaneous abnormalities including folliculitis, nummular eczema, and pyoderma gangrenosum.

This disorder is characterized by defects in axonemal dynein complexes probably secondary to mutations in the genes encoding cytoplasmic heavy chains of the outer arms of the axonemes. Dyneins are large multisubunit adenosine triphosphatases that interact with microtubules in cilia to generate force. The resulting ultrastructural defect causes impaired mucociliary clearance throughout the pulmonary and sinonasal passages. Early recognition and treatment of individuals with these disorders could lead to reduction in irreversible sinus and pulmonary complications with improved survival.[46,47]

REFERENCES

1. Mukherji SK, Fatterpekar G, Castillo M, et al. Imaging of congenital anomalies of the branchial apparatus. Neuroimaging Clin North Am 2000;10:75–93.
2. Jacobsson C, Granstrom G. Clinical appearance of spontaneous and induced first and second branchial arch syndromes. Scand J Plast Reconstr Surg Hand Surg 1997;31:125–36.
3. Srivastava D. HAND proteins: molecular mediators of cardiac development and congenital heart disease. Trends Cardiovasc Medq 1999;9:11–8.
4. Edmonds JL, Girord DA, Woodroof JM, Bruegger DE. Third branchial anomalies: avoiding recurrences. Arch Otolaryngol Head Neck Surg 1997; 123:438–41.
5. Nicollas R, Ducroz V, Garabedian EN, Trigilia JM. Fourth branchial pouch anomalies: a study of six cases and review of the literature. Int J Pediatr Otorhinolaryngol 1998;44:5–10.
6. Thompson LD, Heffner DK. The clinical importance of cystic squamous cell carcinomas in the neck: a study of 136 cases. Cancer 1998;82:944–56.
7. Hong KH, Moon WS, Chung GH. Radiological appearance of primary branchial cleft cyst carcinoma. J Laryngol Otol 1999;113:1031–3.
8. Koeller KK, Alamo L, Adair CF, Smirniotopoulos JG. Congenital cystic masses of the neck: radiologic-pathologic correlation. Radiographics 1999;19: 121–46.
9. Gallagher PG, Mahoney MJ, Gosche JR. Cystic hygroma in the fetus and newborn. Semin Perinatol 1999;23:341–56.
10. Batsakis JG, El-Naggar AK, Luna MA. Thyroid gland ectopias. Ann Otol Rhinol Laryngol 1996;105: 996–1000.
11. Winslow CP, Weisberger EC. Lingual thyroid and neoplastic change: a review of the literature and description of a case. Otolaryngol Head Neck Surg 1997;117:S100–2.
12. Greinwald JH Jr, Leichtman LG, Simko EJ. Hereditary thyroglossal duct cysts. Arch Otolaryngol Head Neck Surg 1996;122:1094–6.
13. Josephson GD, Spencer WR, Josephson JS. Thyroglossal duct cyst: the New York Eye and Ear Infirmary experience and a literature review. Ear Nose Throat J 1998;77:646–51.
14. Organ GM, Organ CH. Thyroid gland and surgery of the thyroglossal duct: exercise in applied embryology. World J Surg 2000;24:886–90.
15. Bardales RH, Shurland MJ, Korourian S, et al. Cytologic findings in thyroglossal duct carcinoma. Am J Clin Pathol 1996;106:615–9.
16. Nguyen Q, de Tar M, Wells W, Crockett D. Cervical thymic cyst: case reports and review of the literature. Laryngoscope 1996;106:247–52.
17. Kacker A, April M, Markentel CB, Breuer F. Ectopic thymus presenting as a solid submandibular neck mass in an infant: case report and review of literature. Int J Pediatr Otorhinolaryngol 1999;49:241–5.
18. Sobol SE, Tewfik TL, Ortenberg J. Otolaryngologic manifestations of neurofibromatosis in children. J Otolaryngol 1997;26:13–9.
19. Koral K, Geffner ME, Curran JG. Transsphenoidal and sphenoethmoidal encephalocoeles: report of two cases and review of the literature. Australas Radiol 2000;44:220–4.
20. Peter JC, Fieggen G. Congenital malformations of the brain—a neurosurgical prospective at the close of the 20th century. Childs Nerv Syst 1999;15:635–45.
21. Hong SW. Deep frontotemporal dermoid cyst presenting as a discharging sinus: a case report and review of literature. Br J Plast Surg 1998;51:255–7.
22. Lowe LH, Booth TN, Joglar JM, Rollins NK. Midface anomalies in children. Radiographics 2000;20:907–22.
23. Brown OE, Pownell P, Manning SC. Choanal atresia: a new anatomic classification and clinical management application. Laryngoscope 1996;106:97–101.

24. Wiener-Vacher SR, Amanou L, Denise P, et al. Vestibular function in children with the CHARGE association. Arch Otolaryngol Head Neck Surg 1999;125:342–7.
25. Tellier AL, Cormier-Daire V, Abadie V, et al. CHARGE syndrome: report of 47 cases and review. Am J Med Genet 1998;76:402–9.
26. Jones JE, Young E, Heier L. Congenital bony nasal cavity deformities. Am J Rhinol 1998;12:81–6.
27. Park AH, Brockenbrough J, Stankiewicz J. Endoscopic versus traditional approaches to choanal atresia. Otolaryngol Clin North Am 2000;33:77–90.
28. Ikushima I, Korogi Y, Makita O, et al. MR imaging of Thornwaldt's cysts. AJR Am J Roentgenol 1999; 172:1663–5.
29. Roncaroli F, Bacci A, Frank G, Calbucci F. Granulomatous hypophysitis caused by a ruptured intrasellar Rathke's cleft cyst: report of a case and review of the literature. Neurosurgery 1998;43:146–9.
30. Borba LA, Al-Mefty O, Mrak RE, Suen J. Cranial chordomas in children and adolescents. J Neurosurg 1996;84:584–91.
31. Munzenrider JE, Liebsch NJ. Proton therapy for tumors of the skull base. Strahlenther Onkol 1999;Suppl 2:57–63.
32. Scimeca PG, James-Herry AG, Black KS, et al. Chemotherapeutic treatment of malignant chordoma in children. J Pediatr Hematol Oncol 1996; 18:237–40.
33. Norris JS, Pavaresh M, Afshar F. Primary transsphenoidal microsurgery in treatment of craniopharyngiomas. Br J Neurosurg 1998;12:305–12.
34. Habrand JL, Ganry O, Couanet D, et al. The role of radiation therapy in the management of craniopharyngioma: a 25-year experience and review of the literature. Int J Radiat Oncol Biol Phys 1999; 44:255–63.
35. Sims E, Doughty D, Macauley E, et al. Stereotactically delivered cranial radiation therapy: a ten-year experience of linac-based radiosurgery in the UK. Clin Oncol (R Coll Radiol) 1999;11:303–20.
36. Coppit GL, Perkins JA, Manning SC. Nasopharyngeal teratomas and dermoids: a review of the literature and case series. Int J Pediatr Otorhinolaryngol 2000;30:219–27.
37. Houdayer C, Bahuau M. Orofacial cleft defects: inference from nature and nurture. Ann Genet 1998;4:89–117.
38. Chung KC, Kowalski CP, Kim HM, Buchman SR. Maternal cigarette smoking during pregnancy and the risk of having a child with cleft lip/palate. Plast Reconstr Surg 2000;105:45–91.
39. Molsted K. Treatment outcome in cleft lip and palate issues and perspectives. Crit Rev Oral Biol Med 1999;10:225–39.
40. LaRossa D. The state of the art in cleft palate surgery. Cleft Palate Craniofac J 2000;37:225–8.
41. Prows CA, Bender PL. Beyond Pierre Robin sequence. Neonatal Netw 1999;18:13–9.
42. Marques IL, Babieri MA, Bettiol H. Etiopathogenesis of isolated Robin sequence. Cleft Palate Craniofac J 1998;35:517–25.
43. Cruz MJ, Kerschner JE, Beste DJ, Conley SF. Pierre Robin sequence: secondary respiratory difficulties and intrinsic feeding abnormalities. Laryngoscope 1996;109:1632–6.
44. Lassaletta-Atienza L, Lopez-Rios F, Martin G, et al. Salivary gland heterotopia in the lower neck: a report of five cases. Int J Pediatr Otorhinolaryngol 1998; 43:153–61.
45. Triglia JM, Nicollas R, Ducroz V, et al. First branchial cleft anomalies: a study of 39 cases and a review of the literature. Arch Otolaryngol Head Neck Surg 1998;124:291–5.
46. Milisav I. Dynein and dynein-related genes. Cell Motil Cytoskeleton 1998;39:261–72.
47. Teknos TN, Metson R, Chasse T, et al. New developments in the diagnosis of Kartagener's syndrome. Otolaryngol Head Neck Surg 1997;116:68–74.

CHAPTER 47

Anatomy and Physiology of the Larynx

Clarence T. Sasaki, MD, Young-Ho Kim, MD, PhD

The human larynx is a complex organ that functions as a sphincter at the junction of the digestive tract and respiratory tract and participates in the diverse physiologic aspects of airway protection, respiration, and phonation. In his classic phylogenic observations, Negus allowed prioritization of the three functions of the larynx.[1] In order of their priority, these are (1) protection of the lower airway, (2) respiration, and (3) phonation. To perform these roles, the internal and external structures of the larynx interact under precise neural control, producing in humans the most complex of laryngeal functions. Thus, the anatomy of the larynx reflects the specialization required by these multiple roles and is best understood in relation to its diverse physiologic behaviors.

ANATOMY

Laryngeal Cartilages

Thyroid Cartilage The thyroid cartilage is a shield-shaped structure that serves to protect the internal anatomy of the larynx.[2–7] It is the largest cartilage of the larynx and is composed of two wings, the alae or laminae. The alae are fused in the midline and open posteriorly (Figure 47–1). In the male, the alae fuse at about 90 degrees, making a laryngeal prominence or Adam's apple. In the female, this prominence is absent owing to the more oblique fusion angle of 120 degrees. Superiorly, the fusion of the alae is deficient, accounting for the thyroid notch. Posteriorly, each ala has a superior and inferior horn or cornu. The inferior cornu articulates with a facet on the cricoid cartilage to form the cricothyroid joint. This is a synovial joint that allows rotation of the cricoid cartilage. This rotation varies the tension placed on the vocal folds. The superior cornu attaches to the greater cornu of the hyoid bone by way of the lateral thyrohyoid ligament. This ligament sometimes contains small triticeal cartilages. The two lateral thyrohyoid ligaments, along with the median thyrohyoid ligament, are condensations of the thyrohyoid membrane; these structures attach the hyoid bone to the thyroid cartilage. At the attachment of the superior cornu to the alae of the thyroid, a protuberance called the superior tubercle is found. About 1 cm anterior and superior to this tubercle, the superior laryngeal artery and the internal branch of the superior laryngeal nerve and associated lymphatics pierce the membrane to supply the supraglottic larynx. At this point, transcutaneous anesthesia of the internal branch can be performed. Running obliquely from the superior tubercle to the inferior tubercle (along the inferior margin of the thyroid cartilage) is a ridge called the oblique line, which serves as the attachment point for the thyrohyoid, sternothyroid, and inferior constrictor muscles.

The relationship of the surface anatomy to internal laryngeal anatomy merits consideration. Most important is the level of the true cords in relation to the thyroid cartilage. An understanding of this relationship is crucial to performing supraglottic laryngectomy and phonosurgery (thyroplasty type I). In this regard, the midline vertical distance from the thyroid notch to the inferior border of the thyroid cartilage ranges from 20 to 47 mm in men and 15.5 to 38 mm in women.[2,3] The anterior commissure is found at the midpoint between these landmarks. The posterior extent of the cords is anterior to the oblique line and is found in the middle third of this line.[3]

The thyroid cartilage is lined by a thick layer of perichondrium on all surfaces except the inner surface at the anterior commissure. At this point are attached five ligaments, which form the scaffolding for the corresponding laryngeal folds. From superior to inferior, they are the median thyroepiglottic ligament (median thyrohyoid fold), bilateral vestibular ligaments (vestibular folds or false cords), and bilateral vocal ligaments (vocal folds) (Figure 47–2, A).

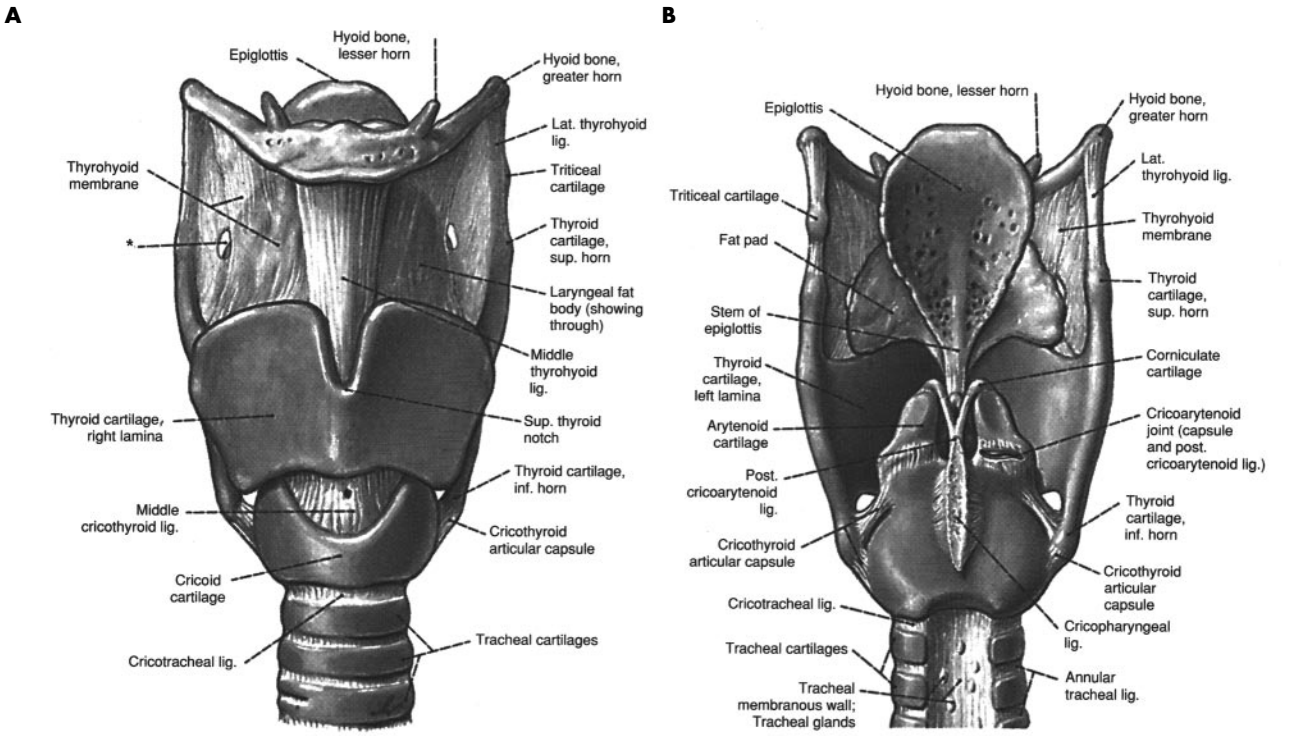


FIGURE 47–1. *A*, Anterior view of the cartilages and ligaments of the larynx and hyoid bone. *B*, Posterior view of the cartilages, ligaments, and articulations of the larynx and hyoid. *Passage for superior laryngeal nerve and vessels. Reproduced with permission from Staubesand J, Taylor AN. Sabotta: atlas of human anatomy. Vol 1. 11th ed. Baltimore: Urban and Schwarzenberg; 1990.

The attachment of these ligaments penetrates the inner perichondrium, forming Broyle's ligament. This ligament contains blood vessels and lymphatics and constitutes an important barrier to the spread of laryngeal neoplasms.

Cricoid Cartilage The cricoid cartilage is a complete ring.[2–7] It is the only supporting structure that completely encircles the airway and serves as the major support for the functioning larynx. Its shape is classically described as that of a signet ring, with the anterior arch measuring 3 to 7 mm in height and the posterior lamina about 20 to 30 mm in height.[4–6] Its inferior border is nearly horizontal and is attached to the first tracheal cartilage by the cricotracheal ligament.

On the posterior surface of the cricoid, the posterior cricoarytenoid muscles are attached in depressions, which are separated by a midline vertical ridge. These muscles are the only abductors of the vocal folds. Attached to this midline vertical ridge are two fasciculi of longitudinal fibers of the esophagus. Housed on the superior surface of the posterior cricoid lamina are the paired arytenoid cartilages.

Posterior to anterior, the cricoid lamina slopes steeply downward to form the anterior cricoid arch. In the midline, between the superior portion of the arch and the inferior border of the thyroid cartilage, is the cricothyroid membrane. It is this structure that must be incised in performing an emergent cricothyrotomy.

Arytenoid Cartilages The arytenoids are paired cartilages that articulate with the posterosuperior portion of the cricoid cartilage.[2–7] Movement of these cartilages and their attached vocal folds allows the larynx its often diverse and complex functions. Each arytenoid is roughly pyramidal in shape, giving it a base, an apex, and three sides. The base of the arytenoid provides the articular facet as well as the muscular and vocal processes. The cricoarytenoid joint is a synovial joint with complex movements that are somewhat debated. It appears, however, that the most important movement of the joint is a rocking motion

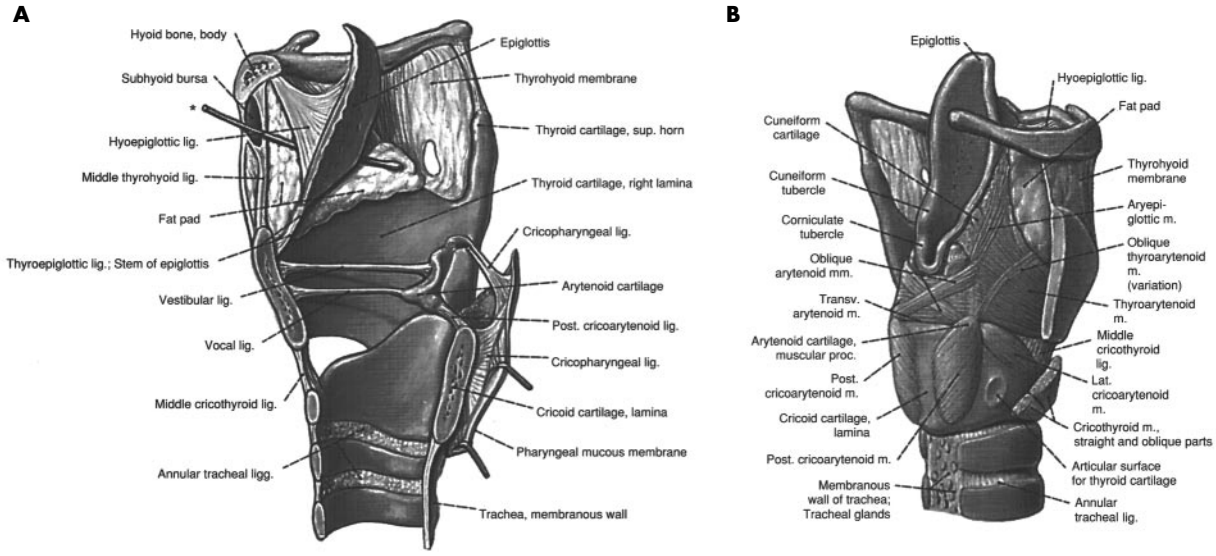


FIGURE 47–2. *A*, Sagittal view of the ligaments and articulations of the larynx. *B*, Posterolateral view of the intrinsic musculature of the larynx. Reproduced with permission from Staubesand J, Taylor AN. Sabotta: atlas of human anatomy. Vol 1. 11th ed. Baltimore: Urban and Schwarzenberg; 1990.

of the cartilage around the long axis of its facet. Laterally, the base forms a broad muscular process, and anteriorly, it forms the thinner vocal process. The anterolateral surface receives the vestibular ligament as well as the thyroarytenoid and vocalis muscles. The posterior surface receives muscular attachments, and to the medial surface is attached the prominent posterior cricoarytenoid ligament. Sitting at the apex of the arytenoid is the corniculate cartilage.

Corniculate and Cuneiform Cartilages These are small, paired fibroelastic cartilages found in the larynx.[4] The corniculate, or cartilage of Santorini, is housed on the apex of the arytenoid cartilage. The cuneiform, or cartilage of Wrisberg, when present, is lateral to the corniculate cartilages and is embedded in the aryepiglottic fold. Although some feel that these cartilages are vestigial, they do appear to add rigidity to the aryepiglottic fold.[7] This rigidity augments the important rampart function of these folds, thus diverting swallowed matter laterally away from the larynx into the piriform sinuses.

Epiglottis The epiglottis is a leaf-shaped elastic fibrocartilage that functions mainly as a backstop against the entrance of swallowed matter into the laryngeal aditus.[2–7] During swallowing, the larynx is raised anterosuperiorly. This action pushes the epiglottis against the base of the tongue, posteriorly displacing it over the laryngeal aditus. The epiglottis has two anterior attachments. Superiorly, it is attached to the hyoid bone by the hyoepiglottic ligament. Inferiorly at the stem or petiole, it is attached to the inner surface of the thyroid cartilage just above the anterior commissure by the thyroepiglottic ligament. The surface of the epiglottic cartilage has multiple pits and is filled with mucous glands; these pits potentially allow the spread of cancer from one surface of the epiglottis to the other.

The epiglottis may arbitrarily be divided into a suprahyoid and an infrahyoid portion. The suprahyoid portion is free on both of its laryngeal and lingual surfaces, with the laryngeal mucosal surface being more adherent than the lingual. As the mucosa of the laryngeal surface is reflected back onto the base of the tongue, three folds result: two lateral glossoepiglottic folds and a median glossoepiglottic fold. The two depressions formed by these folds are known as the valleculae (little depression in Latin). The infrahyoid portion is free only on its laryngeal or posterior surface. This surface contains a small protuberance known as the tubercle. Between the anterior surface and the thyrohyoid membrane and the thyroid cartilage exists

a fat pad within the preepiglottic space. Attached laterally is the quadrangular membrane extending to the arytenoid and corniculate cartilages, constituting the aryepiglottic folds.

Ossification of Laryngeal Cartilages It has long been recognized that incomplete ossification of the laryngeal cartilages can be mistaken for a foreign body on plain roentgenograms of the neck.[8] This particularly applies to ossification of the superior and inferior cornua of the thyroid cartilage and linear ossification of the posterior portion of the cricoid. Thus, the need for understanding the normal ossification pattern of the larynx is self-evident. It is important to realize that only those structures composed of hyaline cartilage will undergo ossification (ie, thyroid, cricoid, and arytenoid cartilages).[8,9] It should be noted that the hyoid bone is completely ossified at 2 years of age and is generally not a point of radiographic confusion.

The thyroid cartilage undergoes ossification in the male about age 20 and in the female a few years later. Ossification begins posteroinferiorly on the lamina. It then extends anteriorly on the inferior border and superiorly at the posterior border. At this time, nuclei of ossification can be seen in the inferior and superior cornua. The cricoid and arytenoid cartilages undergo ossification somewhat later than the thyroid cartilage. Ossification of the cricoid cartilage generally begins at the inferior border, although the superior margin of the quadrate lamina may be an early site of ossification.

Neoplastic invasion of the laryngeal cartilages generally takes place in the ossified portion of the cartilage.[10] The incomplete ossification pattern may make it difficult to appreciate small areas of invasion.

Elastic Tissues

The elastic tissue of the larynx consists of two main parts: (1) the quadrangular membrane of the supraglottic larynx and (2) the thicker conus elasticus and vocal ligaments of the glottic and infraglottic larynx.

The quadrangular membrane attaches anteriorly to the lateral margin of the epiglottis and curves posteriorly to attach to the arytenoid and corniculate cartilages. This structure and the overlying mucosa constitute the aryepiglottic folds. Each fold forms part of the medial wall of each piriform sinus. The inferior edge of the quadrangular membrane constitutes the vestibular ligaments.

The conus elasticus is a thicker elastic structure than the quadrangular membrane. It attaches inferiorly at the superior border of the cricoid cartilage. It then projects upward and medial to its superior attachments, the anterior commissure of the thyroid cartilage and the vocal process of the arytenoid. Between these superior attachments, the conus thickens to form the vocal ligament. Anteriorly, the conus forms the cricothyroid membrane, and in the midline, this membrane condenses to form the cricothyroid ligament (Figures 47–1 to 47–3). The superior extension of the conus (thyroglottic membrane) parallels the superior surface of the true cord. Because it may normally be incomplete, it forms an imperfect barrier to the inferior extension of the transglottic cancers (Figure 47–4).[7]

Muscles

Extrinsic Muscles The extrinsic muscles of the larynx are those muscles of the laryngohyoid complex that serve to raise, lower, or stabilize the larynx.[2–7] Those muscles that elevate the larynx are the thyrohyoid, stylohyoid, digastric, geniohyoid, mylohyoid, and stylopharyngeus. These muscles are important in the elevation and anterior displacement of the larynx during swallowing. They also help to suspend the larynx, via the hyoid bone, from the skull base and mandible. The principal depressors of the larynx are the omohyoid, sternothyroid, and sternohyoid. These muscles displace the larynx downward during inspiration. The middle constrictor, inferior constrictor, and cricopharyngeus muscles are also important extrinsic laryngeal muscles. The proper functioning of these muscles is crucial to the precisely timed swallowing reflex.

Intrinsic Muscles The intrinsic muscles of the larynx are those muscles that are anatomically restricted to the larynx proper. They modify the size of the glottic opening along with the length and tension on the vocal folds. They consist of multiple adductors but only a single abductor. With the exception of the interarytenoid, the intrinsic muscles are paired, and these paired muscles appear to act synchronously (Figure 47–5).

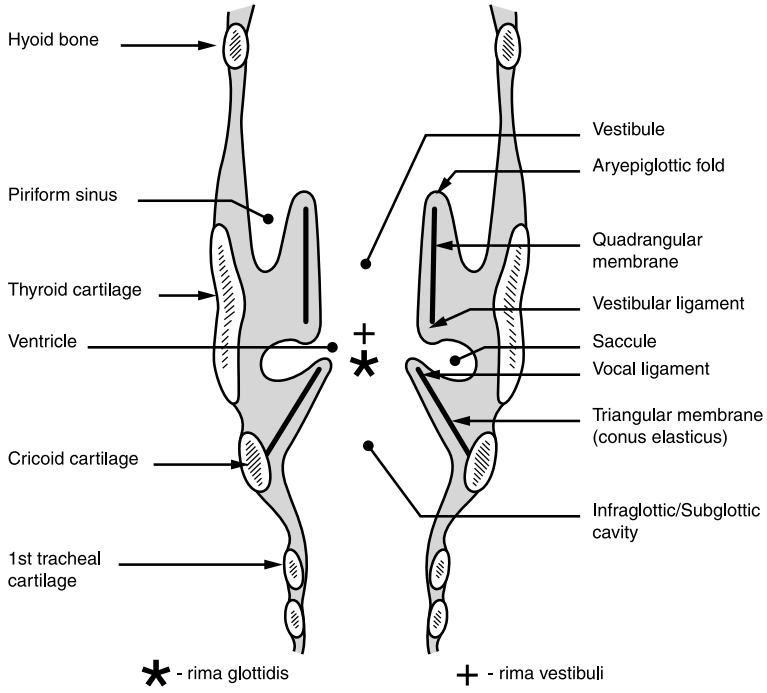


FIGURE 47–3. Coronal section of the larynx, demonstrating the internal cavity and its subdivisions. Note the valve-like nature of the true and false vocal folds. Reproduced with permission from Cooper MH. Anatomy of the larynx. In: Blitzer A, Brin MF, Sasaki CT, et al, editors. Neurologic disorders of the larynx. New York: Thieme Medical Publishers; 1992.

Cricothyroid Muscle. The cricothyroid muscle is located on the external surface of the laryngeal cartilages. It is classically described as consisting of two bellies. The straight portion or pars recta attaches the lateral portion of the anterior arch of the cricoid cartilage to the inferior border of the thyroid cartilage in a fairly vertical direction. The second belly, the pars obliqua, also from the anterolateral border of the cricoid arch, travels obliquely upward to insert on the anterior portion of the inferior cornu. When the right and left cricothyroid muscles contract, they rotate the cricoid at the cricothyroid joint. This action brings the anterior arch of the cricoid superiorly toward the inferior border of the thyroid laminae while displacing the posterior cricoid lamina (and the arytenoid cartilages) inferiorly. This inferior displacement increases the distance between the vocal processes and the anterior commissure; the result of this action is to lower, stretch, and thin the vocal folds while bringing them into a paramedian position. The stretching of the vocal fold also sharpens the edge of the vocal fold and passively stiffens the component layers of the vocal fold (Figure 47–6). Biomechanically, this translates into a higher fundamental frequency produced by the vocal folds.

Posterior Cricoarytenoid Muscle. This muscle is the sole abductor of the vocal folds. It is seated in a depression on the posterior surface of the cricoid lamina, and its fibers run obliquely superior and lateral to attach onto the muscular process of the arytenoid cartilage. It is composed of two compartments: horizontal and vertical bellies. Contraction of these fibers brings the muscular process medial, posterior, and inferior while laterally rotating and elevating the vocal process. This action abducts, elongates, and thins the vocal folds while causing the vocal fold edge to be rounded. The stretching of the vocal fold leads to passive stiffening of its layers. The complex function of this muscle has been studied in the canine in which three distinct neuromuscular compartments are found. It is proposed that the vertical and oblique bellies normally cause vocal fold abduction during respiration, whereas the horizontal belly is primarily used to adjust the position of the vocal process during phonation.[11,12]

Lateral Cricoarytenoid Muscle. This is the main antagonist of the posterior cricoarytenoid. It attaches along the superior border of the cricoid cartilage and then sends its fibers posteriorly to insert on the anterior portion of the muscular process.

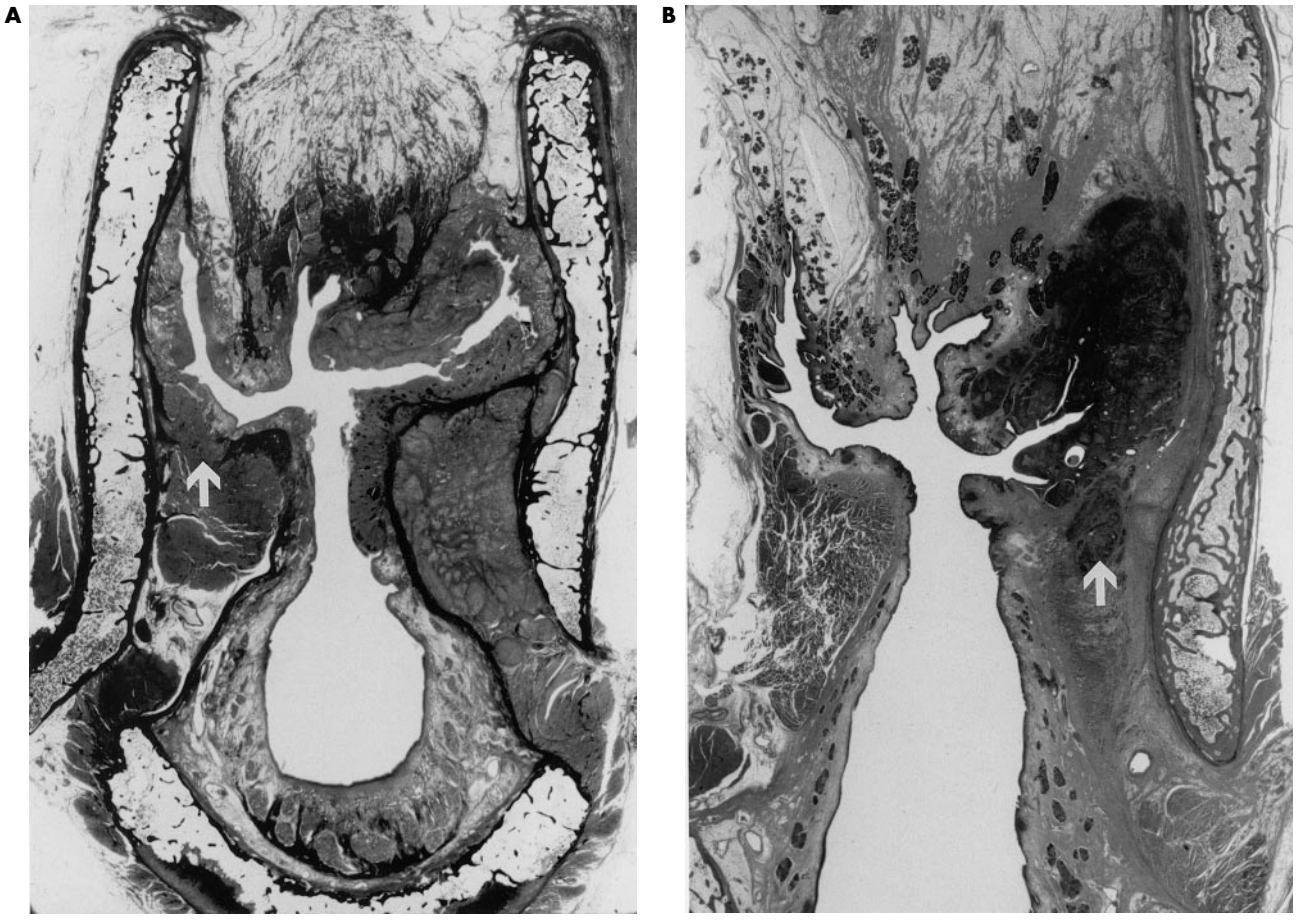


FIGURE 47–4. A, The thyroglottic membrane may be naturally dehiscent (*arrow*). *B*, Dehiscence in the thyroglottic membrane can allow transglottic cancer to extend inferiorly along the paraglottic space (*arrow*). Courtesy of Dr. John A. Kirchner.

Contraction of this muscle brings the muscular process anterolaterally while adducting and lowering the vocal process. This results in adduction, elongation, and thinning of the vocal folds. The edge of the vocal fold becomes sharper, and its component layers are passively stiffened.

Interarytenoid/Aryepiglottic Muscle. The interarytenoid is the only unpaired intrinsic muscle, consisting of two types of muscle fibers. The bulk of the muscle consists of transverse fibers passing from the posterior surface of one arytenoid cartilage to the posterior surface of the other. This muscle contracts to bring together the arytenoid cartilages, thus assisting in closing the posterior portion of the glottis. This does not significantly affect the mechanical properties of the vocal folds. Along with these transverse fibers are oblique fibers. These oblique fibers pass from the posterior portion of the arytenoid on one side to the apex of the arytenoid on the other side, thus crossing in the midline. Some fibers insert at the apex, whereas others travel along the quadrangular membrane. These fibers contract to narrow the laryngeal aditus. Those fibers traveling along the quadrangular membrane (thus the aryepiglottic fold) constitute the aryepiglottic muscle.

Thyroarytenoid Muscle. This muscle is classically divided into the thyroarytenoid internus and externus. These have the same attachments, but the internus lies deep or internal to the externus. In addition, the internus is more well developed than the externus.

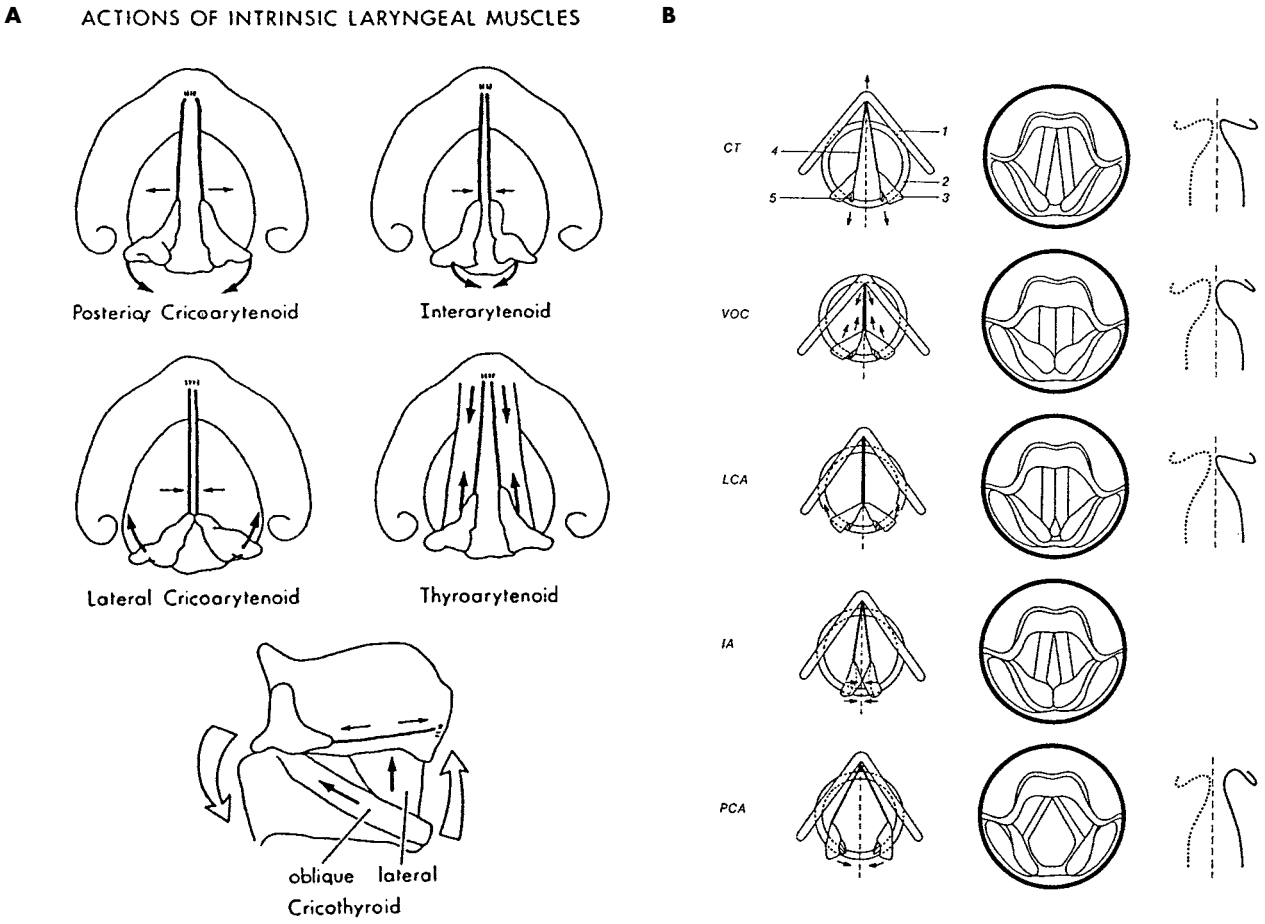


FIGURE 47–5. *A*, The intrinsic muscles of the larynx and their actions. *Heavy arrows* indicate the direction of muscle action, *fine arrows* indicate the motion of vocal ligaments, *open arrows* indicate the motion of cricoid and thyroid cartilages. Reproduced with permission from Ballenger JJ, editor. Diseases of the nose, throat, ear, head and neck. 13th ed. Philadelphia: Lea & Febiger; 1985. *B*, Schematic presentation of the function of the laryngeal muscles. The *left column* shows the location of the cartilages and the edge of the vocal folds when the laryngeal muscles are activated individually. The *arrow* indicates the direction of the force exerted. CT = cricothyroid muscle; VOC = vocalis muscle; LCA = lateral cricoarytenoid muscle; IA = interarytenoid muscle; PCA = posterior cricoarytenoid muscle; 1 = thyroid cartilage; 2 = cricoid cartilage; 3 = arytenoid cartilage; 4 = vocal ligament; 5 = posterior cricoarytenoid ligament. The *middle column* shows views from above. The *right column* presents contours of frontal sections at the middle of the membranous portion of the vocal fold. The *dotted line* shows a control, where no muscle is activated. Reproduced with permission from Hirano M, editor. Clinical examination of voice. New York: Springer-Verlag; 1981. Copyright Urban & Fischer Verlag.

The thyroarytenoid externus arises from the anterior commissure and inserts onto the lateral surface of the arytenoid cartilage. It contracts to bring the vocal process and anterior commissure closer to each other, thus adducting the vocal folds. It also contracts to adduct the false cords. The externus sends a few slips of muscle fibers onto the quadrangular membrane to establish the thyroepiglottic muscle. This muscle, like the aryepiglottic muscle, acts to narrow the laryngeal inlet.

The thyroarytenoid internus or vocalis muscle attaches at the anterior commissure and inserts onto the vocal process, sending a few slips of fibers below the vocal ligament onto the conus elasticus. It contracts to adduct, shorten, thicken, and lower the vocal fold while rounding its edge. The body (muscle) of the vocal fold is actively stiffened, whereas the cover is passively slackened.

Recently, immunohistochemical staining for myofibrillar adenosine triphosphatase reveals that

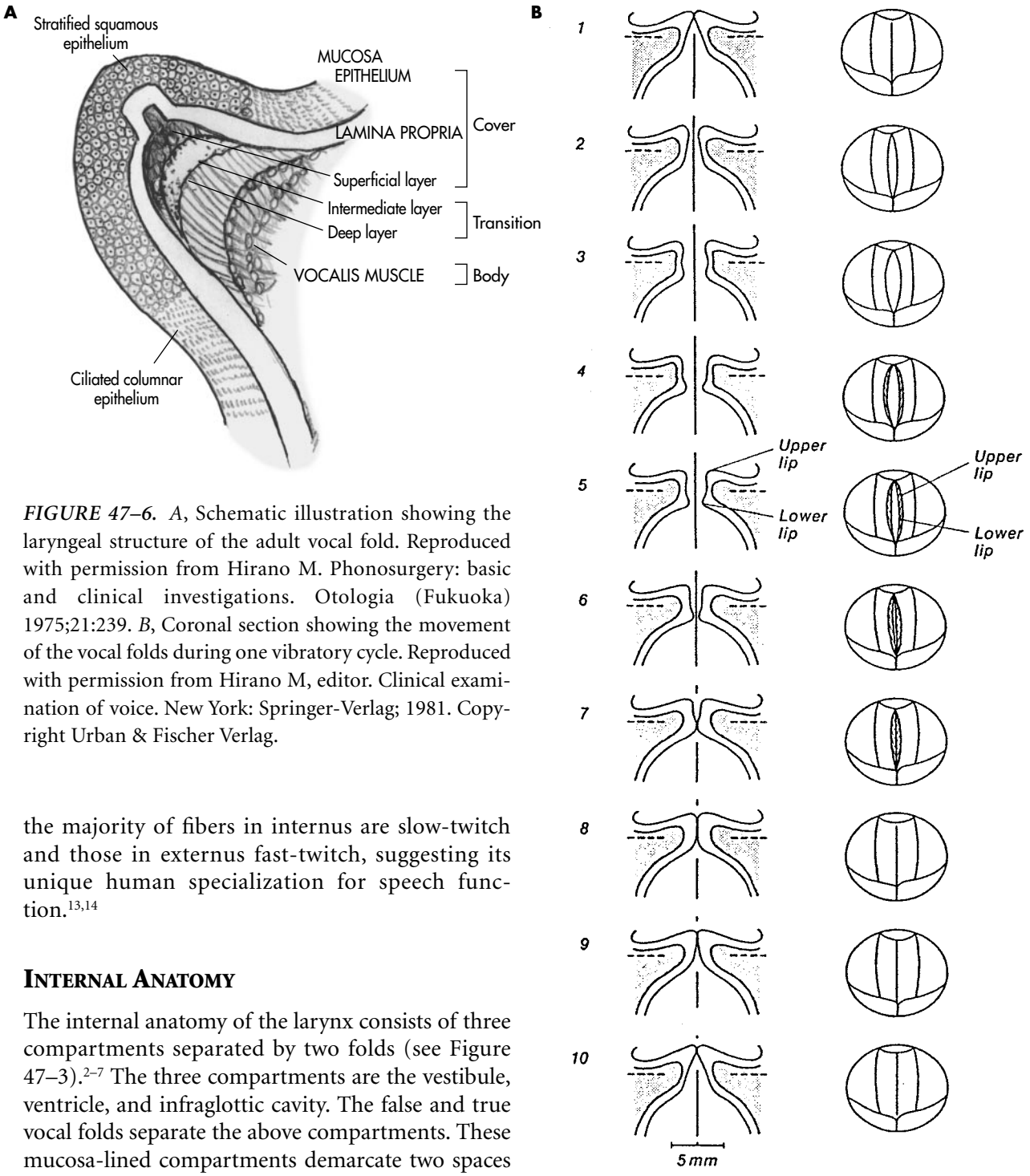


FIGURE 47–6. A, Schematic illustration showing the laryngeal structure of the adult vocal fold. Reproduced with permission from Hirano M. Phonosurgery: basic and clinical investigations. Otologia (Fukuoka) 1975;21:239. *B*, Coronal section showing the movement of the vocal folds during one vibratory cycle. Reproduced with permission from Hirano M, editor. Clinical examination of voice. New York: Springer-Verlag; 1981. Copyright Urban & Fischer Verlag.

the majority of fibers in internus are slow-twitch and those in externus fast-twitch, suggesting its unique human specialization for speech function.[13,14]

Internal Anatomy

The internal anatomy of the larynx consists of three compartments separated by two folds (see Figure 47–3).[2–7] The three compartments are the vestibule, ventricle, and infraglottic cavity. The false and true vocal folds separate the above compartments. These mucosa-lined compartments demarcate two spaces of importance: the preepiglottic space and the paraglottic space.

Vestibule A laryngoscopic view of the larynx reveals the vestibule as that portion of the larynx from the tip of the epiglottis to the false or vestibular folds. Thus, the vestibule is bound by the epiglottis anteriorly, the aryepiglottic folds laterally, and the arytenoid and corniculate cartilages with the interarytenoid muscle posteriorly. In the laryngoscopic view, the anterior commissure is frequently hidden by the protuberance of the epiglottis known as the tubercle.

As previously discussed, the vestibular folds are formed by mucosa overlying the vestibular ligament

(inferior border of the quadrangular membrane). The submucosa of the ventricle contains numerous seromucinous glands. The secretions produced by these exocrine glands provide both mechanical and immune (lysozyme) protection for the vocal folds.[15]

Ventricle The ventricle or sinus of Morgagni is the small space between the false and true vocal folds. The ventricle is often hidden during laryngoscopic examination of the larynx unless exposed by lateralization of the false vocal fold. At the anterior end of the ventricle is a diverticulum known as the laryngeal saccule. The saccule (of Hilton) is lined with mucous glands, which are thought to lubricate the vocal folds. Fibers of the thyroarytenoid muscle line the walls of the saccule and are thought to express mucus from the saccule when they contract. The size of the saccule is quite variable; however, it seldom extends above the superior border of the thyroid cartilage. Abnormal dilatation of the saccule results in an air-filled laryngocele that should be distinguished from a mucocele of the saccule (saccular cyst), which lacks free communication with the ventricle and thus is not air filled.[16] The vocal folds that form the inferior boundary of the ventricle are described in more detail in a separate section.

Infraglottic Cavity The infraglottic cavity extends from the glottis down to the inferior border of the cricoid cartilage. Its lateral boundary is formed by the conus elasticus and walls of the cricoid cartilage.

Piriform Sinus Although the piriform sinus is anatomically part of the hypopharynx, understanding its anatomy and its relationship to the larynx is essential. The piriform sinus is a gutter formed by the aryepiglottic fold, arytenoid, and superior cricoid medially and the thyrohyoid membrane and internal surface of the thyroid lamina laterally. Superiorly, it begins at the lateral glossoepiglottic fold. Inferiorly, the apex of the sinus blends with the esophageal inlet at about the superior border of the cricoid (Figure 47–7). Thus, cancer invasion of the apex implies the necessity of removing a portion of the cricoid if conservation laryngectomy is planned.[7]

There are two important markings within the piriform sinus. Anteriorly in the floor of the sinus, a small fold can be seen, which marks the course of the superior laryngeal nerve. This submucosal course of the nerve makes it possible to anesthetize the nerve topically in the piriform sinus. The second, more variable landmark is the protrusion made into the sinus from the superior cornu of the thyroid cartilage. This smooth protrusion, which usually presents in the elderly, should not be confused with neoplasm.

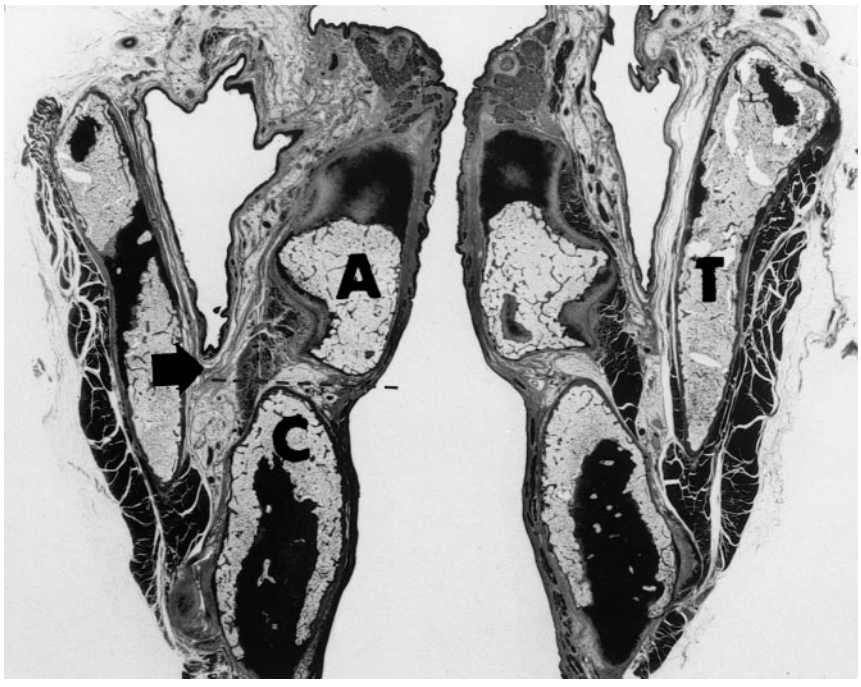


FIGURE 47–7. Coronal section through the human larynx demonstrating the relationship of the cricoid cartilage (C) with the apex of the piriform sinus (*arrow*). A = arytenoid cartilage; T = thyroid cartilage. Courtesy of Dr. John A. Kirchner.

Mucosa The mucosa of the larynx is of two types: pseudostratified ciliated columnar cell (respiratory) epithelium and squamous cell epithelium. Much of the larynx is surfaced by respiratory epithelium; however, the superior portion of the epiglottis, upper portions of the aryepiglottic folds, and free edges of the vocal folds are surfaced by squamous cell epithelium. Beneath this covering epithelium is a variable basement membrane, and separating these two is a layer of loose fibrous stroma. It should be noted that this loose fibrous layer is absent on the true vocal folds as well as the laryngeal (posterior) surface of the epiglottis. The absence of this layer on the posterior surface of the epiglottis accounts for the more intense swelling of the lingual (anterior) surface of the epiglottis in inflammatory conditions of the larynx.

Preepiglottic Space The preepiglottic space, as its name implies, lies anterior to the epiglottis, which serves as its posterior boundary. It is bound superiorly by the hyoepiglottic ligament and mucosa of the valleculae and inferiorly by the thyroepiglottic ligament. The anterior boundaries are the thyrohyoid membrane and the inner surfaces of the thyroid laminae. Laterally, the preepiglottic space opens in the paraglottic spaces. Cancer on the infrahyoid portion of the epiglottis can penetrate this structure and gain access to the preepiglottic space.

Paraglottic Space A paraglottic space, as its name implies, lies on each side of the glottis. This space lies above and below the true and false vocal folds and is important in the transglottic and extralaryngeal spread of neoplasms. Medially, the space is bound by the quadrangular membrane, ventricle, and conus elasticus. Laterally, it is bound by the perichondrium of the thyroid lamina and the cricothyroid membrane. Anterosuperiorly, the space opens in the posterior portion of the preepiglottic space. The mucosa of the piriform sinus forms the posterior boundary. The relationships of this paraglottic space make it important in considering the spread of laryngeal cancer. Supraglottic cancer invading into this space may quickly extend extralaryngeally.

Vocal Folds

The anatomy of the vocal folds is complex and is thus considered separately. The vocal fold is considered the structure between the vocal process of the arytenoid and the anterior commissure. The vocal folds and the slit between them (rima glottidis) constitute the glottis. The glottis can be divided by a horizontal line between the tips of the vocal processes. This imaginary line divides the glottis into an intermembranous portion and an intercartilaginous portion. The anterior to posterior (length) ratio of the intermembranous portion to the intercartilaginous is 3 to 2; however, the ratio of cross-sectional areas defined by them is 2 to 3. Thus, owing to its more rectangular shape, the intercartilaginous portion is larger. Some have called this the respiratory portion of the rima.[4,17] The membranous or vibratory portion of the vocal folds consists of three well-defined structural layers. From superficial to deep, they are the epithelium, lamina propria (three layers), and vocalis muscle. Hirano divided these layers according to a body-cover concept (see Figure 47–6, A).[17] The cover consists of the overlying epithelium and the gelatinous superficial layer of the lamina propria. The body consists of the vocalis muscle, which he likened to thick rubber bands. Between these exists a transition zone composed of the intermediate (elastic) and deep (collagenous) layers of the lamina propria. According to this concept, the vocal folds consist of a multilayered vibrator with increasing stiffness from the cover to the body. Thus, the cover is responsible for most of the vibratory action of the vocal folds. At the anterior and posterior ends of the vocal folds exists an anterior and a posterior macula flava, respectively. These are essentially a thickening of the intermediate (elastic) layer of the lamina propria and are thought to function as "cushions" protecting the ends of the vocal folds from vibratory damage. It should be noted that the same body-cover concept does not apply to the larynx of children because of the more homogeneous nature of the lamina propria. It is not until nearly the end of adolescence that the lamina matures into its adult form. In the senile larynx, the elastic layer and the vocalis muscle tend to atrophy, whereas the collagenous layer thickens. The cover becomes thickened and edematous secondary to changes in the superficial layer of the lamina, whereas the epithelium itself changes little.

The shape of the true and false vocal folds carries biomechanical significance. When seen in coronal section, both appear as valve-like structures, with the leaflets of the false folds pointing inferiorly and those of the true folds pointing superiorly (see Figure 47–3). Thus, the false folds passively impede egress of air, whereas the true folds impede its ingress. Work-

ing with cadaver larynges, Brunton and Cash demonstrated that the false folds offered a resistance equaling 30 mm Hg to the egress of air from below, whereas the true folds offered a resistance equaling 140 mm Hg to the ingress of air from above.[18] Both structures offered little resistance to the opposite flow of air (ie, they act as one-way valves).

Vessels

Arteries and Veins The arterial supply to the larynx consists of the superior and inferior laryngeal arteries.[2–7] After the superior thyroid artery branches off the external carotid, it courses lateral to the laryngohyoid complex and gives off the superior laryngeal artery at approximately the level of the hyoid bone. This artery then runs anteromedially with the internal branch of the superior laryngeal nerve to enter the thyrohyoid membrane inferior to the nerve. It then enters the submucosa of the piriform sinus and is distributed to intralaryngeal structures. The superior thyroid also gives off a cricothyroid branch that courses horizontally below the thyroid cartilage. The inferior laryngeal artery is a branch of the inferior thyroid artery that comes off the thyrocervical trunk branching from the subclavian artery. After coursing posterior to the cricothyroid joint with the recurrent laryngeal nerve, the artery enters the larynx by passing through a gap in the inferior constrictor muscle known as the Killian-Jamieson area. The artery is then distributed to the remainder of the internal larynx, making multiple anastomoses with the superior laryngeal artery. The venous supply parallels the arterial supply.

Lymphatics An appreciation of the lymphatics of the larynx is prerequisite to understanding the spread of cancer of the larynx, as well as the operative procedures designed to eradicate the disease.[2–7,19] The lymphatics of the larynx are divided into superficial (intramucosal) and deep (submucosal) groups. The deep network is further divided into right and left halves, with little communication between them. These two halves can be further divided into supraglottic, glottic, and subglottic, with special consideration given to the ventricle in the supraglottic region. Although the superficial network is richly anastomotic throughout the larynx, it is the deep network that is important in the spread of cancer and will be given further consideration.

The drainage of the supraglottic structures (aryepiglottic folds and false cords) follows the superior laryngeal and superior thyroid vessels. Thus, the lymphatics flow from the piriform sinus through the thyrohyoid membrane to end primarily in the deep jugular chain around the carotid bifurcation. It should be noted that the epiglottis is a midline structure; thus, its lymphatic drainage is bilateral.

The lymphatic drainage of the ventricle is different from the other supraglottic structures. Dye injected into the ventricle enters the paraglottic space and is quickly spread by the lymphatic system through the cricothyroid membrane and also into the ipsilateral lobe of the thyroid (justifying its resection in laryngectomy).

The true vocal folds are devoid of lymphatics, accounting for the high curability of cancer localized to this structure.

The subglottic larynx has two lymphatic drainage systems. One system follows the inferior thyroid vessels to end in the lower portion of the deep jugular chain as well as the subclavian, paratracheal, and tracheoesophageal chains. The other system pierces the cricothyroid membrane. This system appears to receive lymphatics from both sides of the larynx and disseminate bilaterally to the middle deep cervical nodes as well as the prelaryngeal (Delphian) nodes.

PHYSIOLOGY

As previously stated, the three functions of the larynx in order of priority are (1) protection of the lower airway, (2) respiration, and (3) phonation.[1] A flawless performance of these functions requires an intact neuromuscular system to respond to both volitional and reflex signals presented to the larynx.

Innervation

The pattern of innervation to and from the larynx and the type and distribution of its receptors determine the functional capabilities of the larynx. The larynx is innervated by the superior and inferior laryngeal nerves. The superior laryngeal nerve leaves the nodose ganglion to pass between the carotid artery and the laryngohyoid complex. It divides into a larger internal and smaller external branch. The internal branch pierces the thyrohyoid membrane with the superior laryngeal artery and becomes the

sensory supply to the ipsilateral supraglottic larynx, whereas the external branch innervates the cricothyroid and inferior constrictor muscles. The inferior laryngeal nerve originates from the recurrent laryngeal nerve and runs in the tracheoesophageal groove. It enters the larynx posterior to the cricothyroid joint and classically divides into an anterior adductor and a posterior abductor branch. This branching, however, is quite variable, as is the muscular innervation from the branches.[4]

The receptors of the larynx can be divided into mucosal, articular, and myotatic. These receptors mediate much of the reflex activity of the larynx. The mucosal receptors respond to stimuli such as touch, mucosal deformation (mechanoreceptors), and liquids. The articular receptors are located in the joint capsule and respond to deformation of the capsule. The myotatic receptors respond to muscular stretch and appear to be most abundant in the vocalis muscle.[20]

Afferent System The sensory innervation to the mucosa of the supraglottic larynx is carried by the internal branch of the ipsilateral superior laryngeal nerve, which is divided into three divisions. The superior division mainly supplies the mucosa of the laryngeal surface of the epiglottis, the middle division supplies the mucosa of the true and false vocal folds and the aryepiglottic fold, and the inferior division supplies the mucosa of the arytenoid, part of the subglottis, the anterior wall of the hypopharynx, and the upper esophageal sphincter.[21] The cricoarytenoid joint and thyroepiglottic ligament are both innervated by the internal branch of the superior laryngeal nerve. The inferior laryngeal nerve supplies the major portions of mucosa below the glottis as well as muscle spindles of intrinsic muscles. The external branch of the superior laryngeal nerve contains afferent fibers from the cricothyroid joint and from deep muscle receptors.

Histologic examination has revealed the presence of free nerve endings, Merkel cells, Meissner's corpuscles, and taste buds scattered in the larynx. Mechanoreceptors are located either in the superficial mucosa or in the muscles and laryngeal joints. Some of them are spontaneously active, whereas others are silent until stimulated. The supraglottic larynx also has chemical and thermal sensors. A large number of taste buds populate the laryngeal surface of the epiglottis and extend caudally along the aryepiglottic folds, reaching peak density at the caudal extreme of the folds.[22,23] This sensory distribution facilitates the supraglottic larynx in its role of preventing foreign material from penetrating into the larynx.[24] These respond to a number of chemical stimuli and to water. The taste buds of the larynx tend to be most sensitive to the pH and tonicity of the stimulus. In this regard, the water receptors of the epiglottis appear to play a role in the production of prolonged apnea. When stimulated, they lead to a slowing of respiration with an increase in tidal volume. Chemoreceptors of the larynx are adapted to detect chemicals that are not saline-like in composition and can play an important role in modifying reflexes involved in the maintenance of upper airway patency.[23] It is important to note that this same response does not occur with saline. It is felt that the receptors may be responding to the washout of chloride ions.[20] Theoretically, this may be the mechanism by which cold-steam mist assists the child with croup (slowing and deepening breathing, thus decreasing turbulent flow).[25] Interestingly, this response is more potent early in life. This apneic response has also been implicated in sudden infant death syndrome.[20] The vocal folds also have touch receptors that are more abundant posteriorly than anteriorly. The afferent impulses generated are delivered to the tractus solitarius through the ganglion nodosum.

Efferent System The motor innervation is primarily through the inferior laryngeal nerve. This nerve innervates all of the intrinsic muscles of the larynx except the cricothyroid, which is innervated by the external branch of the superior laryngeal nerve. Each nerve is responsible for the muscles on the ipsilateral side of the larynx, with the exception of the interarytenoid muscle. Thus, the only unpaired muscle of the larynx receives its innervation from both inferior laryngeal nerves. Injury to the recurrent laryngeal nerve leaves the injured cord in the paramedian position, resulting from the adductor effect of the intact cricothyroid. Unilateral injury to the superior laryngeal nerve causes the posterior glottic opening to rotate to the paralyzed side, bowing the paralyzed cord.[4,24] These changes in the larynx can be seen on laryngeal endoscopy.

Neurophysiology of Protective Function

The glottic closure reflex is a polysynaptic reflex that allows the larynx to protect the lower airway from

penetration and aspiration. However, when exaggerated, this reflex accounts for the production of laryngospasm. Protective closure of the larynx occurs in three tiers.

In the first tier, the laryngeal inlet is contracted by collapsing the aryepiglottic folds medially. The anterior and posterior gaps are filled by the epiglottic tubercle and the arytenoid cartilages, respectively. In the second tier, the false vocal folds are brought together. The final and most important tier occurs at the level of the true vocal folds. Because the valvular action of the true vocal folds resists ingress of material, they offer the most important level of protection. It should be noted that it is the thyroarytenoid or slips from this muscle that contract at each level of closure. This muscle is one of the fastest contracting of all striated muscles in the body. Classically, the afferent limb of this reflex occurs through stimulation of touch, chemical, or thermal receptors in the supraglottic larynx.[26]

Laryngospasm occurs when stimulation of the superior laryngeal nerve leads to a prolonged adductor response. This response is maintained well after the initiating stimulus is removed, and section of the superior laryngeal nerves abolishes the response. Clinically, this is typically seen in the setting of endotracheal intubation/extubation or after manipulation of the airway, especially if blood has contaminated the laryngeal inlet. The response is dampened in the face of barbiturates, hypercapnia, positive intrathoracic pressure, and severe hypoxia.[27]

Although not classically considered part of the protective reflex, reflex swallowing from stimulation of the superior laryngeal nerve may have protective functions. It has been shown that reflex swallowing occurs with application of hypotonic fluids to the supraglottic larynx, particularly the laryngeal surface of the epiglottis, glottis, and internal larynx. Although this is not the normal mechanism to initiate swallowing, it may serve to protect the larynx against fluid that enters the laryngeal inlet.[20]

Neurophysiology of Respiratory Function

It is intuitive that if the larynx is the sphincter to the lower airway, for respiration to occur, the sphincter needs to be actively opened during inspiration. Also, the opening of the folds must be synchronous with, but slightly precede, the descent of the diaphragm. Through the work of many individuals, this observation is well supported.[24,28,29] The respiratory center of the medulla, with the help of higher central nervous system and peripheral input, maintains eupneic respiration. It drives the synchronous opening of the glottis and descent of the diaphragm during inspiration. The opening of the glottis is primarily through the action of the posterior cricoarytenoid. However, in hyperpneic conditions, the cricothyroid contracts rhythmically with the posterior cricoarytenoid.[28,29] The contraction of both of these muscles increases the glottic aperture in both anteroposterior and horizontal dimensions. During phonation, the cricothyroid lengthens and passively adducts the vocal folds. However, during respiration, when contracted in concert with the posterior cricoarytenoid, the effect is to lengthen the open glottis, thus increasing the cross-sectional area for airflow.

Understanding the role that the cricothyroid plays as an accessory muscle of inspiration underlies the rationale for superior laryngeal nerve section in the face of bilateral recurrent laryngeal nerve paralysis. Bilateral paralysis produces dyspnea, which will lead to cricothyroid contraction, further adducting the paralyzed folds. Unilateral superior laryngeal nerve section reduces glottic resistance by preventing its full adduction.

The rhythmicity of the phrenic nerve and the posterior cricoarytenoid can be increased by hypercapnia and ventilatory obstruction. It is lessened by hypocapnia. The effect of ventilatory resistance on posterior cricoarytenoid activity has been extensively studied in the canine model. In this model, when ventilatory resistance is eliminated, so is the reflex abductor activity of the posterior cricoarytenoid. It is felt that the afferent limb of this reflex resides within the ascending vagus nerve and that the end-organ receptors are located within the thorax, although their precise location is unknown. The longer abductor activity is lost, the more difficult it is to re-establish.[30] This is the rationale for downsizing tracheotomy tubes (thus gradually increasing ventilatory load) prior to decannulation.

The role of the larynx in controlling expiration must also be considered. It is well known that the control of respiratory rate occurs primarily through variation of the expiratory phase. The time of expiration is dependent on the ventilatory resistance pro-

duced by the glottis. As discussed above, the cricothyroid contracting with the posterior cricoarytenoid will give the maximal glottic opening and hence the lowest ventilatory resistance. In this regard, cricothyroid contraction during expiration occurs when the critical subglottic pressure change of 30 cm H_2O/s is exceeded and continues as long as positive subglottic pressure is maintained. As expected, this threshold for activation is reduced in hypercapnia (allowing for quicker expiration and a faster respiratory rate) and increased in hypocapnia.[31]

Neurophysiology of Phonation

The complex mechanisms of phonatory control coordinate central and peripheral components. Electromyographic investigation of the control of peripheral neuromuscular systems involved in phonation has demonstrated specific intrinsic and extrinsic muscle function in humans. Central mechanisms are less well understood and often rely on animal models, which may only infer function in the unique phonatory systems of the human.

As a general model, the larynx, as a system, must respond to central commands from linguistic and motor centers. Signals are then relayed to the motor cortex in the precentral gyrus and on to motor nuclei in the brainstem and spinal cord. These signals are transmitted to the respiratory, laryngeal, and articulatory muscles responsible for speech and voice production. These messages are influenced by the extrapyramidal system, including the cerebral cortex, cerebellum, and basal ganglia, exerting fine control of respiration, phonation, and articulation.[32]

Specific connectivity and the central control of brainstem motor neurons responsible for voluntary control of phonation remain illusive. Laryngeal reflexes, which are key to the coordination of respiration, phonation, and deglutition, are understood primarily from the research focused on respiration and deglutition.[20] Further, central projections to laryngeal mechanisms are not consistent across species and may differ from nonhuman primates to humans as well. The nucleus tractus solitarius (NTS), periaqueductal gray, parabrachial nucleus, locus caeruleus, and ventromedial nucleus of the thalamus are all areas anatomically associated with the laryngeal system.[20] Figure 47–8 illustrates branching of the superior laryngeal nerve with central connectivity and projections of the NTS. However, specific mechanisms of control are not well defined. In some cases, the central terminations of specific sensory receptors and the origin of the motoneuron fibers are known, as in somatotopic organization for the face and limbs. To date, studies of the role of the cerebral cortex in phonation in primates reveal no such individual muscle representation or somatotopic mapping of the laryngeal system.[20]

The role of peripheral mechanisms in phonatory control has been studied more successfully using electromyography, airflow, and pressure studies in the vocal tract as well as imaging techniques, which allow observation of the vocal tract during phonatory postures. It is important to consider that phonation takes place in concert with upper articulators, the lip, tongue, jaw, and velum.

Mechanical tissue deformation and, in particular, the pull of the upper articulators on the larynx via the laryngeal-hyoid complex necessarily influence the phonatory environment. For example, Figure 47–9 illustrates vocal tract shaping for four distinct vowels. Note the apparent pull of the tongue high and forward for production of /i/ (labeled "iy"), in contrast to posterior and low posture for /a/ (labeled "aa"). The consequent constriction or, conversely, the opening of the posterior pharynx plays a key role in the acoustics of sound produced for each of these postures and influences laryngeal posturing for phonation.[33]

Electromyographic studies of specific muscle function indicate firing rates uniquely suited for fine control of laryngeal function. Contraction times for the intrinsic muscles in several species are presented in Table 47–1. Further, the high innervation ratio for human laryngeal muscles, estimated at 100 to 200 cells per motor unit,[34] makes laryngeal muscles capable of a great degree of precise control as required for adjustment of speaking frequency and intensity.

Intrinsic and extrinsic musculature, described previously in this chapter, influence specific action of laryngeal muscles during phonatory shaping of the glottis in sound production as described by Hirano (see Figure 47–5).[32] Table 47–2 summarizes the influence of each of these muscles on the shape and tension of the glottis during phonation.

It has been shown that the laryngeal muscles start to contract about 100 to 200 ms prior to the onset of phonation.[35] Further, the most important muscle in varying the phonation style (from

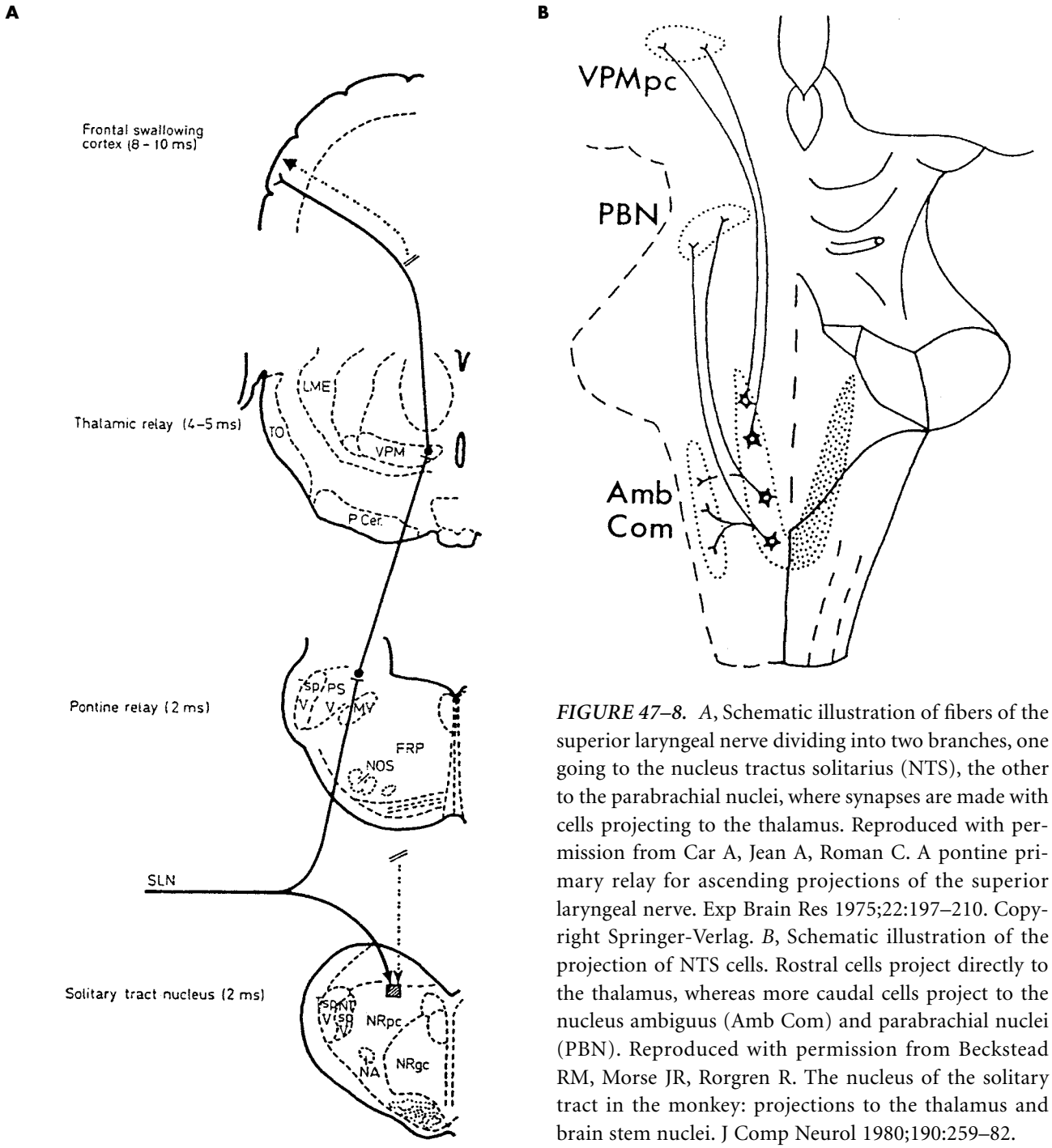


FIGURE 47–8. A, Schematic illustration of fibers of the superior laryngeal nerve dividing into two branches, one going to the nucleus tractus solitarius (NTS), the other to the parabrachial nuclei, where synapses are made with cells projecting to the thalamus. Reproduced with permission from Car A, Jean A, Roman C. A pontine primary relay for ascending projections of the superior laryngeal nerve. Exp Brain Res 1975;22:197–210. Copyright Springer-Verlag. *B*, Schematic illustration of the projection of NTS cells. Rostral cells project directly to the thalamus, whereas more caudal cells project to the nucleus ambiguus (Amb Com) and parabrachial nuclei (PBN). Reproduced with permission from Beckstead RM, Morse JR, Rorgren R. The nucleus of the solitary tract in the monkey: projections to the thalamus and brain stem nuclei. J Comp Neurol 1980;190:259–82.

hypotense to hypertense) is the thyroarytenoid muscle.[36] The frequency of vibration depends on the following: (1) vibratory mass of both vocal folds, (2) the anterior to posterior tension, (3) functional damping at high pitch, and (4) subglottic pressure. In this regard, Gay and associates showed that in the chest register, the cricothyroid and vocalis, with the possible contribution of the posterior cricoarytenoid, most consistently control changes in fundamental frequency.[37]

Lovquist and others later described electromyographic recordings obtained from four

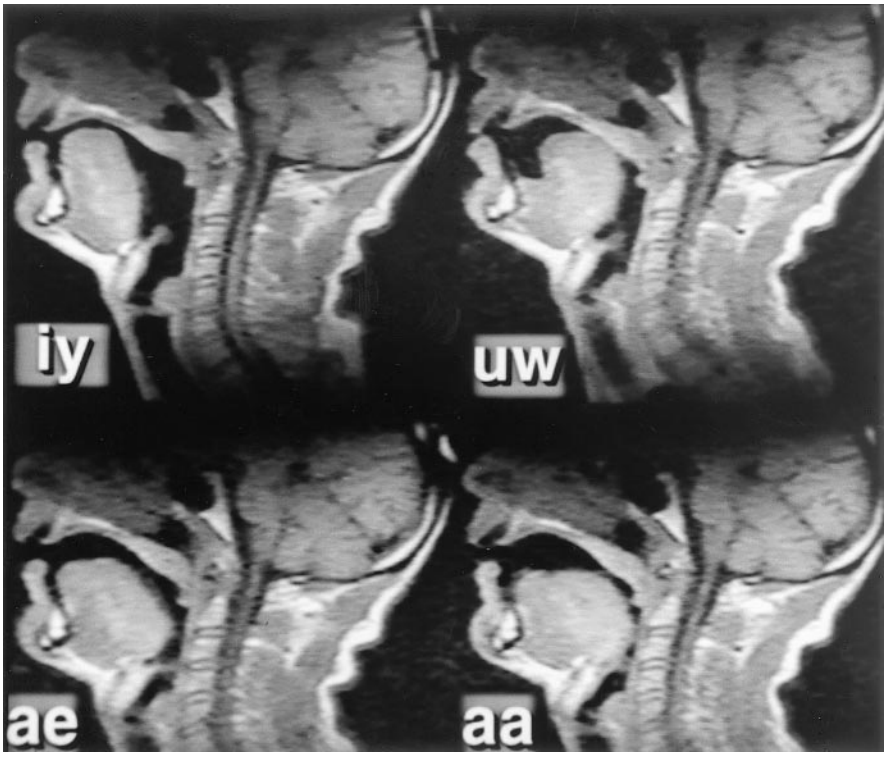


FIGURE 47–9. Magnetic resonance images of the vocal tract during sustained production of four vowels illustrate contrast in shape as viewed in sagittal section. Courtesy of C. Gracco.

intrinsic laryngeal muscles with simultaneous transillumination and acoustic signals.[38] The vocalis and lateral cricoarytenoid muscles were observed to participate in the control of both articulation and phonation. The interarytenoid muscle appeared to be involved only in articulatory adjustments. Activity in the cricothyroid was primarily related to changes in fundamental frequency. This muscle also showed an increase in activity for voiceless sounds. In addition, the vocalis muscle appeared to participate in glottal adduction without complete closure in voiceless clustered sounds, with the lateral cricoarytenoid and the interarytenoid playing no particular roles. Studies such as these underscore the complex interactive nature of laryngeal musculature in phonatory/articulatory interaction.

In considering the phonatory process, a variety of factors necessarily contribute to the acoustic product as defined in Table 47–3. The psychoacoustic parameters of pitch, loudness, quality, and fluctuation are correlated with acoustic qualities of fundamental frequency, amplitude, waveform, and

TABLE 47–1. Contraction Times in Milliseconds for Laryngeal Muscles by Species

	Subject		
Muscle	*Dog*	*Cat*	*Human*
Thyroarytenoid	14	21	35
	12.5	25	—
Posterior cricoarytenoid	30	22	—
	44		
Lateral cricoarytenoid	16	19	—
Cricothyroid	35	44	35

TABLE 47–2. Characteristic Functions of the Laryngeal Muscles in Vocal Fold Adjustments

	CT	*VOC*	*LCA*	*IA*	*PCA*
Position	Paramed	*Adduct*	*Adduct*	*Adduct*	*Abduct*
Level	Lower	Lower	*Lower*	0	*Elevate*
Length	*Elongate*	*Shorten*	Elongate	(Shorten)	*Elongate*
Thickness	*Thin*	*Thicken*	Thin	(Thicken)	Thin
Edge	*Sharpen*	*Round*	Sharpen	0	Round
Muscle (body)	*Stiffen*	*Stiffen*	Stiffen	(Slacken)	Stiffen
Mucosa (cover and transition)	*Stiffen*	*Slacken*	Stiffen	(Slacken)	Stiffen

0 = no effect; ( ) = slightly; italics = markedly; CT = cricothyroid muscle; VOC = vocalis muscle; LCA = lateral cricoarytenoid muscle; IA = interarytenoid muscle; PCA = posterior cricoarytenoid muscle.
Adapted from Hirano M. Clinical examination of voice. New York: Springer-Verlag; 1981.

acoustic spectrum. High-speed photography and observation of vocal fold vibration via videostroboscopic endoscopy reveal much about the behavior of the glottis during phonation.[32,39]

The vibrations of the vocal folds are a passive phenomenon and represent the basis of the aerodynamic theory of sound production. Vibration of the vocal folds changes DC airflow into AC airflow, converting aerodynamic energy to acoustic energy. The mucosal wave produced by these vibrations has been captured on ultrahigh-speed photography by Hirano.[32] The vibratory cycle is described as having

TABLE 47–3. Parameters in the Peripheral Process of the Production and Perception of Voice

	Parameters That Regulate Vibratory Pattern of Vocal Fold		*Parameters That Specify Vibratory Pattern*	*Parameters That Specify Sound Generated*	
Level	*Physiologic*	*Physical*	*Physical*	*Acoustic*	*Psychoacoustic*
Parameters	Neuromuscular control	(Primary)	Fundamental period	Fundamental frequency	Pitch
	Respiratory muscles	Expiratory force	Symmetry Periodicity	Amplitude (intensity)	Loudness
	Laryngeal muscles	Vocal fold Position Shape and size Elasticity Viscosity	Uniformity Glottal closure Amplitude Mucosal wave Speed of excursion	Waveform Acoustic spectrum Fluctuations	Quality Fluctuations
	Articulatory muscles	State of vocal tract (Secondary) Pressure drop across glottis Volume velocity Glottal impedance	Glottal area waveform		

Adapted from Hirano M. Clinical examination of voice. New York: Springer-Verlag; 1981.

three phases: opening, closing, and closed. By convention, the cycle is described as beginning with the vocal folds closed. Frames 1 to 5 (see Figure 47–6, B) represent the opening phase. During this phase, subglottic pressure increases, forcing the vocal folds apart from an inferior to superior direction until the glottis opens, letting air escape and thus releasing subglottic pressure. As the elastic recoil of the vocal folds forces them back together, the portion of the vocal fold that was the last to open (superior portion) is the last to close. Thus, the vocal folds close from inferior to superior (frames 6 to 8). The folds then remain closed until the subglottic pressure builds up enough to force them open again. Anatomically, the movement of the mucosal wave depends on the soft and compliant lamina propria and a healthy layered structure. In extracting complex details of glottal area waveforms from videostrobolaryngoscopic recordings using modern computers and software (Figure 47–10), Woo demonstrated male and female differences as well as influences of pitch and loudness.[39] Women commonly demonstrate small posterior glottal gaps that seem to have no pathologic acoustic significance, whereas their glottal cycles have shorter closed phases than those of men. Mean peak glottal area is smaller in women, not only because of their smaller anatomic size but also owing to the fact that their higher frequency of vibration limits the lateral excursion of the vocal folds, thereby reducing the amplitude of vibration. As intensity increases, the closing phase becomes shorter and the closed period becomes longer, suggesting that the closing patterns of vocal folds are an especially important feature in normal and pathologic phonation.

The coordination of higher cortical centers interacts with specific musculature in the vocal tract

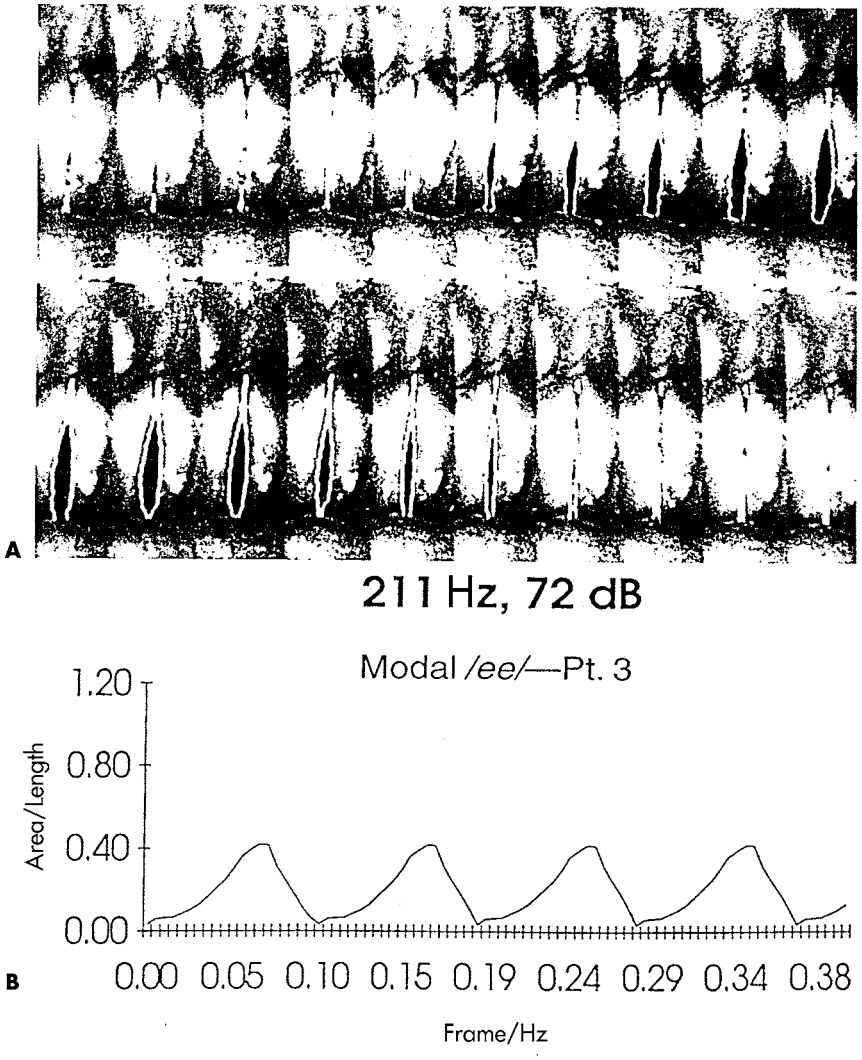


FIGURE 47–10. *A*, Twenty frames of a glottal cycle are digitized. The glottal margin has been traced by computer. *B*, Plot of the glottal area waveform (female, 211 Hz, 72 dB). Adapted from Woo P. Quantification of videostrobolaryngoscopic findings—measurements of the normal glottal cycle. Laryngoscope 1996;106 Suppl 79:1–27. Copyright Lippincott Williams & Wilkins.

to produce acoustic products recognized as speech. The laryngeal contribution to this product continues to be revealed in basic studies of phonation as an interactive process.

REFERENCES

1. Negus VE. The comparative anatomy and physiology of the larynx. London: Heinemann; 1949.
2. Maue WM, Dickinson DR. Cartilages and ligaments of the adult larynx. Arch Otolaryngol 1971;94:432–9.
3. Meiteles LZ, Lin PT, Wenk EJ. An anatomic study of the external laryngeal framework with surgical implications. Otolaryngol Head Neck Surg 1992; 106:235–40.
4. Hollinshead WH. Anatomy for Surgeons. Vol 1: the head and neck. 3rd ed. Philadelphia: JB Lippincott; 1982.
5. Spector GJ. Anatomy of the larynx. In: Ballenger JJ, editor. Diseases of the nose, throat, ear, head and neck. 13th ed. Philadelphia: Lea & Febiger; 1985.
6. Hast MH. Anatomy of the larynx. In: English GM, editor. Otolaryngology. Vol 3. Philadelphia: JB Lippincott; 1993.
7. Sasaki CT, Driscoll BP, Gracco C. Anatomy and physiology of the larynx. In: Ballenger JJ, Snow JB, editors. Otorhinolaryngology head and neck surgery. 15th ed. Baltimore: Williams & Wilkins; 1996.
8. Chamberlain WE, Young BR. Ossification (so-called "calcification") of normal laryngeal cartilages mistaken for foreign body. Am J Roentgen Rad Ther 1935; 33:441–50.
9. Hately W, Evison E, Samuel E. The pattern of ossification in the laryngeal cartilages: a radiological study. Br J Radiol 1965;38:585–91.
10. Kirchner JA. What have whole organ sections contributed to the treatment of laryngeal cancer. Ann Otol Rhinol Laryngol 1989;98:661–6.
11. Sanders I, Jacobs I, Wu BL, Biller HF. The three bellies of the canine posterior cricoarytenoid muscle: implications for understanding laryngeal function. Laryngoscope 1993;103:171–7.
12. Sanders I, Rao F, Biller HF. Arytenoid motion evoked by regional electrical stimulation of the canine posterior cricoarytenoid muscle. Laryngoscope 1994; 104:456–62.
13. Han Y, Wang J, Fischman DA, et al. Slow tonic muscle fibers in the thyroarytenoid muscles of human vocal folds: a possible specialization for speech. Anat Rec 1999;256:146–57.
14. Sanders I. The microanatomy of the vocal folds. In: Rubin JS, Sataloff RT, Korovin GS, Gould WJ, editors. Diagnosis and treatment of voice disorders. New York: Igaku-Shoin; 1995.
15. Gracco C, Kahane JC. Age-related changes in the vestibular folds of the human larynx: a histomorphometric study. J Voice 1989;3:204–12.
16. Holinger LD. Pharyngoceles, laryngoceles, and saccular cysts. In: English GM, editor. Otolaryngology. Vol 3. Philadelphia: JB Lippincott; 1993.
17. Hirano M. Phonosurgical anatomy of the larynx. In: Ford CN, Bless DM, editors. Phonosurgery: assessment and surgical management of voice disorders. New York: Raven Press; 1991.
18. Brunton TL, Cash T. The valvular action of the larynx. J Anat Physiol 1883;17:363.
19. Johner CH. The lymphatics of the larynx. Otolaryngol Clin North Am 1970;3:439–51.
20. Garrett DJ, Larson CR. Neurology of the laryngeal system. In: Ford CN, Bless DM, editors. Phonosurgery: assessment and surgical management of voice disorders. New York: Raven Press; 1991.
21. Sanders I, Mu L. Anatomy of human internal superior laryngeal nerve. Anat Rec 1998;252:646–56.
22. Anderson JW, Sant'Ambrogio FB, Mathew OP, Sant'Ambrogio G. Water-responsive laryngeal receptors in the dog are not specialized endings. Respir Physiol 1990;79:33–43.
23. Bradley RM. Sensory receptors of the larynx. Am J Med 2000;108 Suppl 4a:47–50.
24. Sasaki CT. Physiology of the larynx. In: English GM, editor. Otolaryngology. Vol 3. Philadelphia: JB Lippincott; 1993.
25. Sasaki CT, Suzuki M. The respiratory mechanism of aerosol inhalation in treatment of partial airway obstruction. Pediatrics 1977;59:689–94.
26. Sasaki CT, Suzuki M. Laryngeal reflexes in cat, dog and man. Arch Otolaryngol 1976;102:400–2.
27. Suzuki M, Sasaki CT. Laryngeal spasm: neurophysiologic redefinition. Ann Otol Rhinol Laryngol 1977; 86:150–7.
28. Suzuki M, Kirchner JA. The posterior cricoarytenoid as an inspiratory muscle. Ann Otol Rhinol Laryngol 1969;78:849–64.
29. Suzuki M, Kirchner JA, Murakami Y. The cricothyroid as a respiratory muscle. Its characteristics in bilateral recurrent laryngeal nerve paralysis. Ann Otol Rhinol Laryngol 1970;79:976–83.

30. Sasaki CT, Fukuda H, Kirchner JA. Laryngeal abductor activity in response to varying ventilatory resistance. Trans Am Acad Ophthalmol Otolaryngol 1973; 77:403–10.
31. Sasaki CT. Electrophysiology of the larynx. In: Blitzer A, Brin MF, Sasaki CT, et al, editors. Neurologic disorder of the larynx. New York: Thieme Medical Publishers; 1992.
32. Hirano M. Clinical examination of voice. New York: Springer-Verlag; 1981.
33. Gracco C. Vocal tract imaging. New York: Biomedical Communications; 1994.
34. Kempster GB, Larson CR, Distler MK. Effects of electrical stimulation of cricothyroid and thyroarytenoid muscles on voice fundamental frequency. J Voice 1988;2:221–9.
35. Buchthal F, Faaborg-Anderson K. Electromyography of laryngeal and respiratory muscles: correlation with phonation and respiration. Ann Otol Rhinol Laryngol 1964;73:118–23.
36. Hirano M, Koike Y, Joyner J. Style of phonation: an electromyographic investigation of some laryngeal muscles. Arch Otolaryngol 1969;89:902–7.
37. Gay T, Strome M, Hirose H, et al. Electromyography of the intrinsic laryngeal muscles during phonation. Ann Otol Rhinol Laryngol 1972;81:401–9.
38. Lovqvist A, McGarr N, Honda K. Laryngeal muscles and articulatory control. J Acoust Soc Am 1984;76: 951–4.
39. Woo P. Quantification of videostrobolaryngoscopic findings—measurements of the normal glottal cycle. Laryngoscope 1996;106 Suppl 79:1–27.

CHAPTER 48

Development of the Larynx

Ronan O'Rahilly, MD, DSc, Fabiola Müller, Dr. habil rer nat

Human prenatal life is divided into embryonic and fetal periods. The embryonic period, which comprises the first 8 weeks, is subdivided into 23 developmental stages, the details of which have been described elsewhere.[1] Prenatal age is postfertilizational. Postmenstrual weeks (which are not age) are useful clinically, but the term gestational age is meaningless.[2] A summary of the developing respiratory system in relation to staging was published by O'Rahilly and Boyden,[3] and many further details were provided by O'Rahilly and Müller.[4] Since the classic study of the larynx by Soulié and Bardier,[5] the development of this organ has been investigated by, among others, O'Rahilly and Tucker[6] and Müller et al,[7,8] and there has been an excellent study in the mouse.[9] Many references to the earlier literature can be found in these publications.

It is known that the hyoid bone is derived from the cartilaginous elements of pharyngeal arches 2 and 3, and the thyroid and other cartilages of the larynx are believed to be associated with the more caudally situated arches. The details of the developing larynx have been identified stage by stage. A summary of the development of the larynx is provided in Table 48–1.

EMBRYONIC PERIOD

In a certain sense, the larynx may be considered as a diverticulum of the future laryngopharynx, with which (except for a temporary embryonic closure) it communicates throughout life.

Three to Four Weeks

At 3 to 4 weeks, when the embryo is only about 3 mm long (stage 10), a median pharyngeal groove (internally) and ridge (externally) are discernible, and the groove includes the laryngotracheal sulcus. Within a few days (stage 12), the "lung bud" appears as a respiratory diverticulum that projects from the digestive tube into the mesenchyme ventral to the foregut.

Right and left lung buds are soon detectable (stage 13), and the trachea becomes visible in the more advanced specimens. The bifurcation point of the trachea then begins to descend.

The Tracheo-Esophageal Septum

The term tracheo-esophageal septum was used historically for a supposed fusion between medially growing internal folds in the wall of the foregut, which were believed to advance caudorostrally. Such a septum, however, does not exist.[10] In point of fact, as the respiratory primordium grows caudad, it becomes separated from the digestive primordium by tracheo-esophageal mesenchyme. The summit of this mesenchymal partition is the separation point between the trachea and the esophagus, and it remains at a constant vertebral level during the first 6 postfertilization weeks (Figure 48–1).

Four to Six Weeks At about 5 weeks (stages 14 and 15), the future larynx begins its differentiation. The lateral epithelial walls become apposed in the median plane, thereby forming what is known as the epithelial lamina. This bilaminar plate is situated between the arytenoid swellings, behind which the pharyngeal lumen still communicates with the trachea (pharyngoinfraglottic duct or canal). The hypopharyngeal (formerly called hypobranchial) eminence does not represent the epiglottis, which is identifiable a little later.

The hyoid condensation begins to appear, and by 6 weeks, the cricoid can be identified in mesenchyme (Figure 48–2). This is followed by the appearance of a single cartilaginous cricoid center.[11]

The larynx is clearly definable by 6 weeks (stage 17) (see Figure 48–2, B). At this time, or a few days earlier, the front portion of the epithelial lam-

TABLE 48–1. Initial Appearance of Features of the Larynx and Related Structures

Feature	mm	2	3	4	5	6	8	10	15	20	25	30
	Weeks	4			5				6		7	8
	Stage	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Foregut appears																
Laryngotracheal keel and sulcus (respiratory primordium)																
Pharyngeal arches, clefts, and pouches																
Lung bud and primary bronchi																
Sagittal cleft of larynx																
Arytenoid swellings; epithelial lamina of larynx; trachea; superior laryngeal nerve																
Pharyngoinfraglottic duct connects digestive and respiratory tubes																
Epithelial lamina is complete; coronal cleft (not vestibule) of larynx identifiable																
Hyoid and cricoid condensations																
Skeleton of pharyngeal arches and larynx is dense mesenchyme; infraglottic cavity; ansa cervicalis and recurrent laryngeal nerve																
Hyoid chondrification begins; cricothyroid, posterior cricoarytenoid, arytenoid, and "thyrocricoarytenoid" muscles																
Epithelial lamina closes respiratory canal from pharynx																
Mesenchymal epiglottis																
Thyroarytenoid and lateral cricoarytenoid muscles																
Solid primordium of ventricle of larynx begins to appear; epithelial lamina begins to disintegrate																
Perioral reflex responses															?	
Aryepiglottic and thyroepiglottic muscles; ventricles that are angulated to the coronal cleft; intraepithelial receptors in laryngopharynx and nasal epithelium																
Epiglottis and palate are not yet in contact; vestibule becomes connected to infraglottic cavity; right and left leaves of epithelial lamina begin to separate; infrahyoid and most laryngeal muscles are present, eg, cricothyroid (oblique and straight parts); adult pattern of motor innervation, sensory incomplete																

ina can be regarded as its vestibular part. Lateral expansions of this portion soon form the embryonic vestibule (coronal or transverse cleft) (see Figure 48–2, B). The laryngeal cavity is now T-shaped in transverse section, but the transverse cleft that constitutes the embryonic vestibule corresponds to only a portion of the adult vestibule, which also includes the median cleft.

The appearance of the epithelial lamina at 6 weeks (stage 18) resembles type 1 congenital atresia, which may be essentially a persistence of the embryonic condition.[4] The rostral end of the tracheal

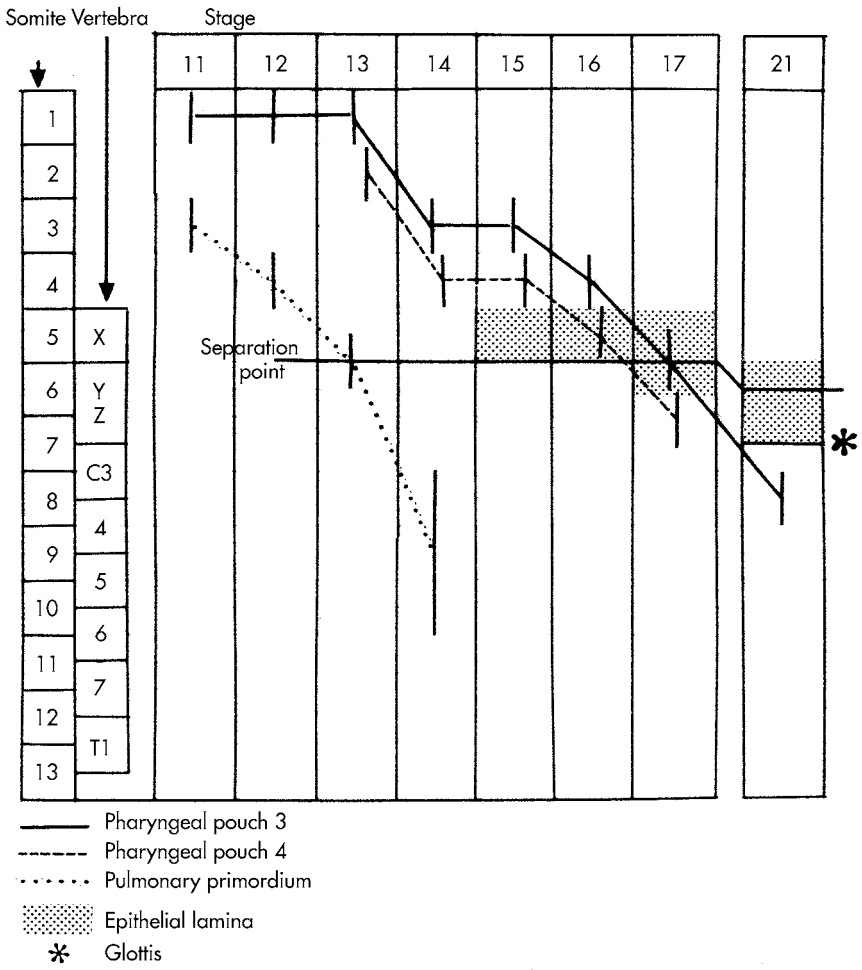


FIGURE 48–1. Graph showing the descent of pharyngeal pouches 3 and 4, and of the respiratory primordium from 4 to 7 postfertilizational weeks with reference to the only reliable landmarks, the somites and the vertebrae. The epithelial lamina and the separation point between trachea and esophagus remain at a constant level during the first 6 weeks. X, Y, Z are portions of the future axis. Based on the authors' reconstructions.

lumen is believed to represent the future infraglottic cavity. A tendency for the vestibule and the trachea to communicate with each other (vestibuloinfraglottic duct or canal) is characteristic. Laryngeal muscles are beginning to develop, but acceptable evidence for the existence of one or two common sphincters initially is lacking.[5,7] The hyoid condensation is undergoing chondrification, and condensation for the thyroid laminae begins to appear.

Importance of Vertebral Levels

In the assessment of levels, the only reliable landmarks are first the somites and then, also postnatally, the vertebrae. Soft tissues, such as pharyngeal pouches and spinal ganglia, are unreliable because they shift in position during development. Pharyngeal pouch 4, for example, used by several authors as a landmark,[10–12] is initially at the level of somite 2 (future basioccipital) but changes to somites 6 and 7 (base of the axis) by 6 weeks (see Figure 48–1). The pouch is then caudal to the epithelial lamina and does not represent the border between the pharynx and the larynx, as claimed.[11] Indeed, it is as low as the level of the future cricoid cartilage, and there is no room for a fifth pharyngeal pouch, supposed to correspond with the laryngeal ventricle.[10] In contrast, what we have termed the separation point between the trachea and esophagus, that is, the summit of the tracheoesophageal mesenchyme,[4] is remarkably stable (see Figure 48–1), which is said to be caused by its catenoid structure.[13]

Seven to Eight Weeks The pharyngoinfraglottic duct is largely, if not entirely, open in most embryos, and the track of the vestibuloinfraglottic duct is visible (see Figure 48–2, C). The ventricles of the larynx begin to appear as right and left laterally projecting epithelial buds. The right and left leaves of the epithelial lamina are beginning to separate, but at

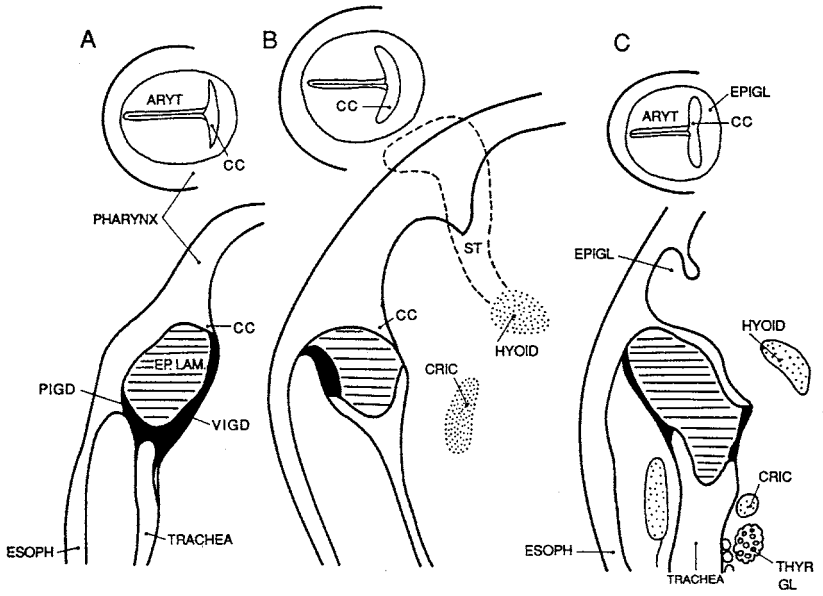


FIGURE 48–2. Development of the epithelial lamina and the first skeletal elements of the larynx in stages 16 to 21 (5½ to 7 weeks). The *upper row* shows the larynx as seen from above, and the *lower row* provides median sections at the corresponding stages (16, 17, and 21). ARYT = arytenoid swelling; CC = coronal cleft; CRIC = cricoid condensation; EPIGL = epiglottis; ESOPH = esophagus; PIGD = pharyngoinfraglottic duct; THYR GL = thyroid gland; VIGD = vestibuloinfraglottic duct.

the end of the embryonic period (stage 23), areas of fusion generally persist rostrally and caudally. The pharyngeal lumen, however, is continuous with the infraglottic cavity and hence with the trachea. Cavitation is beginning in the ventricles.

By the end of the embryonic period (stage 23; Figure 48–3), the hyoid cartilage consists of the body and greater horns, the lesser horns being distinct nodules separated from the styloid processes. The cartilaginous thyroid laminae, which may show a foramen, are united below by mesenchyme ("copula"). The superior horns may or may not be continuous with the laminae, but they arise in common with the greater horns of the hyoid. The cricoid cartilage is a continuous ring that comprises an arch and a lamina. Each arytenoid cartilage possesses a cartilaginous muscular process and a mesenchymal vocal process.

Most of the major laryngeal muscles (cricothyroid, posterior and lateral cricoarytenoids, thyroarytenoid, transverse arytenoid) are now present (although not yet striated), and their innervation follows the adult pattern closely.[7] Motor fibers penetrate the muscles. The vocalis is beginning to differentiate.

The laryngeal cavity now comprises (as in the adult) the vestibule, the ventricles and the part between them, and the infraglottic cavity. The ventricles are not at the level of the future glottis, which lies more caudally. The sensory innervation is well established, although most fibers do not yet reach the epithelium. In the laryngopharynx, however, receptors are present.[8]

FETAL PERIOD

Much information concerning the fetal larynx is available from Soulié and Bardier.[5]

During trimester 1 (Figure 48–4, A), the larynx increases in size from about 3 to 7 mm. The thyroid laminae become fused, and the cricothyroid membrane develops. The vocal ligaments begin to form early (33 mm), and the ventricular ligaments and glottis become increasingly apparent. The ventricles acquire their saccules. The cricoarytenoid joints undergo cavitation, in which they are followed by the cricothyroid joints. By 90 mm, the laryngeal cavity has its adult form.[5]

During trimester 2 (see Figure 48–4, C), the larynx increases in size from about 8 to 15 mm. The thyroid cartilage becomes a single structure. The epiglottis begins to chondrify. The hyoid cartilage commences ossification. The corniculate and cuneiform cartilages develop. The joints acquire ligaments. Glands become well developed. Elastic fibers have been detected in the epiglottis.[14]

During trimester 3, the larynx increases in size to about 15 to 20 mm, and the inlet is about 5 mm in diameter. Laryngeal connective tissue compartments in trimesters 2 and 3 have been studied.[15]

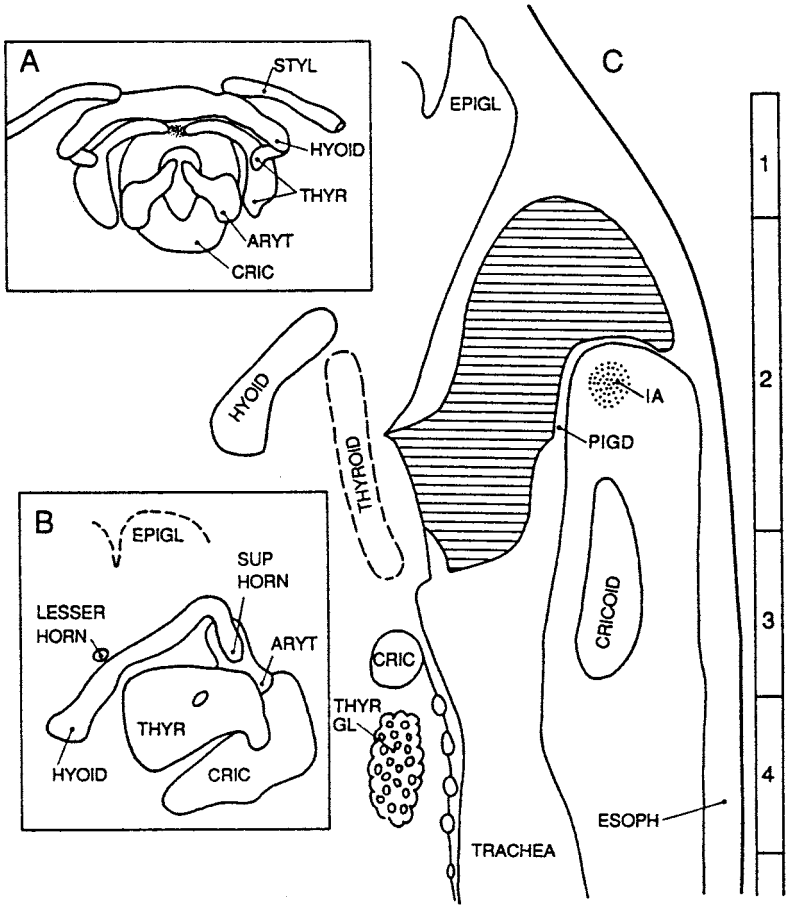


FIGURE 48–3. Larynx at the end of the embryonic period (8 weeks), showing reconstructions of three embryos at stage 23. *A*, Skeletal elements from above. *B*, Lateral view of skeleton. Note the foramen in the thyroid cartilage, through which nerve fibers of the external laryngeal nerve and an artery may enter; the opening can be found as a variation in the adult. The hyoid is remarkably ventral with regard to the thyroid cartilage. *C*, Median section. The numbers 1 to 4 indicate vertebral levels. STYL = styloid process; ARYT = arytenoid cartilage; CRIC = cricoid cartilage; EPIGL = epiglottis; ESOPH = esophagus; IA = interarytenoid muscle; PIGD = pharyngoinfraglottic duct; THYR = thyroid cartilage; THYR GL = thyroid gland.

Measurements of the developing larynx are available: embryonic and fetal,[5] fetal and infantile,[16] infantile,[17] infantile and child's,[18] and prepubertal and pubertal.[19] The larynx, as indicated by the tip of the epiglottis, the hyoid, and the lower border of the cricoid, descends from embryonic and fetal life to adulthood (Figure 48–5). Vertebral levels are available: embryonic,[4] fetal,[20] fetal and child's,[21] and child's.[22] The infantile larynx has been studied,[23,24] and detailed views have been published.[25]

Ossification in the thyroid cartilage frequently begins from about 20 years onward in the posterior

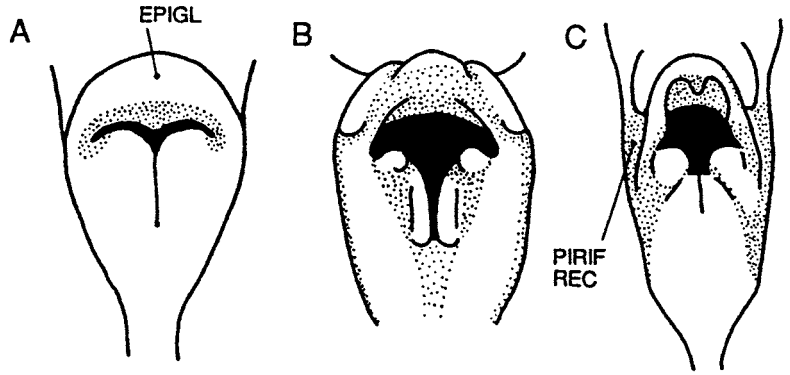


FIGURE 48–4. Dorsal view of the prenatal larynx at (*A*) 32 mm (8 weeks), (*B*) 55 mm (10 weeks), and (*C*) 160 mm (18 weeks). Note the appearance of the corniculate tubercles, the aryepiglottic folds, and the piriform recess (PIRIF REC). Adapted from Soulié A and Bardier F.[5]

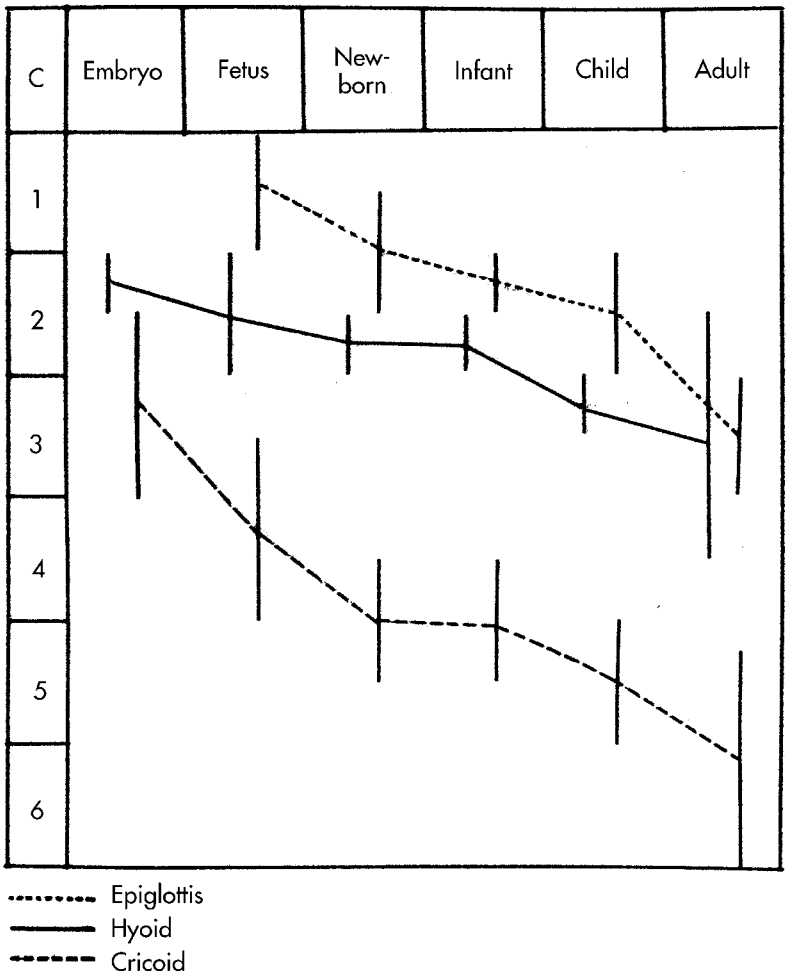


FIGURE 48–5. Graph showing the prenatal and postnatal descent of the hyoid, cricoid, and epiglottis with reference to cervical vertebrae 1 to 6.

border of each lamina. It would be imprudent, however, to attempt to estimate age from the extent of the ossific process.[26]

Function begins in the fetal larynx early in trimester 2, when the larynx opens for the inhalation of amniotic fluid into the lungs, and the diaphragm descends. Phonatory vibrations commence later in the trimester.[27]

CONGENITAL ANOMALIES

Only some congenital anomalies of the larynx are currently amenable to interpretation in developmental terms.

Some laryngeal anomalies are part of a more widespread disturbance. Thus, congenital webs may be accompanied by subglottic stenosis, and structures far removed are affected in some instances, such as anomalies of the limbs. Similarly, cleft larynx may be associated with a congenital cardiac defect or cleft lip and palate and has been considered a median developmental field defect that can occur in causally different syndromes.

Laryngomalacia, said to be the most frequent congenital laryngeal anomaly, produces partial obstruction of the supraglottic airway, but its cause is unclear and cannot be merely immaturity and alterations of the cartilaginous and muscular components.[28]

Laryngeal clefts, which may extend to the trachea and even to the carina, are rare.[29] Interarytenoid clefts develop as a consequence of lack of development of the interarytenoid and aryepiglottic muscles.[30] Cricoid clefts, which may be partial or total, may arise from incomplete dorsal fusion of the limbs of the single ventral chondrific center.[30] Com-

plete clefts, those that involve the trachea, probably develop during separation of the trachea and esophagus at 4 to 5 weeks (stage 13), as suggested by the frequent association with tracheoesophageal fistulae.

Congenital laryngeal atresia probably arises between 6 and 10 postovulatory weeks, the more extensive examples representing embryonic stages and the less severe forms being fetal. They are supraglottic rather than infraglottic.[30] Complete atresia, which is probably the least frequent of congenital laryngeal anomalies, is associated with polyhydramnios.[28] It resembles closely the normal appearance of the larynx at 6 weeks (stage 18) and may be basically a persistence of the embryonic condition when the epithelial lamina has reached its full expansion (see Figure 48–2, C). Incomplete atresia arises later; hence, the vestibule has already developed to a variable degree. Glottic atresia, the most frequent type, probably arises early during the fetal period, when the epithelial lamina has failed to undergo further delamination.

Congenital laryngeal webs, situated ventrally and extending dorsally to a variable degree, are also thought to result probably from a localized failure of splitting of the epithelial lamina into two walls, most likely early in the fetal period. A special instance is a web at the vocal folds, namely at the lower border of the epithelial lamina. It presents, early in the fetal period, as a perforated membrane that obstructs the glottic opening.[30]

Infraglottic (subglottic) stenosis is not associated with the persistence of the epithelial lamina, but some instances seem to be produced by deformities of the cricoid cartilage. Cartilaginous and membranous types would differ in origin.

Congenital laryngeal cysts may be ventricular or infraglottic. They are usually in the region of the saccule but do not communicate with the laryngeal lumen. They probably indicate disturbed development in fetal life.

Nonrecurrent laryngeal nerve is associated with an abnormal origin of the right subclavian artery, usually from the dorsal aspect of the arch of the aorta. The condition is believed to arise from an embryonic failure of the right aortic arch 4, allowing the nerve to proceed directly to the larynx.

REFERENCES

1. O'Rahilly R, Müller F. Developmental stages in human embryos, including a revision of Streeter's "Horizons" and a survey of the Carnegie Collection. Washington (DC): Carnegie Institution of Washington; 1987.
2. O'Rahilly R, Müller F. Prenatal ages and stages—measures and errors. Teratology 2000;61:382–4.
3. O'Rahilly R, Boyden EA. The timing and sequence of events in the development of the human respiratory system during the embryonic period proper. Z Anat Entwicklungsgesch 1973;141:237–50.
4. O'Rahilly R, Müller F. Respiratory and alimentary relations in staged human embryos. New embryological data and congenital anomalies. Ann Otol Rhinol Laryngol 1984;93:421–9.
5. Soulié A, Bardier F. Recherches sur le développement du larynx chez l'homme. J Anat Physiol 1907;43: 137–240.
6. O'Rahilly R, Tucker JA. The early development of the larynx in staged human embryos. I. Embyros of the first five weeks (to stage 15). Ann Otol Rhinol Laryngol 1973;82:1–27.
7. Müller F, O'Rahilly R, Tucker JA. The human larynx at the end of the embryonic period proper. I. The laryngeal and infrahyoid muscles and their innervation. Acta Otolaryngol (Stockh) 1981;91:323–36.
8. Müller F, O'Rahilly R, Tucker JA. The human larynx at the end of the embryonic period proper. 2. The laryngeal cavity and the innervation of its lining. Ann Otol Rhinol Laryngol 1985;94:607–17.
9. Henick DH. Three-dimensional analysis of murine laryngeal development. Ann Otol Rhinol Laryngol 1993;102 Suppl 159:3–24.
10. Zaw-Tun HA. The tracheo-esophageal septum—fact or fantasy? Origin and development of the respiratory primordium and esophagus. Acta Anat (Basel) 1982;114:1–21.
11. Zaw-Tun HA, Burdi AR. Reexamination of the origin and early development of the human larynx. Acta Anat (Basel) 1985;122:163–84.
12. Sanudo JR, Domenech-Mateu JM. The laryngeal primordium and epithelial lamina. A new interpretation. J Anat 1990;171:207–22.
13. Sutliff K, Hutchins GM. Septation of the respiratory and digestive tracts in human embryos: crucial role of the tracheoesophageal sulcus. Anat Rec 1994; 238:237–47.
14. Patzelt V. Über die menschliche Epiglottis und die Entwicklung des Epithela in den Nachbargebieten. Z Anat Entwicklungsgesch 1924;70:1–178.
15. Tucker GF, Smith HR. A histological demonstration of the development of laryngeal connective tissue compartments. Trans Am Acad Ophthalmol Otolaryngol 1962;66:308–18.

16. Schild JA. Relationship of laryngeal dimensions to body size and gestational age in premature neonates and small infants. Laryngoscope 1984;94:1284–92.
17. Eckenhoff JF. Some anatomic considerations of the infant larynx influencing endotracheal anesthesia. Anesthesiology 1951;12:401–40.
18. Peter K, Wetzel G, Heiderich F, eds. Handbuch der Anatomie des Kindes. Munich: Bergmann; 1938.
19. Kahane JC. A morphological study of the human prepubertal and pubertal larynx. Am J Anat 1978; 151:11–9.
20. Magriples U, Laitman JT. Developmental change in the position of the fetal human larynx. Am J Phys Anthropol 1987;72:463–72.
21. Noback GJ. The developmental topography of the larynx, trachea and lungs in the fetus, new-born, infant and child. Am J Dis Child 1923;26:515–33.
22. Roche AF, Barkla DH. The level of the larynx during childhood. Ann Otol Rhinol Laryngol 1965;74: 645–54.
23. Tucker G. The infant larynx: direct laryngoscopic observation. JAMA 1932;99:1899–902.
24. Wilson TG. Some observations on the anatomy of the infantile larynx. Acta Otolaryngol (Stockh) 1953;43:95–9.
25. Bosma JF. Anatomy of the infant head. Baltimore: Johns Hopkins University Press; 1986.
26. Keen JA, Wainwright J. Ossification of the thyroid, cricoid and arytenoid cartilages. S Afr J Lab Clin Med 1958;4:83.
27. Ramón y Cajal CL. Description of human fetal laryngeal functions: phonation. Early Hum Dev 1996;45:63.
28. Willging JP, Cotton RT. Congenital anomalies of the larynx. In: Tewfik TL, der Kaloustian VM, editors. Congenital anomalies of the ear, nose, and throat. New York: Oxford University Press; 1997. p. 383–91.
29. Benjamin B, Inglis A. Minor congenital laryngeal clefts: diagnosis and classification. Ann Otol Rhinol Laryngol 1989;98:417–20.
30. Zaw-Tun HA. Development of congenital laryngeal atresias and clefts. Ann Otol Rhinol Laryngol 1988; 97:353.

CHAPTER 49

Disorders of Voice, Speech, and Language

Fred D. Minifie, PhD

Disorders of voice, speech, and language comprise an important category of problems faced by the physician because treatment of these disorders can have a large influence on vocational, social, and emotional adjustments of the patient. Many otolaryngologists are properly reluctant to provide rehabilitation for communicative disorders beyond the treatment of disease because they believe they are not qualified. Patients, however, inevitably seek answers to questions about communication disorders and, while attempting to be helpful, the physician often finds it necessary to play the role of a part-time speech-language pathologist and counselor. One of the purposes of this chapter is to provide current information about disorders of voice, speech, and language and their remediation.

This chapter is organized around three major categories of communication disorders: disorders of voice, speech, and language. Each category will be discussed separately. Communication disorders stemming from hearing impairments are discussed in Chapter 5.

VOICE DISORDERS

Normal Voice Production

Voiced sound during speech production is called phonation. Depending on the length, tension, and thickness adjustments of the vocal folds, the tone produced by normal vibrating vocal folds can vary in frequency (perceived as variations of the pitch of the voice), intensity (loudness), duration, or quality. Such parameters contribute to changes in intonation and stress during speech production and perception.

Pitch The pitch of sound produced during normal phonation is directly related to the frequency at which air pulses are released through the glottic aperture owing to vocal fold vibration, a rate determined primarily by the mass and elasticity of the vocal folds in relation to their length. Large folds produce a lower pitch than smaller folds because the greater mass causes them to vibrate more slowly. This is why adults have lower vocal pitches than do infants and young children; vocal pitch is inversely related to the size of the vocal folds. However, each talker also has the ability to change the frequency of vocal fold vibration. When the folds in any normal larynx are shortened, their cross-section becomes larger, and their myoelasticity is reduced. Such folds tend to move more slowly. Conversely, when the vocal folds are elongated, they become thinner, and their elasticity is increased. As a result, when the air stream pushes tightened and thinned vocal folds apart, they return to their approximated positions more quickly. This reaction results in increased frequency of vocal fold vibration (higher pitch). Frequency of vibration may also be altered by isotonic tension in the muscles of the vocal folds. This type of contraction increases the stiffness of the folds and consequently causes them to vibrate more rapidly.

Loudness Loudness of the voice is related to the sound pressure created by the released pulsations of air through the glottis. Sound pressure is directly proportional to the volume-velocity of airflow in the glottis. By increasing the thickness and stiffness of the vocal folds, increased resistance can be offered to the flow of air from the lung as the folds remain closed for a longer percentage of time during each cycle of vibration. The longer closed period allows greater subglottal air pressure to build up to overcome the increased resistance offered by the folds. With increased subglottal driving pressure on opening of the glottis, increased volume velocities of air-

flow occur. Moreover, because increased stiffness and thickness of the vocal folds increase the elasticity of the vocalis or internal thyroarytenoid muscles, they snap back to the midline with increased velocity. The rapidly changing volume-velocity of the escaping train of pressure pulses increases the amplitude of sound waves generated, causing greater excursions of the tympanic membrane in the ear of a listener and a sensation of louder sound.

Quality Both vocal fold vibration and resonance determine the quality of voice. The phonatory aspect includes the manner of air pulse release at the glottis. Opening, closing, and closed phases of the glottal cycle can vary in relation to each other. For example, laryngeal lesions can influence the pattern of movement of the vocal folds, independently and in combination (to be discussed later in this chapter). These factors can influence the number and relative intensity of harmonic partials composing the complex vocal sound and, consequently, the quality of the voice. Modification of sound as it travels through the pharynx, mouth, and nose (resonance) results from selective emphasis and damping of the overtones and other partials in the complex sound generated in the larynx.

The contribution of voice to a normal communication process is demonstrated succinctly by its absence; speech without phonation is whispering. In whispered speech, the vocal folds do not vibrate, and the laryngeal sound is aphonic. When this condition is present, the speaker can produce speech and language but has no vocal pitch in the musical sense. The sound is a relatively weak noise composed of breath sounds produced when the air becomes turbulent as it rushes past the irregularities that jut from the walls of the respiratory tract (particularly at the true and false vocal folds). This turbulent air noise constitutes a sound source that is acted on by the resonances of the mouth, pharynx, and nose much as resonance occurs when vocal sound arises at the glottis. When the vocal folds vibrate, there is vocal sound that has pitch and many of the distinguishing vocal qualities that identify individual voices, whereas in whisper they are absent.

Causes of Voice Disorders

Voice disorders always have causes. Something must be abnormal or atypical in the way in which the vocal folds function to produce disordered voice. A clinical appreciation of vocal fold vibration can begin with the concept of an ideal larynx. In such an organ, the two vocal folds would have the same dimensions, they would move symmetrically and regularly, and each vibratory cycle would include three phases: glottal opening, glottal closing, and closed glottis. The vocal sound from this ideal larynx would be judged "excellent"; it would be smooth and free of all hoarseness, have an appropriate pitch range for the age and gender of the talker, and would be capable of wide pitch and loudness variation. When vibration of the vocal folds varies from the ideal, the characteristics of the voice also vary from the ideal—the extent of vocal difference depending on the type and amount of alteration in the vibratory pattern.

The assumption that no two faces are exactly alike is commonly accepted. Observation of living and dissected larynges leads to a parallel assumption that no two larynges are exactly alike. Each larynx reflects its uniqueness in its behavioral physiology, yet most larynges are capable of producing a normal voice that is judged to be within the accepted ranges of pitch, loudness, and quality found in a majority of individuals of the same age and sex.

Abnormal vocal fold vibration takes many forms, most of them not visible to the unaided eye. Ultra–slow-motion films of abnormal voices reveal that one fold may move faster than the other, one may have a greater lateral excursion, vibration may be limited to one vocal fold, there may be no glottal closure or incomplete glottal closure, closure may occur at a paramedian position, vibratory patterns may be dissimilar at different regions along one or both folds, and the vibratory periods and amplitudes may be randomly variable in consecutive glottal openings.[1] Acoustic indicators of such patterns may include pitch period perturbations (jitter), amplitude perturbations (shimmer), decreased signal-to-noise ratio (more glottal noise), altered fundamental frequency (F0), increased standard deviations of F0, and many others. The potential complexity of vibratory patterns resulting from combinations of these cyclic abnormalities and sequential irregularities is almost endless.

Pitch Disorders Pitch disorders are present when the voice is consistently higher or lower (in relation

to the musical scale) than would be expected for an individual of a given gender and age or when the sound is tremulous, monotonous, or bizarre (eg, when the pitch patterns do not convey the ideas being expressed). An involuntary, high-pitched voice in males does not necessarily interfere with communication, and if a woman produced it, it would probably not be considered unpleasant. A similar example could be provided for women with very low-pitched voices. The "voice disorder," therefore, is not the sound itself, which may be quite appropriate for the size and shape of the larynx that produced it, but the inappropriateness of the sound to the circumstances (the age and gender of the talker).

Functional Pitch Problems. In contrast to the foregoing example, high-pitched male voices frequently have functional origins. Verdolini described mutational falsetto as "the continued use of a high pitched (falsetto) voice by a postpubertal male, without any apparent physical basis. For some patients, psychological conflict is implied as a basis for the disorder. In other patients, the disorder may simply reflect a learned pattern."[2] To tell a person with this type of voice that he has nothing wrong with his larynx and that he should go home and speak normally may reassure him, but it does not help him, and usually it serves only to increase his frustration. He is already speaking "normally" and needs to be taught how to phonate in a way that is "abnormal" for him. Most of these patients respond readily to voice therapy and adopt the new voice after a brief period of self-consciousness. Those with normal structure who voluntarily maintain the high pitch reflect an abnormal self-image and should be referred to a psychiatrist or clinical psychologist.

Another form of functional pitch problem is ventricular phonation wherein the talker produces a low-pitched, gravely voice through vibration of the false vocal folds. This type of phonation most often develops as a compensatory mechanism.

Organic High-Pitch Problems. Although most abnormally high-pitched voices are probably caused by functional problems, cases with organic causes are not uncommon and may be classified into four categories: underdeveloped larynx, laryngeal web, structural asymmetry, and swelling in the anterior commissure.

Underdeveloped Larynx. The underdeveloped larynx has small vocal folds that vibrate more rapidly than larger folds and consequently create a high pitch. The small larynx may accompany a general structural retardation or may be part of a hereditary familial body size. The clinician should be particularly sensitive to genetic syndrome effects from various forms of dwarfism, short neck, an abnormality associated with syndromes of short stature, and other possible syndromic effects when an underdeveloped larynx is observed.[3] The larynx may also fail to develop as a result of hormonal imbalance, which is usually revealed in simultaneous retardation of other secondary sex characteristics.

Laryngeal Web. Another organic cause of high-pitched voice is a laryngeal web, which may be congenital or cicatricial. When this structure is small enough not to interfere with breathing or when it does not contribute to stridor, it may go undetected until the voice is observed as abnormal. A web need not extend far along the borders of the vocal folds to create a voice problem. Its effect is to shorten the free portions of the folds and thereby produce faster vibration and a resulting higher pitch. The vocal tone in children with webs resembles falsetto and tends to be weak. Hoarseness may be present also, accompanied by the audible evidence of vocal strain, particularly with men who have attempted to force a lower pitch.

Structural Asymmetry. A third organic cause for chronic high pitch is abnormal approximation of the vocal folds. In this condition, a structural asymmetry may cause the vocal process of one arytenoid cartilage to slide on top of its opposite member in such a manner that posterior parts of the membranous folds are pressed together, thereby effectively shortening their vibrating portions. The adjustment approximates that which is often observed in the male larynx when a high-pitched, falsetto voice is produced.

Swelling of the Anterior Commissure. A fourth condition, and one that may be overlooked or ignored as a cause of excessively high pitch, is enlargement of the glottal margin of one or both vocal folds adjacent to the anterior commissure. A localized protrusion no more than l mm high and 2 mm long on the glottal margin is sufficient to shorten the vibrating length of the folds and to produce a higher than normal pitch. This lesion can be

seen only when the anterior commissure is visualized, while the folds are widely abducted. During adduction, the protruding areas are compressed into the underlying tissue and create the illusion of a prominent commissural attachment without significance. The location of the offending lesion at the anterior attachment where the folds converge enables it to modify the vibratory pattern more extensively than it would if it were located more posteriorly. When a protruding mass is located where it is not pinched between the vocal folds, it usually does not shorten the vibrators and consequently does not raise vocal pitch; instead, it often causes a lower pitch accompanied by hoarseness or breathiness.

Organic Low-Pitch Problems. Defects of vocal pitch can also encompass the abnormally low voice. Excessively low pitch in both men and women is usually associated with organic change; however, functional disorders are observed in persons who attempt to speak at a pitch below that which is optimum for the structures involved. The most common organic origins of low-pitched voices are Reinke's edema, virilization, glottalization or "vocal fry," and tremulousness.

Reinke's Edema. Verdolini suggested that

> Reinke's edema involves the widespread accumulation of fluid or swelling, usually in both vocal folds, in the layer of tissue right below the vocal fold epithelium. Although the causes for this edema are not entirely clear, some clinicians think that chronic exposure to irritants such as smoke may play a role in the development of the condition. The most obvious effect of Reinke's edema on voice typically is an extremely low speaking pitch (due to the increased mass of the folds) and loss of the ability to produce high notes. The voice is also described as raspy.[2]

Virilization. An abnormally low pitch or masculine type of voice in a woman is as distressing to the woman possessing it as high pitch is in the male, and for the same kinds of psychosocial reasons. Virilization of the female voice after hormone therapy in young women has been reported with increasing frequency.[5,6] Recovery of normal pitch in these patients does not occur with interruption of medication, and at present, no specific counteragent seems to exist. The larynges of women with virilized voices appear to be normal in all aspects except perhaps for some general enlargement of the vocal folds toward their glottal margins. The effects of compensatory vocal adjustments on size and condition of the folds are not known.

Glottalization or "Vocal Fry." A form of excessively low pitch characterized by a low-frequency popping or ticking sound is referred to as vocal fry or glottalization. It is produced normally when only a little breath pressure is exerted against the vocal folds, such as during the ending of a phrase. This low-pitched voice is adopted occasionally by adolescent boys and young men in an effort to sound "more masculine." Although glottalization is often functional, it also has organic causes. Edematous vocal folds from whatever cause are frequently associated with this voice. Other alterations of the vocal folds, such as dry mucosa following irradiation therapy, or absence of a flexible covering membrane, may be observed.

Tremulousness. Vocal pitch deviation such as persistent tremulousness and its opposite, extreme monotony, are rare in young persons and are usually considered psychogenic. In older individuals, however, these vocal characteristics may be symptomatic of deterioration or injury to the nervous system.[7] In instances of brainstem involvement, such vocal changes may appear among the early symptoms and consequently deserve careful consideration in diagnosis.[8,9] On the other hand, when either tremulousness or monotony develops as a secondary symptom, other impairments are usually so prominent that the voice, as an entity, is relatively unimportant as a priority in therapy.[10,11] One of the most common diseases of the extrapyramidal system is Parkinson's disease, in which chemical deficiencies (specifically a lack of dopamine) produce a generalized movement disorder characterized by small movements (hypokinesia) and slow movements (bradykinesia) that impose low-frequency variations on voice production.[2,12]

Loudness Disorders Disorders of loudness may also be classified into categories that parallel those used with pitch deviations; some are functional and some are organic. The voice may be too loud or not loud enough in relation to the place and circumstances, or there may be loudness variations that are not appropriate for conveying the meaning of the utterance. Moreover, any generalized swelling of the

glottal margins that improves approximation without stiffening the folds contributes to glottal closure and sound generation. Patients whose voices become louder when they have laryngitis or mild edema illustrate this fact.

Functional Loudness Problems. Personal Adjustment. Atypical loudness is often an indicator of personality aberrations (eg, overly aggressive, shy, or socially insecure persons). Voices may be too loud or too soft depending on the personality. Fortunately, these adjustment problems are amenable to counseling procedures by a speech-language pathologist or clinical psychologist. Voice improvement is usually enhanced by direct training of voice production along with counseling.

Environmental Stress. Some persons are required to speak loudly in their occupations. This vocal requirement often creates laryngeal trauma, subsequent changes in the vocal organs, and consequent voice disorders. Excessively loud speaking and loud cheering at sports events are recognized as principal types of vocal abuse. These causes create difficult problems to manage when the basic occupational conditions persist. Voices can be retrained, however, if the patient is willing to learn how to use the voice. An adjunct to voice training and a substitute when circumstances are not conducive to such therapy is the use of a speech aid. Small, portable amplifying systems consisting of a clip-on microphone, amplifier, and loudspeaker are available and, when used properly, can provide adequate loudness of voice in noisy environments. The focal problem with the occupational voice disorder, when work requires daily vocal abuse, is the lack of opportunity to recover from traumatic effects and to institute a period of training to prepare the voice for unusual demands. Such patients benefit from periods of complete vocal rest (silence) prior to participating in vocal retraining therapy. Ideally, it is helpful to the professional speaker as well as to workers in noisy environments to acquire "big voices" during therapy. Fortunately, amplification can satisfy the job needs and reduce the vocal abuse for some patients.

Organic Loudness Problems. Paralysis or Paresis. When a slowness, weakness, or absence of complete glottal closure occurs, the voice is weak and breathy. The basic causes are as follows: impaired neural control from cerebral cortex lesions (stroke) or peripheral factors such as local trauma (including surgery), viruses, and heart disease; myasthenia gravis; arthritis; or general debility from anemia or other diseases. In the larynx, peripheral paralyses and pareses usually affect only one vocal fold. The unaffected vocal fold usually compensates to some degree by crossing the midline of the glottal space to achieve closure during phonation.[2] When the disease and organic conditions permit, direct therapy for voice improvement often produces good results. Suggestions about therapy are presented later in this chapter.

Bowed Vocal Folds. This muscular deformation stems from long-term heavy use of the voice, particularly in elderly patients. Patients produce a weak voice because of an impaired ability to develop closure of the vocal folds during voicing.

Sulcus Vocalis. A groove parallel to the vocal fold margin may be present in one or both vocal folds as a result of congenital or possibly developmental causes. Although this condition is not well understood because data are limited, the weak voice may result from an inability to achieve adduction of the vocal fold along its entire length.[2]

Hearing Impairment. When a loss of hearing is sufficient to cause an individual to speak more loudly than circumstances warrant, a hearing problem is usually evident, and the offending vocal symptoms (similar to those described above under the topic of functional loudness problems) are approached through treatment of the hearing loss.

Voice Quality Disorders

Voice quality disorders encompass resonance and phonatory components. The phonatory deviations can be placed along an auditory continuum extending from aphonia to hoarseness, with such intervening and intermixing qualities as breathiness and harshness. In this chapter, phonatory disorders are presented under five headings: aphonia, breathiness, harshness, hoarseness, and spasmodic dysphonia. Authors reporting in the medical literature tend to classify all voice disorders under the headings of aphonia and hoarseness. These terms are useful but not adequate for a discussion of the variety of existing phonatory defects.

Aphonia and Its Causes. Aphonia, the absence of phonated sound, is revealed as a whispered voice,

which indicates that the vocal folds are not vibrating. One should be careful not to confuse aphonia with elective mutism, in which the patient has the potential to phonate but simply chooses not to speak. Aphonia exists when the folds do not approximate sufficiently to be activated by the air stream or when the folds themselves are not capable of vibrating even when approximated. Causal factors in aphonia include lateral positioning of the arytenoid cartilages with the associated open glottis, massive intrusion of some neoplasm, atrophy or absence of vocal fold tissue, excessively stiff folds, or nonflexible mucosal covering of the folds. Aphonia can also result from organic disease. The associated organic diseases range from paralyses and other neural impairments to various tumors, inflammatory disease, scarring, and other localized laryngeal lesions. Aphonia also may be indicative of a functional disorder of psychogenic origin, most commonly conversion aphonia/dysphonia. Treatment of such functional disorders must be sensitive to the underlying psychogenic bases of the disorders. Aphonia may be present continuously but is more frequently intermittent. Intermittent aphonia varies considerably, sometimes even within the speech of a single patient. Some persons whisper most of the time; others may be aphonic momentarily within words or parts of sentences.

Breathiness. Chronic breathiness can be recognized by excessively audible breath-flow noise that is accompanied by a relatively low vocal loudness level (a low vocal signal-to-noise ratio). The problem is common and varies from scarcely discernible levels of breathiness in the voice to the almost aphonic. The vocal folds vibrate during the production of a breathy voice but do not impede the airflow sufficiently to allow much increase in subglottal pressure. Air flows alternately more and less but is not completely interrupted during the vibratory cycle. Several organic conditions modify the glottal configuration sufficiently to create incomplete glottal closure associated with a breathy voice. Muscular tension dysphonia (simultaneous contraction of laryngeal adductor and abductors) can create incomplete closure during each glottal cycle and produce a breathy voice.[2] Moreover, various conditions of paralysis (unilateral or bilateral) as well as growths protruding into the glottal space (these include discrete mass lesions such as nodules, sessile polyps, peduncular polyps, cysts, papillomas, keratosis, leukoplakias, cancers, contact ulcers, and edema) can prevent the glottal borders from touching firmly during the vibratory cycle. Even complete closure of the muscular borders of the glottis can be associated with a breathy voice if glottal resistance is markedly reduced (eg, when flaccid vocal folds result from a lower motoneuron disease such as myasthenia gravis), or in the presence of hypokinetic neurologic disorders such as Parkinson's disease, or if a chronic posterior glottal chink (intercartilaginous portion) is observed.[13]

In many respects, the causes of a breathy voice are similar to, or may be the same as, those of aphonia. This fact stresses the observation that the voice is not a reliable indicator of the type or extent of laryngeal impairment or disease. At any given moment, the degree of vocal fold approximation that can be achieved is influenced by the vowel sound that is being produced,[14] the vocal effort used by the talker,[15] and the muscle tension in the larynx. These factors can determine whether the voice is normal, aphonic, breathy, or hoarse. Although many attempts have been made to identify laryngeal disease based on acoustic measurements of the vocal signal-to-noise ratio, particularly by investigators in Japan, such measures appear to be more related to perceived levels of breathiness than to specific diseases of the larynx.[16–20]

Breathy voice has been associated traditionally and sometimes exclusively with the availability of breath. This association is not valid apart from a consideration of the concurrent condition of the larynx. Breath supply may be expressed in two ways: the amount of air available and the force with which it is expelled. Observation reveals that a reduced air supply, such as that resulting from the removal of one lung, is not in itself a cause of breathiness. Substantial clinical evidence also indicates that reduced breath pressure alone is not a cause of breathiness. If low pressure produced the defect, a person with a normal voice would necessarily have a breathy voice whenever a tone of low intensity was produced. Many persons with low vital capacity have normal voices.

Harshness. Harshness is a physiologic opposite of breathiness. When the vocal folds remain in contact for a disproportionately long time in the vibratory cycle, a voice quality known as harshness results.

This vocal problem is usually functional. Harshness is accompanied by increased air pressure and tighter glottal closure. When the harsh voice is produced loudly, as it is sometimes by hyperactive, aggressive individuals, the speech is staccato, and the listener perceives hyperactive contraction in the entire respiratory, phonatory, and articulatory musculature. This behavior is called hyperfunctional voice production. Organic causes of harshness include edematous vocal folds, neoplasms, and any other structural alteration that may prolong the closed phase of the vibratory cycle.

Hoarseness. Hoarseness and harshness are often confused with each other, partly because both terms are relatively imprecise. Hoarseness varies from mild to severe and is so common in the population as a result of acute upper respiratory infections that it evokes little concern unless it is severe or chronic. The differentiating, audible feature of hoarseness is a roughness that results from random variations in the periodicity of the glottal waveform and/or random variations in the intensity of consecutive sound. Such variations may contribute to, or be accompanied by, noise in the phonation. Consequently, it is understandable that the perceptual conditions of hoarseness and breathiness often coexist.

Normal vocal sound results from repetitive vibrations that are similar to each other in time and intensity or that vary progressively as pitch or intensity shifts upward or downward. Vocal sound is generated by the motions of the folds in conjunction with the breath stream that activates them, and any condition that alters their regular, repetitive, synchronous vibration, causing randomly timed or randomly intense pressure pulses, creates a voice quality called hoarseness. Physical conditions that cause random aperiodicity probably result from a combination of transient and interference factors and may include any disease or condition in the larynx that changes the size, stiffness, or surface characteristics of one or both vocal folds (laryngitis or any of the discrete mass lesions, eg, nodules, polyps, cysts, papillomas) or that causes excessive squeezing of one fold against the other (eg, pressed phonation by a patient with contact ulcers). Any of these factors may create the conditions for hoarseness. Enlargements caused by tumor or edema, reductions resulting from surgery or atrophy, flaccidity subsequent to neural involvement, and functional hypercontraction of the adductor muscles represent the kinds of changes to which reference has been made. Under certain circumstances, swollen tissue or secretions above or on the vocal folds may be set into more or less independent vibration and may thereby create hoarse or rough sounds. Descriptions of the many laryngeal diseases and anomalies potentially related to hoarseness occupy much of this volume. The possible effects of some of these conditions are summarized in a later section on voice therapy.

Spasmodic Dysphonia. The term spasmodic dysphonia (sometimes called spastic dysphonia) designates one of the most disabling of communication disorders. It is probably an organic movement disorder with an unknown cause. It prevents or seriously limits the person who has it from holding positions in which speaking is required. The problem originates in the larynx and is heard most frequently as a sudden, momentary interruption of the voice caused by brief, spasmodic glottal closure. In some patients, instead of closing, the glottis spasmodically opens to allow air to escape, as in a whisper.[21] The closure form of the disorder is often referred to as adductor spasmodic dysphonia and the open form as abductor spasmodic dysphonia.[22] In addition to these symptoms, some patients describe difficulty with breathing in the adductor type. Also, as patients attempt to "speak through" the spasmodic closure, the voice may be described as "squeezed," "effortful," or "struggle strained." Occasionally, when a vowel sound is prolonged, as in singing, a prominent intermittent pulsing occurs, suggesting a tremor. The adductor form is more common than the abductor, but both may be heard in a single patient. The spasmodic episodes occur most frequently on vowel sounds, particularly at the beginning of words. The abductor form is most prominent on vowels that follow unvoiced fricative sounds. Almost without exception, persons with spasmodic dysphonia can whisper normally without interruption in the flow of speech. Superficially, spasmodic dysphonia resembles stuttering, but significant differences exist between these disorders. One of the puzzling features of this type of dysphonia is that the severity and type of symptoms can vary within as well as among patients. Inconsistency of symptoms undoubtedly contributes to the traditional assumption that spasmodic dysphonia is psychogenic.[23,24] Indeed, it may be, but research evidence indicates

the possibility of a neural etiologic component.[25–27] A discussion of the several theories is not appropriate here but may be explored through references at the end of the chapter.[28–31] Differing concepts are mentioned simply because the neurogenic theory has led to surgical therapeutic procedures that have been more successful in treating the adductor type than other approaches.[32] The tenacious persistence of spasmodic dysphonia has caused patients to seek help from psychiatry, clinical psychology, speech therapy, hypnosis, drugs of many varieties, and bizarre remedies of every type, all unsuccessfully.[33]

The observed freedom from spasmodic dysphonia during whispering, and almost complete fluency by patients when they produce a breathy voice, demonstrates that the spasmodic adductory speech symptom is absent when glottal closure does not occur. Short-term clinical relief of the dysphonia can be achieved by chemically blocking one recurrent laryngeal nerve to create a temporary unilateral vocal fold paralysis. This injection procedure is also useful in diagnosis because it demonstrates the effect of a unilateral paralysis on the speech symptoms. When the injected anesthetic results in immediate elimination of both the tight glottal closure and the symptoms of spasmodic dysphonia, the diagnosis of the disorder is strongly supported. The resulting voice is not normal, although it is useful for communication purposes.

Patients with spasmodic dysphonia usually have had the disorder for many years and have become frustrated with the uniform lack of assistance from various therapies. They are generally eager to prolong the fluent speaking they experience during the nerve block and request some way to accomplish that end. The three options that are available are (1) the creation of a permanent unilateral paralysis by sectioning of one of the recurrent laryngeal nerves, a therapy that has become rather widely used[34–36]; (2) the establishment of a temporary unilateral paralysis by crushing one recurrent laryngeal nerve[37]; and (3) the production of another temporary laryngeal nerve paralysis by the injection of botulinum toxin.[38] The crushing and toxin procedures are usually followed in 3 to 6 months by a return to pretreatment function. The underlying rationalization for the temporary paralysis procedures is that if spasmodic dysphonia is psychogenic, the fluent speech practiced during the period of paralysis should persist after recovery from the paralysis. Unfortunately, the speech symptoms tend to recur with return of function. When unilateral paralysis is permanent after section of one recurrent nerve, the voice is almost always normally fluent for at least the first year, but as time passes, more and more patients report recurrence of their symptoms.[39–41] The several reported failures signal the need for increased caution with the surgical approach and careful assessment of the conditions associated with the occurrences.[42] In contrast, some patients are known to have maintained normal fluency for more than 10 years. These varying results lead to the speculation that spasmodic dysphonia symptoms are probably the result of multiple causes.

The abductor type of spasmodic dysphonia is not helped by intentional unilateral vocal fold paralysis. This procedure simply exaggerates the undesired glottal opening. Patients with this problem report considerable variation in symptoms, with the disorder worsening with anxiety and worry.[43] Psychiatry, clinical psychology, speech-language pathology, medicine, and many exotic approaches have been as unsuccessful with this form of spasmodic dysphonia as with the adductor form. Unfortunately, there is no known temporary alleviation of the abductor symptoms that parallels the anesthetic blocking of one recurrent laryngeal nerve used with the adductor form of the disorder. There is substantial evidence of neurogenic disorder with the abductor as well as the adductor form of the dysphonia. Support for neural involvement has been found in electromyography,[44] electroencephalography,[45] magnetic resonance imaging,[46] auditory brainstem response,[47] and various clinical neurologic tests.

Resonance and Resonance Disorders When the shapes and adjustments of the resonance spaces do not conform to the customary configurations, resonance disorders are apt to be present. The two most common resonance defects are too much nasal resonance (hypernasality) and insufficient nasal resonance (hyponasality). The first is caused by incomplete closure of the velopharyngeal valve or by a fistula in the structures separating the oral and nasal spaces, and the second is caused by blockage of the nasal passageway. Incomplete velopharyngeal closure results from one or more of five possible causes: (1) congenital deformity, such as cleft palate, submucous cleft, excessively deep pharynx, or short palate; (2) paralysis of pharyngeal or palatal mus-

cles; (3) destructive disease; (4) surgical procedure in which adenoidal tissue vital to velopharyngeal closure has been removed; and (5) imitation of hypernasal speech.

The second general condition, blockage of the nasal passageway, causes denasality (cold-in-the-head speech), which also results from one or more of five causes: (1) growths such as adenoid tissue, papillomas, polyps, and nasal spurs; (2) hypertrophy resulting from chronic disease; (3) swollen mucosa associated with allergies; (4) trauma to the nose; and (5) imitation. A blockage situated anteriorly in the nasal passageway may contribute to hypernasality instead of hyponasality because the posterior nasal areas act as resonators if an opening is present at the velopharyngeal valve. This form of hypernasality can be demonstrated by pinching the nose closed and trying to speak "through the nose."

A resonance disorder identified as muffled quality or "hot potato speech" is often difficult to identify because it is similar to some rural dialects. It results from space-occupying lesions such as masses of lymph tissue or tumors in the valleculae between the epiglottis and the base of the tongue. Confusion with dialects occurs because some speakers habitually retract the base of the tongue into the pharynx on sounds that customarily do not use that adjustment.

Combined Sources of Voice Quality Disorders. The two separate sources of voice quality disorders, the larynx and resonance cavities, are capable of functioning independently; consequently, quality deviations can be generated in either one separately or in both concurrently. This possibility provides the means by which many combinations of phonatory and resonance problems may be produced simultaneously and presents a basis for the variety of voice quality disorders that exist. The potential array of vocal problems may be compounded by the simultaneous presence of both organic and functional deviations. It is apparent that vocal disorders deserve careful evaluation and diagnosis; it is equally clear that vocal rehabilitation may require more than one type of therapy.

Summary of Conditions Affecting Laryngeal Health and Function

To stress the need for careful and complete medical evaluation and treatment of the individual who has a voice disorder is probably redundant. Otolaryngologists, particularly, are aware that voice problems are among humankind's most subtle disorders; they reflect certain aspects of the patient's thinking, behavior, health, and diseases. The causes of voice disorders may be arbitrarily grouped into the following nine categories:

1. Structural modifications that result from misuse of the voice, such as thickened vocal fold tissue, myasthenia laryngis, vocal nodules, and contact ulcer. Furthermore, vocal abuse combined with infection may cause such chronic conditions as laryngitis, corditis with hypertrophic and hyperplastic laryngitis, and atrophic laryngitis.
2. Anxiety, emotional turmoil, frustration, and similar psychogenic problems that cause excessive contraction of the laryngeal, thoracic, and associated musculature during speech but do not cause observable organic disease
3. Diseases and growths, including infections not related to vocal abuse; paralyses of central, peripheral, and myopathic origin; cysts; and both benign and malignant tumors
4. Mechanical and chemical irritants affecting the mucosa: fumes, irritating vapors and gases, dry air, dusts, allergens, caustic fluids, and gastric reflux
5. Substances causing noninflammatory edema, such as internal medicaments, mechanical compression of venous blood flow, endocrine imbalance, and allergy
6. Irradiation therapy, antihistamines, and other medications that cause drying of the mucosa
7. Congenital anomalies
8. Destruction of laryngeal tissue by surgery, trauma, or disease
9. Systemic disorders leading to chronic fatigue, such as anemia, metabolic disturbances, malnutrition, and inadequate rest, can also affect the voice adversely

Therapy for Voice Disorders

The long history of medicine has demonstrated that medical and surgical treatment may eliminate some types of voice disorders, but such treatment cannot always restore normal function. Nonmedical rehabilitative measures may be necessary to help compensate for altered anatomic and physiologic

conditions, and re-educative procedures are usually indicated in the treatment of habitual or functional disorders. The preceding descriptions of voice disorders and their causes have indicated the potential complexity of voice problems. Rehabilitative measures used in the management of such problems follow the approved convention of adapting the therapy to the specific patient and his or her disorder. Because this presentation must play a relatively minor role in a volume on otorhinolaryngology, suggestions for therapy have been limited to basic or universal recommendations that can be used by the physician who cannot devote much time to vocal rehabilitation. When greater depth is needed, one should refer to detailed books and extensive reports in the periodical literature to support systematic therapy. The purpose of this volume is best served by directing the physician's attention to the following four aspects of diagnosis and voice therapy: (1) conditions affecting the patient's general health and the function of the larynx, (2) environment, (3) psychological adjustment, and (4) voice.

Environmental Factors A program of voice therapy that does not recognize the demands of the environment on the communication needs of the patient is incomplete and may be doomed to failure. The person who must talk more or less continuously in a noisy environment may develop detrimental vocal habits and may also abuse the larynx and create tissue changes. In some families, excessive and harmful use of the voice is so common or so subtle that the persons involved are not aware of the excess and do not associate the behavior with the voice problem. Therapy for a voice disorder that is generated and nourished within an environment often encompasses complex personal management problems. Frequently, the patient cannot move to another job or otherwise modify his/her living situation; when such changes can be made, however, the restoration of the voice is simplified. When the environmental conditions cannot be altered, the voice therapy must include a review of the situation with the patient, or with the parents if children are involved, to provide insight and a rationale for adjustment to the environment. To ignore the patient's work setting, recreation routines, and living patterns is to invite failure.

Psychological Factors An individual's attitude toward self and the environment is reflected in such vocal elements as the rate of speaking, choice of words, vocal pitch, loudness of the voice, and vocal quality.[48] These factors often indicate the degree of poise, anxieties, emotional states, feelings of friendship or hostility, and belief about acceptance or rejection. When these concepts cause the individual to use an unpleasant, inadequate, or defective voice, any successful modification of the problem must include a consideration of the person's concepts about self, the environment, and his or her speech. In Western culture, the low-pitched male voice seems to be identified strongly with concepts of masculinity, and unless the young man realizes that a tenor voice is just as masculine as a bass one, he is apt to develop detrimental vocal habits that are carried into adulthood. Children as well as adults sometimes use loud voices or rough-sounding, hoarse voices to dominate, control, or compete in their family or social groups. The focus of attention here is on the fact that the attitudes and needs of the individual that led to vocal nodules or thickened vocal fold tissue must be considered in any program of vocal rehabilitation. Individuals occasionally demonstrate overwhelming anxieties in vocal disorders (eg, hysterical aphonia, intermittent dysphonia, and tremulousness). In these conditions, psychiatric assistance is usually desirable, and if any work on the voice is recommended, it is carried on as a supporting activity in close cooperation with the psychiatrist. The preceding paragraphs have suggested that psychogenic factors of many degrees of severity may cause voice disorders. It follows that successful vocal rehabilitation must include appropriate psychotherapeutic procedures when indicated.

Analysis of the Voice Problem Defective voices are usually not consistently abnormal in the same degree under all circumstances. A voice that is hoarse at a low pitch may be completely clear in falsetto. The vocal sound may vary during the production of different vowels, and the characteristics of vocal defects often change with loudness or vocal effort. The individual who provides voice therapy must know the capacities and limitations of the patient's voice because this information not only provides diagnostic data but also determines the pattern of rehabilitation. To determine the dimensions of a voice disorder, the clinician should evaluate systematically and individually the pitch, loudness, and quality characteristics of the voice.

The patient should be asked to sing a vowel, such as /a/, up and down a musical scale to the limits of the individual's range. It may be necessary to provide a tone to imitate because many patients with voice disorders have poor pitch discrimination. The adequacy of the pitch range should be noted, as well as any changes in vocal quality that occur at different positions on the scale. The patient should also be required to read a few paragraphs of simple material to reveal typical vocal habits. Generous samples of the patient's vocal production should be recorded during the evaluation process. Recordings are essential for objective listening to the voice and serve as a basis for determining improvement or regression during the course of therapy. The vocal evaluation should answer the following questions: Is the pitch of the voice that the patient uses in conversation and while reading aloud normal for a person of that age, sex, and size? Are there any unusual inflections or atypical pitch variations? Is the pitch range satisfactory, such as a minimum of five to six tones? Can the patient match the pitch of a given tone and voluntarily go up and down a scale? Is the loudness of the voice appropriate to the circumstances? Is there any evidence of a hearing loss? Is the quality of the voice predominantly normal? Is it aphonic? Breathy? Harsh? Mildly hoarse? Severely hoarse? Does the quality change when the pitch is high or low? Does the speaker appear to be using too much or too little effort during speaking? Does too much sound seem to come out through the nose? Does too little sound pass through the nose on the /m/, /n/, and /ng/ sounds? Is there any tremulousness? Is the speech commensurate with the socioeconomic background of the patient? Under what conditions of pitch, vowel sounds, and loudness levels does the subject produce the best vocal sound?

Patient's Evaluation of the Voice Problem Two basic premises support the concept that one must evaluate one's own voice problem if one wishes to modify it. First, an individual does not hear his or her own voice as others hear it; second, the clearer and more specific a goal or task can be made, the faster it will be reached. A corollary to these concepts is that unless an individual can recognize the voice problem and can form a clear concept of the improved vocal sound, remedial efforts will be relatively ineffective. The process by which a person evaluates his or her own voice is systematic listening, sometimes called "ear training." Samples of the patient's voice should be recorded at appropriate intervals during therapy, and these should be studied carefully with frequent comparison of the changing vocal features over time. The clinician should choose one vocal element at a time for special attention and must avoid both discouraging and overwhelming the patient with the re-educative task. This purposeful listening requires great clinical skill to make it meaningful and to keep the patient motivated. Inexperienced listeners do not hear many of the elements of vocal sound, particularly their own, until they are instructed; this is the function of the clinician. The temptation to send the patient home to listen to a voice recording as a therapeutic measure should be resisted until one is assured that the patient can identify specific faults. Another aspect of evaluative listening is detailed analytic study of the vocal characteristics of other persons with both good and poor voices. This exercise improves the patient's ability to identify specific vocal features and use them in production through imitation and experimental variations.

Voice Therapy *Direct Voice Therapy.* Two types of instruction relate specifically to direct voice therapy: the first can be classified as recovery, the second as training. Recovery procedures presume a need for healing, for a return of the structures to normal. They are based on the premise that the vocal organs will restore themselves if abusive behavior is discontinued. Recommendations commonly given to achieve these goals include complete vocal silence for a week or two (or sometimes longer) with no whispering, limited vocal use in which speaking is allowed only when absolutely necessary, reduced vocal intensity, elimination of all singing, limitation of physical exercises and activities that cause the breath to be impounded by the closure of the glottis, and avoidance of coughing and clearing of the throat whenever possible. These recommendations are essentially passive but are often desirable early in therapy. However, if recovery procedures have allowed the larynx to become more normal, the resumption of habitual patterns of phonation ordinarily causes a relapse. To be successful with these patients, voice therapy must include a period of training that modifies previous habit patterns and replaces them with more efficient phonatory behavior.

Imitation and Experimentation. Soon after learning to hear specific flaws in the voice, the patient should

be encouraged to experiment with voice production in an effort to modify specific faults and also to develop greater control over the voice. It is often helpful to use computer software that permits two voice samples to be displayed at once. One channel can display a previously recorded voice sample or the clinician's live voice sample and the other displays the patient's efforts to match the model.

Relaxation. With voice disorders in which vocal abuse is present, the patient usually has excessive tension in the muscles of the larynx as well as elsewhere throughout the body. Unless this tension can be controlled, vocal therapy cannot progress. The patient needs to relax. The term relaxation commonly has two meanings: it may refer to the absence of muscle contraction, or it may signify coordination in which the opposing sets of muscles exert just enough reciprocal tension on each other to accomplish a desired movement with perfect control. Both types of relaxation can be, and sometimes must be, learned for the successful alleviation of a voice disorder. Learning to relax involves both muscle training, which can be approached directly through exercise, and emotional control, which is managed through modification of attitudes, anxiety, worry, and comparable problems.

Pitch Adjustment. Many persons, particularly men, attempt to use a vocal pitch that is lower than that which is normal for their laryngeal structures. This behavior not only produces an unpleasant voice but also traumatizes the laryngeal structures. If this vocal tendency is observed, its inadequacy should be explained and assurance provided that a more normal pitch will supply a satisfactory masculine image. Subsequently, direct instruction can be given to establish a habitual pitch that averages four or five musical notes above the lowest tone that the patient can produce. An average shift of as little as one tone upward "feels" much higher to the patient and causes reluctance to accept the change. Furthermore, the patient is usually surprised to discover little change in the recording and is startled when the average listener does not recognize a change of as much as several tones in his speaking pitch.

Speech Articulation in Patients with Voice Disorders. Although articulatory disorders are discussed later in this chapter, a close relationship exists between disorders of articulation and certain voice problems that need attention in a discussion of voice therapy. Two types of phonatory problems are effectively treated through the management of articulation. One of these is observed when generalized excessive muscle tension is chronic during speaking, and the other is found in indistinct speech caused by imprecise or inaccurate movements of the lips and tongue. Where excessive muscle tension exists, mouth opening is usually minimal, talking is of the "clenched jaw" variety, and the vocal sound may be hoarse, harsh, or hypernasal. Easier, more efficient speaking usually accompanies increased opening of the mouth, reduction of rate, and relaxation of the muscles of the head and neck. When imprecise, inaccurate, or lax speech exists, the tongue and lips do not make firm contact with their opposing structures, that is, movements are often incomplete in a phonetic sense. The speech is frequently called "slurred" or "careless." Actually, the atypical sounds are usually learned in an environment where such speech is typical. When speakers with these habits are requested to talk so that they can be more easily understood, they attempt to comply by speaking more loudly and by increasing laryngeal and respiratory muscle contraction. Training in more appropriate production of speech sounds usually increases intelligibility and reduces the detrimental, excessive effort.

Vocal Practice. If the physician finds it necessary or desirable to direct the vocal practice of a patient, the focus should be on the use of a quiet voice. More specifically, practice sessions of 5 to 10 minutes should be advised in which vowel sounds are prolonged gently at various pitches, and short selections from magazine stories and other sources are read aloud. If the subject has experimented successfully with various voice qualities and is able to relax adequately, such practice will prove beneficial. If the patient uses detrimental vocal behavior when he/she reads aloud, however, this practice should be postponed. Vocal practice requires patience and constancy. Patients are often disturbed when they are told that they may not detect much improvement in 3 months or more and that 6 months may be required to achieve some skill in the use of new vocal habits. Persons preparing their voices for singing or acting accept the concept of lengthy training, but those correcting voice defects seem to expect a mag-

ical change. Consequently, one of the most difficult and important aspects of voice therapy is motivation of the patient.

Laryngeal Cancer and Vocal Rehabilitation

Hoarseness, laryngeal cancer, and speech rehabilitation are intimately related. The sound of the voice is often the first evidence of disease and consequently is important in diagnosis. The choice of medical treatment for the disease, whether it be irradiation, chemotherapy, partial laryngectomy, total laryngectomy, or combinations of these procedures, influences the type and amount of speech therapy to be used after control of the disease. Although cancer in the larynx may not influence the voice, in more than half of the patients, hoarseness is present and is accompanied frequently by pitch change and reduction of loudness. A retrospective study of 260 patients with laryngeal carcinomas by Lowry and others reported that "hoarseness was the predominant symptom in all four stages of glottic and supraglottic lesions."[49] In 1974, the American Cancer Society anticipated 9,500 new cases of laryngeal cancer for that year: 8,300 expected in men and 1,200 in women.[50] In 1988, the estimate was 12,200 new cases, with 9,900 in men and 2,300 in women.[51] The estimated new cases in 1994 totaled 12,500: 9,800 men and 2,700 women. A slight downturn in the number of laryngeal cancers was estimated for the year 2000: 10,100 new cases, with 8,100 men and 2,000 women.[52] These data have not been adjusted to general population growth. There appears to be little change in the proportion of cases in the general population over the 25-year period, but the female-to-male ratio has increased progressively from 0.144 in 1974 to 0.246 in 2000.[53] A report by Gates and associates in 1982 revealed that "of 103 people with clinical diagnosis of laryngeal cancer studied by the authors, 53 eventually were treated by total laryngectomy and, in some cases, radical neck dissection (43 cases), preoperative radiation therapy (15 cases), postoperative radiation therapy (29 cases), and postoperative chemotherapy (7 cases)."[54] The implication of these reports is that in approximately half of the cases, laryngeal cancer will be treated by total laryngectomy. Subtotal laryngectomy and irradiation will be used with the other half, only a relatively small number of whom will require extensive vocal rehabilitation. Training in communication skills is almost always desirable for patients who have had a total laryngectomy.

Presurgical Considerations Presurgical management of the person who must have the larynx removed is extremely important in the total therapy and should be handled with great care. Most laryngeal surgeons have developed effective procedures for informing and advising their patients about the surgery and its sequelae. A study by Snidecor of approximately 150 laryngectomees, however, indicated that presurgical management could be improved.[55] Of the patients studied, 85% believed that they had been counseled well and at the right time, but the majority of these patients also stated that they would have benefited from more specific counseling for themselves and their spouses about the operation, the postsurgical physical conditions, and the impending speech impairments. Only 53% learned from their surgeons about the kinds of devices available for replacement of voice. The remainder obtained their information from other laryngectomees, speech-language pathologists, and the American Cancer Society. At the time when the diagnosis of laryngeal cancer was reported to the patients, about three-fourths of them were accompanied by their spouses, which was considered helpful and desirable, but more than half of the spouses were "shocked, anxious, and deeply concerned."[55] The same study confirmed what most laryngologists know: that patients who have been told they have cancer and need surgery are filled with anxiety and many fears. They fear the cancer and the possibility of death, the surgery, and the loss of voice and its consequences, and they are distressed about the potential effects of vocal impairment on their work, family life, and social relationships. Such fears and anxieties are to be expected, but the sensitive counselor who acknowledges such feelings and whose experience allows anticipation of the postsurgical conditions can often reduce the patient's apprehension. The importance of patient attitudes and their improvement as part of the counseling of patients was revealed also in a study by Gates and associates.[56] In their second article showing an analysis of factors related to success or failure in rehabilitation, they stated, "Factors that correlated with successful rehabilitation were less postoperative radiotherapy, more vigor, greater imaginativeness, independence and

self-assuredness, and absence of denial."[57] The latter four factors are amenable to counseling and reflect the importance of this process in both the pre- and postsurgical periods.

Presurgical Speech Training Some laryngeal surgeons regularly introduce a skilled laryngectomized speaker to the patient before surgery. Others believe this practice to be unwise. Gates and associates stated that "… a preoperative visit by a laryngectomee is not associated with success in learning esophageal speech or in rehabilitation and in fact, is statistically significantly correlated with a poorer outcome."[56]

Reliable estimates of the prevalence of presurgical speech training are not available, but few speech-language pathologists currently seem to advise this practice. Perhaps the most satisfactory way to plan for postsurgical speech in the presurgical period and to handle the many questions about work, family, and social relationships is to introduce a mature speech-language pathologist who has had extensive experience with laryngectomized persons. The speech-language pathologist's function is to encourage and to motivate the patient to make plans for the postsurgical period. The speech-language pathologist does not promise fluent speech or that the learning will be easy but expresses confidence that verbal communication is usually possible with esophageal speech, a tracheoesophageal prosthesis, or an artificial larynx.

Early Postsurgical Period Within a few days after surgery, the patient begins to assess the new condition. The patient is relieved that the operation is successfully past and is glad to be alive but discovers new problems and fears to add to those that linger. The patient is bothered somewhat by breathing and coughing through the stoma and is dismayed by the reality of voicelessness. Attempts to whisper are frustrating. During convalescence, anxieties often grow about employment, personal appearance, family relationships, and association with friends. If the patient or the patient's associates have impaired hearing, communication becomes doubly difficult. The inability to talk leads to depression and withdrawal from social situations, a condition that becomes progressively more difficult to reverse. The distinction between vocal tone production and speech sound formation is clearly illustrated by the laryngectomized person. Without the larynx (voice generator), the patient can simply mouth the words. The removal of the larynx usually leaves the articulatory mechanism intact. Consequently, when the patient substitutes some other sound for the missing laryngeal tones, speech is possible. The options open to the laryngectomized person to develop a new voice are discussed later in this chapter. If the person who loses the larynx could resume talking immediately after the laryngectomy, concern would be no greater than that associated with any other major surgery. Unfortunately, when the means of communication is disrupted, the patient feels isolated and at a great disadvantage. The early restoration of the ability to communicate is a major consideration in successful rehabilitation. Every possible effort should be made while the patient is still in the hospital to establish communication by computer, writing, picture boards, signing, or the use of an artificial larynx. In addition, plans should be made during this period for instruction at home. This program should include direct work with the patient. Of equal importance is instruction of the spouse and other family members. Because individuals vary greatly in their ability to understand alaryngeal speech, the problem listeners in the family group should be identified and taught to comprehend. Some of the instruction will include speech reading skills and the use of hearing aids where indicated.[58]

Hospital care of laryngectomees, whether in the immediate postsurgical period or at some later admission for any reason, deserves special attention from all service levels. The inability of the patient to communicate creates serious problems, especially when specialized nursing care is not available. Many horror stories have been related by laryngectomized persons, ranging from aides splashing water into the stoma during bathing and application of oxygen to the nose to failure to open a plugged stoma. The International Association of Laryngectomees recommends that a special "warning" sign at the bed and on the patient's chart identify every hospitalized laryngectomee and that the care providers be given specific instructions.[59]

Restoring Speech Communication Following Laryngectomy It is a truism that any form of substitute voice is inferior to that produced normally. Similarly, any usable speech is superior to mutism, whispering, writing, or signing. Clinicians must

search for the best substitute voice sources in relation to the needs and capacities of the individual patient. Because structural adjustments above the larynx regulate the acoustic filtering of the laryngeal sound source, they are primarily responsible for creating the sounds of spoken language. Thus, any complex sound that can be put into the upper airway can be substituted for the laryngeal sound source and molded into speech. The need to restore speech to the laryngectomized patient has enabled engineers, surgeons, speech-language pathologists, and others to devise various means for producing sound that can substitute for the normal voice.

Engineering Approach. The engineering approach involves the development of artificial sound generators. These can be separated into two general types according to the driving force used in the production of sound: one is powered by the breath stream, the other by electricity from batteries.

Breath-Activated Instruments. The various forms of breath-activated instruments have three elements in common: (1) a flexible tube that conveys the breath from the tracheal stoma to a small capsule held in one hand; (2) the capsule, which contains a reed or membrane that is vibrated by the breath, thereby generating the sound; and (3) a second, smaller tube that carries the sound from the capsule into the mouth, where it is articulated into speech. Breath-powered, handheld artificial larynges are inexpensive and offer the patient an early means of talking with little or no special instruction. The speech sound can be relatively loud, phrasing and sentence length are normal, and speech is intelligible, even though it has a monotonous pitch. Unfortunately, the disadvantages attending the use of these simple instruments are substantial; consequently, few of them are currently in use. The problems can be listed as follows: moisture from the exhaled breath condenses relatively quickly at the vibrator, causing malfunction and the necessity of draining or drying it; the tube conveying air from the stoma to the vibrator becomes congested with mucus from the trachea, causing both hygienic and esthetic problems; and saliva seeps into the sound-conducting tube, contributing to vibrator problems and necessitating frequent changes and washing of the tube to keep it clean. Because the unit is held by one hand, it interferes with two-handed activities; furthermore, because it must be transported by the user and kept readily available for speaking, it becomes a nuisance.

Electrical Artificial Larynges. Two major types of electrical instruments are in common use. In one, the sound is generated by a battery-powered buzzer, about the size of a hearing aid phone, that is held in the hand and from which the sound travels by way of a tube into the mouth. The sound is formed into speech by articulatory movements as it is with the breath-powered unit. The second type of instrument has several forms, but basically it is a handheld unit, about the size of a two-cell flashlight, that transmits a buzzer type of sound through the tissues of the neck into the pharynx, mouth, and nose, where the sound is articulated into speech in the customary manner. The vibrating element of these units is a disk that is placed against the skin of the neck. It is activated by a battery and controlled by a switch similar to that on a flashlight. Battery-powered artificial larynges share with the breath-activated units such advantages as low cost, early and easy restoration of speech, and good intelligibility. The electrical tissue-vibrating instruments have the additional positive features of freedom from hygienic problems, moisture condensation, and instrument care that accompany the breath-powered units. On the negative side, electrical artificial larynges require one hand for operation and must be carried continuously by the user. Furthermore, the vibrating disk models generate a field noise that is transmitted directly to the surrounding air, where it competes with the sound that comes from the speaker's mouth. The ambient noise and the monotonous speech sound produce an artificial, machine-like utterance. Breathing and phrasing are interrelated in normal speech. When the basic sound is produced more or less continuously, as with the electrical instruments, utterance tends to become "mechanical." Experienced users of these units, however, learn to start and stop the sound to approximate the appropriate phrasing of the sentences.

Surgical Approach. Surgical efforts to produce substitute vocal sound have extended over many years and can be classified into three forms.

Tracheoesophageal Shunts. A primary surgical approach to the construction of a substitute sound source is the formation of a tracheoesophageal shunt.[60,61] Basically, the shunt is a small channel extending from

the upper, posterior wall of the trachea through the anterior wall of the esophagus. When the tracheal stoma is closed with a finger, exhaled air passes through the shunt into the esophagus, where it vibrates the pharyngoesophageal sphincter or other esophageal tissue to create a sound. At least four advantageous features are associated with the surgical creation of a shunt for the production of voice. When the air resistance in the shunt is relatively small, the manner of speaking is almost effortless; the voice is produced on the exhalation of pulmonary air as in ordinary speaking; phrases and pauses occur at linguistically appropriate places; occasionally, slight pitch variation is achieved; and little training in the use of the method is needed. The Blom-Singer and Panje prostheses are comfortable, reliable, and relatively unobtrusive.[62] Detailed descriptions of these prostheses and the procedures for placing them are presented in Chapter 54. Additional information appears in the literature.[63,64] The major limitation, from a speech standpoint, is that a finger must be used to close the tracheal stoma whenever sound is produced unless the patient is able to wear a tracheostoma valve.[65] Vocal intensity is usually reduced from that of normal speech, and the prosthesis must be removed, cleaned, and reinserted. The new prostheses have greatly reduced problems of leakage, infection, and stenosis.[66] Patients must be selected carefully and motivated to want to improve their speech and must be capable of managing the hygienic care of the prosthesis as well as having the dexterity to remove and replace it.

Construction of a Pseudoglottis. Theoretically, the ideal solution to the loss of voice through the removal of the larynx would be the construction of a living tissue substitute that would produce sound and have no detrimental sequelae. In this approach, the surgeon constructs a pseudoglottis at approximately the location of the original larynx. Muscle and cartilage that can be spared are formed into an airway continuous with the trachea. At the junction with the pharynx or esophagus, a vibrator "buttonhole" is formed that can generate a sound when the breath is exhaled through it. Some success has been achieved through this approach by highly skilled surgeons with carefully selected patients. Unfortunately, the pseudoglottis is usually not a liquid-tight valve. Fluids leak from the pharynx or esophagus into the trachea, leading to pneumonia and other diseases. In patients with whom the procedure has been successful, the voice has been good. When a tracheal stoma has been maintained, however, the patient has the disadvantage of using a finger to cover the opening to divert the air stream.

Combined Surgical and Artificial Devices. In this approach, an artificial, breath-powered, sound-generating reed is mounted in a small box and is activated by air conducted through a tube from the tracheal stoma to the box. The sound is discharged through a second tube that extends from the generator to a surgically created fistula that passes through the skin and deeper structures at the side of the neck into the pharynx. The small enclosure containing the sound source lies on the patient's chest just below the stoma.[67] The surgical creation of a fistula into the pharynx, permitting the connection of an external, artificial sound source, appeals to some laryngectomized persons, particularly those who have been unable to learn esophageal speech and who do not want to use one of the handheld artificial larynges. The advantages of the surgical fistula and attached mechanical device include the capacity to produce highly intelligible, fluent speech with customary phrasing and the utterance of sentences of normal length. Both hands are freed for normal activity. The disadvantages of this system are (1) the obvious esthetics, the hygienic and clothing annoyances caused by a device attached at both the tracheal stoma and the fistula; (2) the special surgical procedure that is necessary to create the fistula; (3) occasional rupture of a neck artery when the fistula is located incorrectly; (4) the hygienic care of the fistula and instrument; and (5) the need to insert a stopper into the fistula to prevent leakage when the tube that passes through the fistula is removed for sleeping or other reasons.

Esophageal Speech Approach. The third approach to speech recovery for laryngectomees is the development of esophageal speech. This procedure allows the patient to control and refine the natural eructation sound as the basis for speech. The advantages of esophageal speech are that (1) natural physiologic structures and functions are used, thereby obviating the need for an artificial device; (2) both hands are free for normal activities during speech; and (3) the good esophageal speaker presents a relatively normal appearance and speaking manner. The disadvantages of esophageal speech are that (1) it is often dif-

ficult to learn (according to Gardner, approximately 30% of laryngectomized patients do not learn to speak well with esophageal speech[68]), and frequently, 3 months or more are required to achieve fluency; (2) phrasing is usually changed, so fewer words than normal are uttered in a sequence without pause for phonatory air; and (3) the voice lacks sufficient loudness to be heard easily over common environmental noise. All of the substitute methods of creating voice, except the battery-powered, disk-type artificial larynges, encounter maximal problems when the patient eats. Swallowing food and the stimulation of salivary secretion complicate speaking and frequently embarrass the patient in social situations. Esophageal voice production and the artificial larynges with oral tubes are probably more adversely affected while eating than with the other methods discussed.

The air used in ordinary eructation comes from the stomach, and the sound usually does not result from voluntary acquisition and expulsion of air. Many persons, however, can learn to take air into the esophagus and expel it without allowing it to reach the stomach. In this method, air in the esophagus is put under slight pressure by respiratory movements and escapes in vibratory pulses through the sphincter mechanism at the upper end of the esophagus. This organ does not produce sound as effectively as the larynx, but its phonation function is similar, and it is subject to considerable regulation. The ability to control the air charge and consequently the moment of sound production is the primary problem in learning esophageal speech. Before air can be expelled from the esophagus, it must be taken in. If it is allowed to enter the stomach, it cannot be returned readily for voice; consequently, the air charge must not be swallowed in the sense that water is swallowed. Instead, it must either be "inhaled" or "injected" into the esophagus. Experience proves that either of these procedures is serviceable for each has been used by good esophageal speakers. Indeed, many of these speakers use both methods interchangeably.

INHALATION METHOD. If the esophagus were held open as the trachea is, the ordinary respiratory expansion of the thorax would draw air into the esophagus as readily as into the lungs. Many laryngectomized persons can learn to relax the upper esophageal sphincter at the moment of inhalation and thereby allow air to be pulled into the esophagus. In the initial stages of learning this procedure, however, before the laryngectomee is able to relax the esophageal valve readily, he or she can facilitate the intake of air by momentarily covering the stoma with a finger. This action suddenly decreases the pressure in the thorax, causing the air in the pharynx to be drawn into the esophagus. As soon as the air has been taken in, it can be expelled by increasing the thoracic pressure, as in exhalation. This return flow is interrupted intermittently by the vibration of the tissues of the pseudoglottis, thereby producing sound.

INJECTION METHOD. The "injection" method of putting air into the esophagus uses a different principle from that used in the inhalation procedure just described. In injection, the air is forced, pumped, or squeezed from the mouth and pharynx into the esophagus. Injection is accomplished by closing the lips and the velopharyngeal valve to provide an airtight oropharyngeal area while simultaneously thrusting the back of the tongue upward from the /d/ position. This movement pumps the air out of the mouth (oral cavity) into the pharynx, and because it is not permitted to escape through the nose and no longer has access to the larynx and trachea, it passes into the esophagus. Squeezing the air from the mouth into the esophagus can also be achieved by compressing the lips and buccal spaces, as in making a /b/ sound, and simultaneously lifting the mandible.

Several manuals and instructional booklets are also available through local cancer societies. It is often necessary for the person working with a laryngectomee to help with suggestions about everyday hygiene and appearance. The instructor should be able to advise on such items as care of the tracheal tube and prostheses, cleanliness, coverings for the stoma, suitable neckwear for men and women, precautions related to bathing, difficulties of speaking while eating, and similar personal problems. The clinician should also be able to help the laryngectomee join a "Lost Cord Club" for instruction, practice, and recreation. Addresses can be obtained from the International Association of Laryngectomees* or from a local chapter of the American Cancer Society.

*International Association of Laryngectomees, 8900 Thornton Road, Box 99311, Stockton, CA 95209, USA.

Selecting a Method of Speech Rehabilitation
The welfare of the patient is, of course, the first consideration in the rehabilitation process. The laryngologist, speech-language pathologist, and others involved in restoration of communication should assess the patient's needs and potential for rehabilitation as early as possible, preferably while the patient is still in the hospital. If the team determines that a shunt with a Blom-Singer or Panje prosthesis is appropriate, the steps toward spoken communication are clearly evident. When surgery is not indicated, attention can be given immediately to the use of an artificial larynx. Subsequently, when healing and the effects of postsurgical irradiation have progressed satisfactorily, esophageal speech instruction can be introduced. Most clinicians agree that good esophageal speech is more desirable than speech with artificial devices. The rehabilitation team, however, can determine when or whether such instruction should be introduced. If the decision is made to use an artificial larynx, either temporarily or permanently, the patient should be taught how to manipulate it most effectively and how to care for it.[69] Manufacturers customarily include printed instructions, but usually these need to be supplemented by demonstration and personal experimentation.

DISORDERS OF SPEECH

Speech-language pathologists recognize that substantial segments of the adult population as well as younger groups have serious speech, language, and voice defects, the problems that involve otolaryngologists most directly. Prevalence of communication disorders in children is high, with a gradual decline through adolescence. At least 5% of school-age children (up to age 17) have significant speech impairment, although most professionals consider this estimate extremely conservative. Some researchers have reported a prevalence as high as 10% for certain samples of children.[70] Most of these difficulties in children involve functional articulation disorders; faulty language formulation and dysfluency (stuttering) comprise the remainder. The presence of communication problems appears to stabilize during adulthood and tends to increase after the age of 40 years. There seems to be a higher percentage of disorders in men than women. Fewer disorders are assumed to occur among adults than among children. The ever-increasing size of our geriatric population, however, with its associated communication disorders, may require us to change that assumption. In addition, many neurogenic speech and language disorders, unlike stuttering and many articulatory disorders, develop in adulthood.

Speech and language disorders can be divided into three major categories. First and most common are the *articulation disorders*. These problems involve the production of defective speech sounds and sound combinations that may be distorted, omitted, substituted, or added as accessory sounds. Sometimes articulatory disorders are the result of neurogenic disorders, the dysarthrias. A second type of disorder is the *impairment of speech fluency*, called *stuttering* or *stammering*. Repetition of sounds, syllables, words, or phrases; sound prolongations; atypical pauses (hesitations); word substitution; and use of word fillers characterize dysfluent behavior. A third category is variously labeled *language impairment, linguistic disability,* or *faulty symbolization*, which refers to disorders in both expression of thought through verbal language and its comprehension. The category includes various disorders, ranging from delayed/deviant language development to neurogenic disorders known as aphasia.

Normal Speech Development in Children

The normal acquisition of speech may be expected to follow much the same pattern as the motor, adaptive, and personal-social behavior of the child. As in all other types of learned behavior, speaking depends on the process of maturation. A "speech readiness period" extends from birth to the fifth year of life, when the child acquires the ability to develop speech as a method of communication. This capacity, however, represents the cumulative learning of many bits and pieces of language that begin within the first few weeks of life.[71] Although endowed by nature with the capability of learning language, it appears that a child learns to talk only because people in the immediate environment speak to each other and to the infant. In the first few months of life, infants are capable of perceptually differentiating a wide variety of speech sound contrasts, giving children the capability of learning any of the languages of the world.[72,73] However, by 6 months of age, children have already begun to lose some of the capability of perceptually differentiating among sound contrasts that are not used by talkers within their immediate language environ-

ment.[74] Also, by 6 months of age, infants show an enhanced ability to distinguish perceptually those sounds used by talkers within their native language environment.[75] Infants and young children are typically "bathed in language" by their caregivers during the first months of life. Such stimulation, often called motherese, serves to nurture proper language development.[76] This potential for language learning is graphically highlighted by cases of abandoned, nonstimulated children who manifest severe language impairments. Before the child can learn to use oral symbols as a form of social behavior, he or she normally lives through a series of prelinguistic overlapping stages of sound production. The child's physical, emotional, and intellectual needs are served and expressed in each stage.[77] Furthermore, the emergence of speech in the "normal" child serves as an excellent index of his or her physical, intellectual, and emotional status. From a body systems perspective, the general sequence of speech development involves attaining proficiency with breath control, then laryngeal control, and finally articulatory control.[78] The development of speech can be classified into a sequence of five stages, each of which has distinctive characteristics that blend cumulatively.

The Cry The first stage is crying, which includes both undifferentiated and differentiated cry. During the first month or two, the cry does not differentiate sensations of hunger or pain, heat, cold, or other kinds of discomfort. Differentiated crying appears within the first few months, and the mother soon learns to interpret variations in the cry as signals of immediate needs or desires.[79]

Babbling In the second stage, babbling, the infant does not abandon differentiated crying but usually by the third or fourth month produces a variety of sounds at random and may appear to be "listening" as well as responding to sounds in the environment. Most of these are "cooing" sounds that are like vowels. Prosodic features (pitch, stress, and duration) begin to develop. Later the child begins to use the lips and tongue to make consonant-like sounds that combine with those vowels, producing what we call babbling.[80] Because the child's sensory feedback loop (primarily auditory) is beginning to develop now, the child's babbling becomes progressively imitative of the prosodic and phonemic aspects of the input (adult) language. Children at 3 to 10 months of age progressively play with adult forms of speech.

Reduplicated Babbling Reduplicated babbling (repeated syllables such as pa, pa, pa, ba, ba, ba, ku, ku, ku, etc) begins during the second 6 months. At first, the sounds produced are carried over from babbling, but later the infant appears to be responding more selectively to the sounds of others, as well as to those he or she has just produced. Reduplicated babbling appears to be a good index of normal development. Virtually all normal-hearing children produce reduplicated babbling before 12 months of age, whereas children with significant hearing impairments rarely do.[81] The speech development of the congenitally deaf child differs from that of the child with normal hearing beyond the babbling stage. The deaf child does not continue babbling for as long as the hearing child, and speech development usually follows a different pattern from that of a hearing child.[82]

Echolalia Echolalia, the name given to the fourth stage, begins somewhere around the ninth to tenth month, when the baby repeats sounds that he or she hears in the environment. The child seems to produce sounds that are pleasurable and repeats sounds that another individual produces. This is the stage when the child demonstrates that speech and language comprehension are emerging. Up to this stage, the baby has been acquiring an inventory of sound complexes that later will be used in learning to speak and in developing a vocabulary. The ages at which children typically master the sounds of English are shown in Table 49–1 (after Sander[83]).

Intentional Speech The fifth stage is the acquisition of speech as a practical tool for communication. Before entering this final but ever-expanding stage, the child must have established a functional understanding of conventionalized speech patterns. Providing meaningful time sequences for the mastery of speech sounds is difficult because many factors influence and regulate their acquisition. Girls generally surpass boys in learning speech sounds and tend to be accelerated in articulation skill from about $4\frac{1}{2}$ years on. Girls normally approximate mature articulation by the age of 7, whereas boys usually take an additional year to reach the same degree of proficiency. Whereas development of articulation skills is nearly complete for most normal children by age 8 years, subtle performance differences still exist for several years. Research evidence emphasizes that the accuracy of speech motor control continues to improve until age 11 or 12 years, at which point,

TABLE 49–1. Guidelines of Normal Acquisition of Consonant Sounds

	Customary Articulation	
Sounds	*50% of Children (yr)*	*90% of Children (yr)*
p, m, h, n, w, b	1.5	3.0
k, g, d	2.0	4.0
t, ng	2.0	6.0
f, y	2.5	4.0
r, l	3.0	6.0
s	3.0	8.0
ch, sh	3.5	7.0
z	3.5	8.0
j	4.0	7.0
v	4.0	8.0
th	4.5	8.0
zh	6.0	8.0+

Adapted from Sander EK.[83]

adult-like articulation performance is achieved.[84] In general, research indicates that the average child of 36 months should have acquired a vocabulary of approximately 1,000 words and should be using short, meaningful sentences.[72] Errors in the articulation will probably occur; nevertheless, 80 to 90% of conversational speech should be intelligible to an interested listener. If a child is not speaking acceptably by the age of 7 years, it may be because of any one or a combination of the following: (1) defective hearing, (2) poor motor coordination, (3) physical defects, (4) poor speech models in the family environment, (5) insufficient language stimulation, (6) mental retardation, or (7) emotional disturbance. As Table 49–1 highlights, however, occasional difficulty with later developing sounds should not prompt undue concern by parent, clinician, or physician.

Since the mid-1970s, a major shift has occurred regarding the theory behind speech sound development. Rather than tracking the developmental timetable of individual speech sounds, contemporary researchers investigate the emergence and disappearance of phonologic processes (eg, cluster reduction, unstressed syllable deletion, stop for fricative substitutions, etc).[85,86] A detailed discussion of the chronology of phonologic processes is beyond the scope of this chapter, but developmental charts/graphs are available to the interested reader in the professional literature.[78,87,88]

Articulation Disorders

Disorders of speech articulation may be described as faulty or atypical production of sounds in the spoken language. These problems encompass many kinds of articulatory defects, both functional and organic, and constitute the most frequent type of speech disorder observed within the population.

Functional Articulation Disorders When a disorder is designated functional, it does not necessarily indicate that it is a simple problem. Some functional articulation disorders are truly psychogenic, but most represent problems stemming from faulty learning and/or habits of misuse. Because successful therapy may depend on discovering causal factors, however, one must explore certain aspects of the individual's physical and social development. Habit-strength is a learning theory concept that relates to the automaticity of both appropriate

and inappropriate behavior. Applied to speech, habits can account for the persistence of a problem and the resistance to its successful treatment.

The child's mental age or intelligence must also be considered when articulation is being evaluated. Intelligence, reflected in cognitive functioning, underlies one's ability to learn. Because learning is fundamental to speech and language acquisition, intelligence factors can influence speech performance. It is not surprising, then, that numerous studies have demonstrated that children of low mentality produce more articulation errors than children of normal intelligence. Moreover, children with severe articulation defects frequently have depressed scores in reading and spelling, which may help to explain their low scores on verbal intelligence tests.

The word delay implies normal sound development as a reference, and impairment of this type is manifested as (1) a delay in onset, (2) slowness of development, and/or (3) termination of development before the average adult skills are achieved.[78,89] Furthermore, delay assumes that between the onset and termination of a particular child's development, the course of events will follow the developmental sequence observed with children who are acquiring speech normally.[90]

Answers to the following questions may supply useful insight into the problems: Did the child follow a normal pattern in sitting, walking, eating, talking, and using the toilet? Did the child have any periods of extreme illness or temporary hearing loss? Did the child have extended absences from the family environment? Did emotional conflicts arise between the mother and father or between the parents and the child? Did any undue penalties and frustrations occur that may have been associated with the speech readiness period? Did or do persons in the environment have similar speech problems that could have been imitated?

In summary, a child with a delay profile demonstrates performance similar to that of a normal but younger speaker. In contrast, deviant speech refers to an articulation profile unlike that of a normally developing child, even a younger one.[78] Although both appropriate and delayed characteristics may still persist in the speech of these children, the deviant speaker's articulation reveals error patterns that cannot be accounted for on the basis of a "delayed" concept.

The traditional way of describing the nature of articulation disorders is in terms of the following four classic error types:

Substitutions. One standard sound is substituted for another in the initial, medial, or final position within the word. For example, the sound *w* is frequently substituted for *r,* producing *wed* for *red; t* or *d* may replace *th,* and *mother* becomes *mudder* and *something* changes to *someting; th* may substitute for *s,* and *miss* becomes *mith.* The most frequently misarticulated sounds are often those that are mastered later in the course of speech development: */s, z, r, 1, th, ch, sh, zh/.*

Omissions. Sounds are omitted in a word and are not replaced by a substitute. The child may produce *tep* for *step, air* for *where,* or *daw* for *dog.*

Distortions. The target sounds are produced but are modified so that the perceptual result is inaccurate. A distortion might be considered a special case of substitution error in that a standard sound is replaced by a nonstandard sound. The /s/ and /z/ are among the most misarticulated speech sounds because adjustments of the tongue, teeth, and bite relationships during normal growth and development during the preschool and primary years are so exacting that minor deviations easily occur. A common form is lisping, when one or more of the sibilant consonants /s, z, *sh, zh/* is distorted.

Additions. Extra sounds are added that are not part of the word: examples are *warsh* for *wash* or *plass* for *pass.* Additions are not as common as substitutions, omissions, and distortions.

Organic Articulation Disorders in Children

Many speech-language pathologists now approach the subject of organic articulation disorders cautiously. As Bernthal and Bankson stated,

> Although individuals with oral structural deviations frequently experience articulation problems, the relationship between structural deficits and articulation skills is not very predictable. The literature cites many instances of individuals with structural anomalies who have developed appropriate compensatory gestures to produce acoustically acceptable speech. Why some individuals are able to compensate for relatively gross abnormalities and others are

unable to compensate for lesser deficits has not been resolved.[91]

These observations clearly emphasize that the relationship between oral structure and articulation proficiency is neither simple nor automatic. The remarkable compensatory potential of the oral-peripheral mechanism seems to account for the clinically observed inconsistencies in speakers whose performance is relatively better than would be expected from the individual's structural status. In general, structural deviations can affect articulation by interfering with articulatory contact points, breath stream flow (force and direction), oral breath pressure, and oral-nasal cavity coupling. Disorders such as cleft lip and palate can involve all of these aspects, whereas the normal loss of front teeth in a growing child might create involvement of only the first two factors. As Carrell pointed out, a good rule of thumb is to "consider any structural deviation a possible cause of misarticulation provided the nature of the speech defect is consistent with the deviation."[92]

Structural and Syndrome-Related Speech Disorders. Many structural deviations (eg, clefts of the lip and/or palate) are associated with speech, hearing, and language disorders. The physician should be sensitive to the possibility that structural anomalies concomitant to speech disorders can result from dysmorphisms of genetic origin, commonly referred to as syndromes, or a "pattern of things that run together."[3] When multiple dysmorphisms appear in the same patient, consultation with a medical geneticist is prudent to determine whether evaluation and treatment of the family as well as the patient are necessary. Shprintzen identified some syndromic patterns commonly associated with language and speech disorders. Contiguous gene disorders give rise to Prader-Willi, Beckwith-Wiedemann, velocardiofacial, and Rubinstein-Taybi syndromes.[3] Mechanically induced syndromes that cause deformations from tearing or mutilation of the embryo include sequences in which multiple anomalies stem from a single factor. Such sequences that result in speech and language disorders include ADAM sequence (amniotic deformations, adhesions, and mutilations—the craniofacial complex and limbs are particularly susceptible to amniotic adhesions resulting in cleft lip, facial clefts, and anencephaly), Pierre Robin syndrome (micrognathia, large U-shaped cleft palate, and upper airway obstruction), and Stickler's syndrome (clefting of palate, micrognathia, and perhaps Robin sequence). Other teratogenic disorders are associated with multiple anomalies. Examples include the effects of thalidomide, German measles virus, and fetal alcohol syndrome. Myotonic dystrophy (Steinert's disease), neurologic diseases associated with lysosomal storage diseases (Hunter's syndrome), cleft of the larynx (Opitz syndrome), and vocal cord atrophy (owing to Werner's syndrome) can have specific consequences on the voice and speech production capabilities of the patient. Although a discussion of these syndromes is well beyond the scope of this chapter, interested readers should consult the excellent references provided.[3,93–95]

Tongue-tie (Ankyloglossia). Parents frequently take children with articulatory defects to physicians to determine whether the child is "tongue-tied." The term "tongue-tied" (ankyloglossia) is used when the lingual frenum appears to be abnormally short or taut, restricting the movement of the tongue so that it cannot move upward to the alveolar ridge. Although this condition is known to exist, its interference with articulation is considered comparatively rare. McEvery and Gaines examined 1,000 children who had short frenums and found only 4 with articulatory defects.[96] They recommended that the frenum not be clipped because of possible infection, hemorrhage, and residual scar tissue. A more recent study confirmed the minimal relationship between ankyloglossia and articulation disorders, finding that an individual who presents a tongue-tie and an associated articulatory problem can usually be re-educated successfully without surgical removal or clipping of the frenum.[97] The exception is when the tip of the tongue is almost completely immobilized. Even in these cases, however, surgical intervention does not independently remedy the speech problem. The freedom of the tongue that results from clipping the frenum provides a potential for improvement, but the tongue habits that have developed will persist during speaking until the individual has been retrained.

Dental Abnormalities and Tongue Thrust. Normally, the teeth serve an important function in the production of speech sounds. Research and clinical observations have indicated, however, that many

individuals with dental abnormalities are able to make compensatory adjustments that produce acceptable speech. Professional debate related to the developmental nature of swallowing (normal and abnormal), the impact of tongue thrust on dentition and articulation, and the efficacy of speech and/or myofunctional therapy has created an ongoing controversy that is still unresolved. Review of the various positions in the literature led Bernthal and Bankson to the following conclusions[91]:

1. Currently, no data support the existence of a distinct clinical entity sometimes referred to as "tongue thrust." Behavior so identified is probably "normal," especially in preadolescent children.
2. Speech sound errors, particularly sibilant distortions, are present in greater number in those who evidence reverse swallow patterns than in those who do not.
3. The speech-language pathologist should focus on the correction of speech sound errors. On the basis of developmental data, techniques designed to change swallow patterns as part of an overall remediation program appear ill advised before a client has reached adolescence.

Treatment of Articulation Disorders in Children The complexity of speech production requires awareness of many factors that could influence speech sound learning/monitoring/modification and performance. One must evaluate the articulation problem in terms of (1) the degree of sound variation, (2) the consistency of misarticulation, (3) the effect of the disorder on intelligibility, and (4) the total number of sounds affected. This clinical strategy naturally results in capturing an essential component of any treatment plan; it must be prescriptive. The goal in treating articulation disorders is to identify and correct defective sounds. Most treatment is directed at improving performance accuracy and consistency. The four general steps of remediation are (1) training the speaker's perceptual skills (phonologic process or sound identification and discrimination), (2) establishing the new response (correct sound production), (3) strengthening and generalizing the correct sound production or phonologic process to connected speech levels, and (4) carrying over the new response to conversational speech both within and outside the clinical setting. The last three stages are really designed to habituate the new response.[78] Although the steps are straightforward, the complexity of the behavior and the clinical process dictate that treatment be guided by a professional speech-language pathologist rather than by the client or parents attempting a self-improvement program. Additionally, the parents should guard against continually correcting the child in his or her speech production. When proper therapy for articulatory problems is administered, the prognosis is good.

Disorders of Speech Motor Control

"Neuromotor speech disorders are disturbances of movement of the speech production system that result in someone being imperfectly understood or creating the impression that something is unusual or bizarre about his or her speech pattern."[98] The two major types of neuromotor speech disorders are dysarthria and apraxia.

Dysarthria Dysarthria may be defined as a defect of articulation resulting from a lesion in the central or peripheral nervous system, which directly regulates the muscles used for speaking. It involves an impairment in the control and execution of speech movements because of muscle weakness, slowness, incoordination, or altered muscle tone.[98] Depending on the type of damage to the neuromuscular control system, dysarthria can result in impaired respiration, phonation, articulation, resonance, and prosodic aspects of speech. Largely through the work of Darley and his colleagues at the Mayo Clinic in the 1970s, differential diagnosis of the types of dysarthrias can now be made.[99] Table 49–2 presents a list of recognizable dysarthria types along with the neurologic conditions or level of brain damage associated with them. Dysarthrias associated with Parkinson's disease[7,9,11–13,100,101] and amyotrophic lateral sclerosis[10,102–106] have probably received more research attention during the past 5 years than other types of speech disorders.[107,108]

Apraxia In contrast, apraxia of speech represents an impairment in the programming of speech movements in the absence of muscle impairments associated with dysarthria.[98] Apraxia is considered to involve damage to the speech programming area of the brain in the left frontal lobe (Broca's area).[109] With both

TABLE 49–2. Types of Dysarthrias

Dysarthria Type	*Neurogenic Condition or Level of Damage*	*Speech Deviations*
Flaccid	Brainstem, lower motoneuron	Hyponasality, breathiness, poor articulation
Spastic	Bilateral motor strip	Strained-strangled voice, low pitch, very poor articulation
Ataxic	Cerebellum	Irregular speech and syllable repetition
Hypokinetic	Basal ganglia, parkinsonism	Reduced loudness, rushes of speech
Hyperkinetic	Many levels of extrapyramidal motor system	Unsteady rate, pitch, loudness; sudden or slow variations in speech or voice; stoppages; tics; grunts

conditions, dysarthria and apraxia, articulation proficiency of the speaker is adversely affected. As would be expected, the degree to which speech intelligibility is compromised depends on the type and severity of the underlying neuromuscular impairment. Darley and associates suggested that a comparison of dysarthria and apraxia of speech reveals distinctive performance profiles.[99] According to their extensive study of these clinical entities, the "most characteristic error made by dysarthric patients is imprecise production of consonants, usually in the form of distortions and omissions." These are considered errors of "simplification." In contrast, the patient demonstrating apraxia of speech reveals errors of "complication" and relatively few simplification errors. These complication errors include "substitutions of other phonemes, often unrelated substitutions, as well as additions of phonemes, repetitions of phonemes, and prolongations of phonemes."

Treatment of the Motor Speech Disorders
Although distinctive articulation profiles can be described for both motor speech disorders, differential diagnosis between these two clinical entities is usually difficult to make with confidence. The prognosis for improvement with therapy for both of these motor speech disorders should be guarded, but judgment varies with the type and severity of the underlying neuromuscular impairment.[110] Because patients with either disorder may benefit from a formal, structured remediation program, rehabilitation should be provided. According to Darley and associates,[99] any treatment program should incorporate five fundamental principles: (1) the patient must be assisted to develop functional compensation from the healthy body systems; (2) the patient must develop the perspective that speech production must now become a highly conscious, deliberate effort; (3) the patient must develop the ability to monitor speech performance continuously; (4) remediation must begin as soon as possible because waiting is not beneficial; and (5) the patient must receive continued support and reassurance. In an attempt to enumerate various treatment options with neurogenic speech disorders, LaPointe and Katz included medical (alleviate cause), behavioral (modify neuromuscular and aerodynamic events), palliative (make the behavior, as well as the reaction to it, more moderate), and alternate mode (implement alternative communication systems) intervention strategies.[111] Typically, the speech-language pathologist uses one or all of the final three options in conjunction with the appropriate medical intervention.

Defects of Fluency (Stuttering)

The terms stuttering and stammering are used synonymously around the world, although the term stuttering is preferred in the United States. Stuttering is one of the most enigmatic speech defects encountered by the speech-language pathologist. This phenomenon is difficult to define because "no two stutterers are alike" and no single stutterer is the same from one time until the next. A workable def-

inition of stuttering that has been acceptable to all speech-language pathologists has not yet been formulated. Wingate, however, provided a useful description of the disorder that enumerates behaviors common to all stutterers and indicates kinds of accessory behaviors shown only by some:

> I. (a) Disruption in the fluency of verbal expression, which is (b) characterized by involuntary, audible or silent, repetitions or prolongations in the utterance of short speech elements, namely: sounds, syllables, and words of one syllable. These disruptions (c) usually occur frequently or are marked in character and (d) are not readily controllable. II. Sometimes the disruptions are (e) accompanied by accessory activities involving the speech apparatus, related or unrelated body structures, or stereotyped speech utterances. These activities give the appearance of being speech-related struggles. III. Also, there are not infrequently (f) indications or report of the presence of an emotional state, ranging from a general condition of "excitement" or "tension" to more specific emotions of a negative nature such as fear, embarrassment, irritation, or the like. (g) The immediate source of stuttering is some incoordination expressed in the peripheral speech mechanism; the ultimate cause is presently unknown and may be complex or compound.[112]

Although no conclusive prevalence figures currently exist, it has been estimated that stuttering has its highest occurrence in the preschool years (above 4%), declining thereafter to an unstable value of less than 1%. This declining pattern appears to relate to two factors, successful therapy and spontaneous recovery. Research has confirmed the clinical observation that stuttering is more common among boys than girls in a ratio of approximately 3 to 1.[113] The literature presents a variety of theories regarding the origin of stuttering; unfortunately, none represent professional consensus. Based on extensive case study, Van Riper concluded that one cannot determine or account for the onset of stuttering in terms of the conditions surrounding it for the vast majority of cases studied soon after onset.[114] At a basic physiology level, however, Van Riper stated that it seems reasonable to attribute stuttering to the "difficulty some children are bound to experience in mastering the synchronized timing of the motor coordinations required for speech." Van Riper's conclusion has current consensus in light of the emerging research evidence that differences in speech physiology do exist between stutterers and nonstutterers.[115–117] In fact, research into the nature of stutterers' *fluent* utterances indicates subtle motoric differences (not auditorily perceptible) that suggest some underlying difficulty with motor control and timing.[118] Substantial agreement exists on the general "facts" regarding the onset and development of stuttering. The typical onset occurs during the preschool years, usually 1 or 2 years after the child first learns to speak (onset is rare in adulthood). The onset is usually gradual and most often characterized by an excessive amount of repetitive speech (sounds or syllables), usually without tension or effort. Although the exact nature of development varies among stutterers, virtually all demonstrate a change of behavior as long as the disorder persists. The initial speech characteristics of part-word repetitions are often supplemented with sound prolongations and hesitations, usually associated with increasing struggle and avoidance behavior. Reactions by listeners, important "others," and the stutterer tend to create fear, anxiety, and self-doubt. These observations tend to confirm the accuracy and utility of Wingate's "definition" presented earlier. Moreover, the stability of these behaviors is not observed until the advanced stage of the disorder. Until that time, the severity of stuttering has been described as cyclic, changing from better to worse with no apparent logic. Despite differences in severity and symptom profile, there is clear evidence that as long as the disorder persists, the stutterer must face exposure of the disability to an ever-growing number of situations that come with social and professional growth in adulthood. Stuttering is truly a disorder of social living, not just a speech impediment.

Research regarding the development of speech and language skills makes it clear that most children experience periods of "normal nonfluency" during their early years that look and sound similar to the behaviors described as onset characteristics of stuttering. The number and duration of such episodes vary among children, but their occurrence does *not* automatically confirm the existence of stuttering. Therefore, great caution should be exercised to avoid the unnecessary stigmatizing effect of labeling a child's nonfluencies as clinically significant (stuttering) when they could be variations of *normal* speech development. Although differentiation between normal nonfluencies and those reflective of incipient stuttering is difficult, research has gener-

ated behavior guidelines to assist the speech-language pathologist in making these clinical distinctions.[114,119,120] Almost all stutterers have periods of relative freedom from hesitancies, repetitions, blocks, or prolongations. Stutterers can usually sing or speak in unison without disruption in their flow of speech and usually have no difficulty when speaking aloud to themselves or to a pet. The degree of communicative stress in a speaking situation appears to govern the degree of stuttering severity.

Common Questions about Stuttering What is primary stuttering? This term is sometimes used to describe a child's speech when it is marked by effortless repetitions or prolongations of words, phrases, or syllables without an awareness on the child's part that these mannerisms are different or abnormal. Johnson stated that "practically all stutterers are originally diagnosed (regarded as "stutterers," "not talking right," or "having difficulty saying words," and so forth) by … their parents, more often than not the mother being the first to become concerned."[121] Is stuttering hereditary? No clear evidence supports the view that stuttering is a characteristic that can be transmitted, in a biologic sense, from one generation to another. Does stuttering run in families? Research studies pertaining to the families of stutterers and nonstutterers indicate that, to a limited extent, stuttering does tend to run in families. The reason appears to involve tradition more than heredity. When parents have had a background of experience with stuttering, they appear to react to the speech imperfections of their children differently from parents who are unfamiliar with the condition. Parents with stuttering in their background may be so conditioned in attitude, policies, and concern that they view their child's normal speech imperfections as stuttering. Is stuttering caused by imitation? Parents are frequently concerned about the possibility of a child becoming a stutterer if their child associates with a stuttering child. Imitation in and of itself does not cause stuttering. If this were true, the incidence of stuttering would be considerably higher than that currently reported. Children have been exposed to stuttering for years in public and private schools, and no "epidemic" has ever been recorded.

Why do more boys stutter than girls? Studies have shown that male stutterers outnumber female stutterers by about 3 to 1. From a physical standpoint, the sex ratio might reflect a less stable neuromuscular control system for speech in boys, at least during the early years of development.

Treatment of Stuttering Parents of a young child who demonstrates minor hesitations and repetitions in his or her speech should remain unemotional about these speech patterns to prevent the development of undue awareness and concern on the part of the child. Being an attentive, thoughtful listener is the best response. The young child with incipient stuttering should not be subjected to direct speech therapy or any type of therapy related to speech production. According to Shames, the family is a critical factor in dealing with the young child because "the family can either reinforce or counteract the efforts of the speech-language pathologist."[122] An advanced, or secondary, stutterer should seek the assistance of a qualified speech-language pathologist. The stutterer or the family should be warned against treatment employing unethical procedures, such as gimmicks, devices, pills, or the guarantee of a "cure for stuttering." The treatment should always be conducted by individuals who have had academic training and experience in the areas of speech-language pathology, clinical psychology, or psychiatry. A multitude of different therapy strategies are currently available, and each reflects a theoretic view of the cause and nature of stuttering. Recent theories and treatment strategies have centered on fluency disorders as a motor speech disorder. In general, however, clinical management of stuttering involves changing the stutterer's attitude, method of talking, and/or environment. The advanced or secondary stutterer must perceive the need for therapy. Frequently, the family of the secondary stutterer is also in need of appropriate treatment.

DISORDERS OF SYMBOLIZATION (LANGUAGE)

Language may be defined as an organized symbolic representation of thought and action used as a means of communication on an abstract level by human beings. Therefore, preceding the act of speaking is the process of symbolization, which involves the comprehension and formulation of language. When this process is disturbed, the result may be classified as a communication disorder that is manifested in the inability, or limited ability, to use linguistic symbols as a means of oral communication. This difficulty is

common to many different types of children (including children with "specific language impairment," children with aphasia, mentally retarded children, children with autistic-like characteristics, and hearing-impaired children) and markedly influences normal language acquisition. Language disorders are also central to the adult aphasic patient.

Children with Specific Language Impairment

Children with specific language impairment generally have a disorder profile limited to the area of language. The particular difficulty may be in some or all of the following components of language: (1) lexicon (the concepts and labels of our vocabulary), (2) syntax (word order), (3) morphology (word forms), (4) semantics (word meaning), and (5) pragmatics (use of language in social contexts). According to Leonard's description of children with specific language impairment, "the most notable feature of these children's language is that the large majority of the linguistic features reflected in their speech are slow to emerge and to develop."[123] This is descriptive of a delay profile. Although no consensus exists on the causes underlying the problems of this group of children, causes most often cited are perceptual difficulties, cognitive development problems, environmental and personal interaction difficulties, and brain damage. In many respects, language intervention is similar to that used for other disorders of communication. A prescriptive program, based on comprehensive language testing, should be directed to the necessary area of oral language production and/or comprehension.

Children with Acquired Aphasia

The general term aphasia (dysphasia) refers to the difficulty experienced in the comprehension (receptive-sensory) and the use of linguistic symbols (expressive-motor) in verbal communication. A useful operational definition, presented by Myklebust, is as follows:

> Childhood aphasia refers to one or more significant deficits in essential processes as they relate to facility in use of auditory language. Children having this disability demonstrate a discrepancy between expected and actual achievement in one or more of the following functions: auditory perception, auditory memory, integration, comprehension, and expression. The deficits referred to are not the result of sensory, motor, intellectual, or emotional impairment, or of the lack of opportunity to learn. They are assumed to derive from dysfunctions in the brain, though the evidence for such dysfunctioning may be mainly behavioral, rather than neurological, in nature.[124]

Leonard emphasized that children with acquired aphasia differ from other language-disordered children in that they have an initial period of normal language development.[123] They then lose that ability, usually because of some known brain damage caused by either illness or trauma. The most prominent disorder feature has been described as a noticeable reduction in the amount of speech generated, sometimes approaching mutism.[125] Apraxia of speech and word retrieval problems are not uncommon; however, comprehension deficits are less severe. In general, language recovery is usually more rapid and complete than for adult aphasics, particularly when focusing on production-side symptoms. The age of damage onset and location of the lesion are two critical factors that dictate the impairment profile and the prognosis. For example, if the damage occurs before age 9 and is limited to one cerebral hemisphere, the child often demonstrates normal or near-normal language development and performance. In general, the best prognosis is associated with a discrete lesion of late onset, whereas the worst prognosis emerges with a widespread lesion early in life. Parents frequently seek professional assistance after their child has been "diagnosed" erroneously as mentally retarded, deaf, or emotionally disturbed. The child may appear mentally retarded because of poor performance on standardized intelligence tests; the label deaf may be applied because of an inability to use verbal symbols that require hearing for response. The emotionally disturbed classification may seem appropriate because of exhibited behavioral problems, yet more inclusive evaluation of all of the data usually differentiates several problems and identifies the aphasic child. Diagnosing aphasia in children is a complex process requiring the united efforts of many specialists, including the neurologist, pediatrician, psychologist, otologist, audiologist, speech-language pathologist, and educational consultant. Only through the combined efforts of these specialists, using a systematic, diag-

nostic approach, will the child be led to a more prescriptive type of educational training.

Because many aphasic children recover quickly after head injury or illness, the role of the speech-language pathologist and other professionals may be minimal. When residual deficits persist, however, a formal remediation program appears justified.[126] Holland and Reinmuth remind us that for an unfortunate minority of these children, the prognosis is not good.[127] Extensive residual problems require intensive, specialized, interdisciplinary remediation available only through special rehabilitation centers or special schools. In those settings, "restitution of linguistic, cognitive and physical skills is the primary goal." The diagnosis and prescribed training program for the aphasic child would be woefully incomplete if the emotional factors of the family environment were not considered. The family may protect or shield the child excessively, or else coerce or reject the child. The presence of an aphasic child in the family unit may be deeply disturbing; the situation requires a thorough appraisal of both the family structure and the need for therapeutic counseling.

Aphasia (Dysphasia) in Adults

Aphasia is associated with cortical disturbances or lesions resulting from vascular impairment (thrombosis, embolism, and hemorrhage), tumors, degenerating and infectious disease, and trauma. Holland and Reinmuth stated that aphasia is a general term applied to different but related syndromes that impair the ability to formulate, retrieve, and/or decode the symbols of language.[127] As a common adult disorder of communication, it disrupts speech and verbal output, comprehension of speech, and reading and writing skills. The onset of aphasia is usually abrupt and typically occurs in adults with no previous speech or language difficulty. According to Holland and Reinmuth, most syndromes of aphasia share several concepts:

1. Aphasic symptoms are not bizarre but rather are extreme extensions of language problems demonstrated daily by normal speakers (ie, losing one's train of thought, forgetting how to spell a word, not understanding a spoken word or concept).
2. All aphasics have some fundamental difficulty with both comprehension and word retrieval, regardless of the presence of other deficits.
3. There is a continuum of severity within each syndrome.

The linguistic disturbances of the aphasic have been classified in numerous ways that reflect the clinical experience of the classifiers and their concepts of cortical function related to language.[128] These classifications have also attempted to recognize the major types and specific forms of language disturbance. One classification system presently in wide use is the one developed by a Boston-based group of professionals that highlights fluent versus nonfluent types.[129] The six options were summarized by Boone as follows[70]:

1. Broca's or motor aphasia—marked reduction in speech output, common oral-verbal apraxia, relatively good speech comprehension, word retrieval problem, writing performance matches speech output
2. Transcortical motor aphasia—a fairly rare type with significant struggle in producing an utterance, telegraphic speech is common, relatively good speech comprehension, impaired writing skills
3. Anomic aphasia—word retrieval problems predominate for both speech and writing, nearly normal comprehension for speech and reading
4. Wernicke's or jargon aphasia—significant comprehension deficit for speech and reading, fluent but jargon-like verbal output, writing parallels speech patterns
5. Conduction aphasia—nearly normal comprehension for speech and reading, marked deficit in speech repetition, fluent speech output with numerous sound and word errors
6. Global aphasia—the most severe and most common type; profound problems with both verbal output and comprehension, and both reading and writing functions are poor

The emotional language of the adult aphasic is usually better than his or her propositional language. The patient may find it easier to swear, count, or use other forms of automatic speech or nonpropositional forms of speech but is at a loss when requested to develop this emotional language into abstract or propositional language situations. The patient experiences difficulty in combining simple linguistic symbols into more complex linguistic units. The patient is not without words but rather cannot

quickly command the response that is appropriate to the situation. Boone emphasized that the brain damage (regardless of its cause) can produce other deficits or symptoms.[70] These include hemiplegia (gross motor paralysis), hemianopsia (visual field defect), intellectual deficits, seizures, dysarthria (motor speech impairment), and apraxia (inability to execute voluntary movement).

Prognosis for the recovery of the aphasic patient must be based on many factors, including the site and extent of the lesion, general health, attitude toward self and environment, age and educational attainment, vocation and avocations, family attitude, and degree of cooperation. Assuming that these and other factors are relatively positive, considerable recovery may be expected from the overall linguistic and physical impairments. Spontaneous improvement may be noted during the first few months after the onset without structured training. This may give false hope to the patient and family for additional recovery without seeking professional assistance. Depending on the degree of involvement, the family, with the physician's guidance, should enlist the services of a physical therapist, occupational therapist, speech-language pathologist, psychologist, and vocational counselor in formulating a personalized rehabilitation program. Ideally, the adult aphasic should be started on a program of rehabilitation as soon after the traumatic episode as possible.

REFERENCES

1. Ludlow CL. Research needs for the assessment of phonatory function. In: Ludlow CL, Hart MO, editors. Proceedings of the Conference on the Assessment of Vocal Pathology. ASHA Reports No. 11. Rockville (MD): American Speech-Language-Hearing Association; 1981.
2. Verdolini K. Voice disorders. In: Tomblin JB, Morris HL, Spriestersbach DC, editors. Diagnosis in speech-language pathology. San Diego (CA): Singular Publishing Group; 1994. p. 253–4.
3. Shprintzen RJ. Syndrome delineation and communicative impairment. In: Minifie FD, editor. Introduction to communication sciences and disorders. San Diego (CA): Singular Publishing Group; 1994. p. 439–80.
4. Bastian RW, Keidar AK, Verdolini-Marston K. Simple vocal tasks for detecting vocal fold swelling. J Voice 1990;4:172–83.
5. Damste PH. Virilization of the voice due to anabolic steroids. Folia Phoniatr Logop 1964;16:10.
6. Damste PH. Voice change in adult women caused by virilizing agents. J Speech Hear Disord 1967;32:126.
7. Ramig LO, Scherer RC, Klasner ER, et al. Acoustic analysis of voice in amyotrophic lateral sclerosis. J Speech Hear Disord 1990;55:2–14.
8. Hanson K, Gerratt B, Ward R. Cinegraphic observations of laryngeal function in Parkinson's disease. Laryngoscope 1984;94:348.
9. Coutryman MA, Ramig LO, Pawlas AA. Speech and voice deficits in parkinsonism plus syndromes: can they be treated? In: NCVS Status and Progress Report 1994;6:99.
10. Ramig LA, Scherer RC, Titze IR, Ringel SP. Acoustic characteristics of voice in amyotrophic lateral sclerosis: a longitudinal study. Ann Otol Rhinol Laryngol 1988;97:164–72.
11. King JB, Ramig LO, Lemke JH, Horii Y. Parkinson's disease: longitudinal changes in acoustic parameters of phonation. J Med Speech Lang Pathol [In press].
12. Strand E, Yorkston K. Description and classification of dysarthric individuals: toward a new taxonomy. In: Till J, Beukelman D, Yorkston K, editors. Clinical dysarthria. Baltimore (MD): Paul H. Brookes; 1994.
13. Ramig LO. Speech therapy for Parkinson's disease. In: Koller W, Paulson G, editors. Therapy of Parkinson's disease. New York: Marcel Dekker; 1994.
14. Orlikoff RF, Huang ZD. Influence of vowel production on acoustic and electroglottographic perturbation measures. Presented at the American Speech-Language-Hearing Association Convention, Atlanta, GA, 1991.
15. Huang ZD, Minifie FD, Kasuya H, Lin XS. Measures of vocal function during changes in vocal effort level. J Voice 1995;9:429–38.
16. Koike Y. Application of some acoustic measures for the evaluation of laryngeal dysfunction. Studia Phonol 1973;7:17.
17. Yoshida M. Study on perceptive and acoustic classification of pathological voices. Pract Otol (Kyoto) 1979;72:249.
18. Sawashima M, Hirano M. Clinical evaluation of voice disorders. Ann Bull Res Inst Logop Phoniatr 1981;15:164.
19. Yumoto E, Gould WJ, Baer T. Harmonics-to-noise ratio as an index to harshness. J Acoust Soc Am 1982;71:1544–9.

20. Yumoto E, Sasaki Y, Okamura H. Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. J Speech Hear Res 1984;27:2–6.
21. Zwitman DH. Bilateral cord dysfunctions: abduction type spastic dysphonia. J Speech Hear Disord 1979;44:373–8.
22. Freeman F, Cannito M, Finitzo-Hieber T. Classification of spasmodic dysphonia by perceptual-acoustic means. In: Gates GA, editor. Spastic dysphonia: state of the art 1984. New York: The Voice Foundation; 1985.
23. Brodnitz FS. Spastic dysphonia. Ann Otol Rhinol Laryngol 1975;85:210–4.
24. Spiegel H. Psychiatric aspects of spasmodic dysphonia. In: Gates GA, editor. Spastic dysphonia: state of the art 1984. New York: The Voice Foundation; 1985.
25. Aronson AE, Brown JP, Litin EM, Pearson JS. Spastic dysphonia: I. Voice, neurologic and psychiatric aspects. J Speech Hear Disord 1968;33:203–18.
26. Feldman M, Nixon J, Finitzo-Hieber T, Freeman F. Abnormal parasympathetic vagal function in patients with spasmodic dysphonia. Ann Intern Med 1984;100:491–5.
27. Dedo HH, Townsend JJ, Izdebski K. Current evidence for the organic etiology of spastic dysphonia. Otolaryngology 1978;86:875–80.
28. Izdebski K, Shipp T. Model of spastic dysphonia. In: Gates GA, editor. Spastic dysphonia: state of the art 1984. New York: The Voice Foundation; 1985.
29. Fink M. Neurobiology of spastic disorders. In: Gates GA, editor. Spastic dysphonia: state of the art 1984. New York: The Voice Foundation; 1985.
30. Malmgran L. Neuromuscular anatomy of the larynx. In: Gates GA, editor. Spastic dysphonia: state of the art 1984. New York: The Voice Foundation; 1985.
31. Bocchino JV, Tucker HM. Recurrent laryngeal nerve pathology in spasmodic dysphonia. Laryngoscope 1978;88:1274–8.
32. Dedo HH, Shipp T. Spastic dysphonia—a surgical and voice therapy treatment program. Houston: College-Hill Press; 1980.
33. Freeman F, Cannito MP, Finitzo-Hieber T, Ross E. Spasmodic dysphonia: myths and facts. In: Rosenfield D, editor. Speech news. Houston: Baylor College of Medicine; 1984.
34. Dedo HH. Recurrent laryngeal nerve section for spastic dysphonia. Ann Otol Rhinol Laryngol 1976;85:451–9.
35. Barton TT. Treatment of spastic dysphonia by recurrent laryngeal nerve section. Laryngoscope 1979; 89:244.
36. Levine HL, Wood BG, Batza E, et al. Recurrent laryngeal nerve section for spasmodic dysphonia. Ann Otol Rhinol Laryngol 1979;88:527–30.
37. Biller HF, Som MI, Lawson W. Laryngeal nerve crush for spastic dysphonia. Ann Otol Rhinol Laryngol 1979;88:531–2.
38. Miller RH, Woodson GE, Jankovic J, Izdebski R. Botulinum toxin injection of the vocal fold for spasmodic dysphonia. Arch Otolaryngol Head Neck Surg 1987;113:603–5.
39. Dedo HH, Izdebski K. Intermediate results of 306 recurrent laryngeal nerve sections for spastic dysphonia. Laryngoscope 1983;93:9–16.
40. Aronson AE, DeSanto LW. Adductor spasmodic dysphonia: three years after recurrent nerve section. Laryngoscope 1983;93:1–8.
41. Wilson FB, Oldring DJ, Mueller K. Recurrent laryngeal nerve dissection: a case report involving return of spastic dysphonia after initial surgery. J Speech Hear Disord 1976;45:112–8.
42. Fritzell B, et al. Experiences with recurrent laryngeal nerve section for spastic dysphonia. In: Phoniatric and logopedic progress report no. 3. Stockholm: Department of Logopedics and Phoniatrics, Huddinge University Hospital, Karolinska Institute; 1981.
43. Hartman DE, Aronson AE. Clinical investigations of intermittent breathy dysphonia. J Speech Hear Disord 1981;46:428–32.
44. Blitzer A. Electromyographic findings in spastic dysphonia. In: Gates GA, editor. Spastic dysphonia: state of the art 1984. New York: The Voice Foundation;1985.
45. Robe E, Brumlik J, Moore P. A study of spastic dysphonia: neurologic and electroencephalographic abnormalities. Laryngoscope 1960;70:219.
46. Schaefer S, Freeman F, Finitzo T, et al. Magnetic resonance imaging findings and correlations in spasmodic dysphonia patients. Ann Otol Rhinol Laryngol 1985;94:595–601.
47. Finitzo-Hieber T, Freeman FJ, Gerling IJ, et al. Auditory brain stem response abnormalities in adductor spasmodic dysphonia. Am J Otolaryngol 1982;3: 26–30.
48. Boone DR, McFarland SC. The voice and voice therapy. Englewood Cliffs (NJ): Allyn & Bacon; 1999.
49. Lowry LD, Marks JE, Powell WJ. 260 laryngeal carcinomas. Arch Otolaryngol 1973;98:147–53.
50. Silverberg E. Cancer statistics, 1981. Reprinted by American Cancer Society from CA Cancer J Clin 1981;37:13–28.

51. Cancer facts and figures. New York: American Cancer Society; 1988.
52. Cancer facts and figures. Atlanta: American Cancer Society; 1994.
53. American Cancer Society. Cancer facts and figures. Year 2000. Available at: www3.cancer.org (accessed Nov 2000).
54. Gates GA, Ryan W, Cooper JC, et al. Current state of laryngectomee rehabilitation: I. Results of therapy. Am J Otolaryngol 1982;3:1–7.
55. Snidecor JC. The family of the laryngectomee. In: The family as supportive personnel in speech and hearing remediation. Proceedings of a postgraduate short course. Santa Barbara (CA): University of California, Santa Barbara, US Department of Health, Education and Welfare; 1971. p. 12–24.
56. Gates GA, Ryan W, Lauder E. Current status of laryngectomee rehabilitation: IV. Attitudes about laryngectomee rehabilitation should change. Am J Otolaryngol 1982;3:97–103.
57. Gates GA, Ryan W, Cantu E, Hearne E. Current status of laryngectomee rehabilitation: II. Causes of failure. Am J Otolaryngol 1982;3:8–14.
58. Ryan W, Gates GA, Cantu E, Hearne E. Current status of laryngectomee rehabilitation: III. Understanding of esophageal speech. Am J Otolaryngol 1982; 3:91–6.
59. Ignorance of laryngectomee care confirmed by medical authorities. The IAL News 1982;27:June.
60. Singer MI, Blom ED. An endoscopic technique for restoration of voice after laryngectomy. Ann Otol Rhinol Laryngol 1980;89:529–33.
61. Zwitman DH, Calcaterra TC. Phonation using the trachea-esophageal shunt after total laryngectomy. J Speech Hear Disord 1973;38:368–73.
62. Weinberg B. Airway resistance to the voice button. Arch Otolaryngol 1982;108:498–500.
63. Singer MI, Blom ED, Hamaker RC. Further experience with voice restoration after total laryngectomy. Ann Otol Rhinol Laryngol 1981;90:498–502.
64. Singer MI, Blom ED. A selective myotomy for voice restoration after total laryngectomy. Arch Otolaryngol 1981;107:670–3.
65. Blom ED, Singer MI, Hamaker RC. Tracheostoma valve for postlaryngectomy voice rehabilitation. Ann Otol Rhinol Laryngol 1982;91:576–8.
66. Andrews J, Mickel RA, Hanson DG, et al. Major complications following tracheoesophageal puncture for voice rehabilitation. Laryngoscope 1987; 97:562–7.
67. Sisson G, McConnell FM, Logemann JA, Yeh S. Voice rehabilitation after laryngectomy. Arch Otolaryngol 1975;101:178.
68. Gardner WH. Laryngectomee speech and rehabilitation. Springfield (MA): Charles C. Thomas; 1971.
69. Lauder E. The laryngectomee and the artificial larynx —a second look. J Speech Hear Disord 1970;35:62–5.
70. Boone D. Human communication and its disorders. Englewood Cliffs (NJ): Prentice Hall; 1987.
71. Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. Speech perception in infants. Science 1971;171:303–6.
72. Kuhl PK. Speech perception. In: Minifie FD, editor. Introduction to communication sciences and disorders. San Diego (CA): Singular Publishing Group; 1994. p. 77–148.
73. Gopnik A, Meltzoff AN, Kuhl PK. The scientist in the crib: minds, brains, and how children learn. New York: William Morrow; 1999.
74. Werker JF, Tees RC. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. Infant Behav Dev 1984;7:49.
75. Kuhl PK, Williams KA, Lacerda F, et al. Linguistic experience alters phonetic perception in infants by 6 months of age. Science 1992;255:606–8.
76. Greiser DL, Kuhl PK. Maternal speech to infants in a tonal language: support for universal prosodic features in motherese. Dev Psychol 1988;24:14.
77. Bloom L. Language development. In: Minifie FD, editor. Introduction to communication sciences and disorders. San Diego (CA): Singular Publishing Group; 1994. p. 189–235.
78. Stone JR, Stoel-Gammon C. Phonological development and disorders in children. In: Minifie FD, editor. Introduction to communication sciences and disorders. San Diego (CA): Singular Publishing Group; 1994. p. 149–87.
79. Lewis MM. How children learn to speak. New York: Basic Books; 1957.
80. Locke JL. Phonological acquisition and change. New York: Academic Press; 1983.
81. Oller DK, Eilers RE. The role of audition in infant babbling. Child Dev 1988;59:441–9.
82. Yoshinaga-Itano C. Language development of deaf and hard of hearing children. San Diego (CA): Singular Publishing Group; 1999.
83. Sander EK. When are speech sounds learned? J Speech Hear Disord 1972;37:55–63.
84. Kent RD. Anatomical and neuromuscular maturation of the speech mechanism: evidence from acoustic studies. J Speech Hear Res 1976;19:421–47.

85. Ingram D. Phonological disability in children. London: Edward Arnold; 1976.
86. Dyson AT. Phonetic inventories of 2- and 3-year-old children. J Speech Hear Disord 1988;53:89–93.
87. Crary MA, Fokes J. Phonological processes in apraxia of speech: a systemic simplification of articulatory performance. Aphasia Apraxia Agnosia 1980; 4:1.
88. Grumwell P. Clinical phonology, 2nd ed. Baltimore: Williams & Wilkins; 1987.
89. Smit AB. Speech sound disorders. In: Tomblin JB, Morris HL, Spriestersbach DC, editors. Diagnosis in speech-language pathology. San Diego (CA): Singular Publishing Group; 1994. p. 179–99.
90. Stoel-Gammon C, Dunn C. Normal and disordered phonology in children. Austin (TX): Pro-Ed; 1985.
91. Bernthal JE, Bankson NW. Articulation and phonological disorders. 2nd ed. Englewood Cliffs (NJ): Prentice-Hall; 1988.
92. Carrell J. Disorders of articulation. Englewood Cliffs (NJ): Prentice-Hall; 1968.
93. Cohen MM. The child with multiple birth defects. New York: Raven Press; 1982.
94. Fraser FC. The genetics of cleft lip and cleft palate. Am J Hum Genet 1970;22:336–52.
95. Siegel-Sadewitz VL, Shprintzen RJ. The relationship of communication disorders to syndrome identification. J Speech Hear Disord 1982;47: 338–54.
96. McEvery ET, Gaines FP. Tongue-tie in infants and children. J Pediatr 1941;18:252.
97. Fletcher SG, Meldrum JR. Lingual function and relative length of the lingual frenulum. J Speech Hear Res 1968;11:382–90.
98. LaPointe LL. Neurogenic disorders of communication. In: Minifie FD, editor. Introduction to communication sciences and disorders. San Diego (CA): Singular Publishing Group; 1994.
99. Darley F, Aronson A, Brown J. Motor speech disorders. Philadelphia (PA): WB Saunders; 1975.
100. Hartelius L, Svensson P. Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: a survey. Folia Phoniatr Logop 1991;46:9–17.
101. Yorkston KM, Hammon VL, Beukelman DR, Traynor CD. The effect of rate on the intelligibility and naturalness of dysarthric speech. J Speech Hear Disord 1990;55:550–60.
102. Hartelius L, Nord L, Buder EH. Acoustic analysis of dysarthria associated with multiple sclerosis. Clin Linguist Phonet 1994;8:143.
103. Kent JF, Kent RD, Rosenbek JC, et al. Quantitative description of the dysarthria in women with amyotrophic lateral sclerosis. J Speech Hear Res 1992;35: 723–33.
104. Kent RD, Kent JF, Weismer G, et al. Impairment of speech intelligibility in men with amyotrophic lateral sclerosis. J Speech Hear Disord 1990;55:721–8.
105. Kent RD, Sufit RL, Rosenbek JC, et al. Speech deterioration in amyotrophic lateral sclerosis: a case study. J Speech Hear Res 1991;34:1269–75.
106. Weismer G, Martin R, Kent RD, Kent JF. Formant trajectory characteristics of males with amyotrophic lateral sclerosis. J Acoust Soc Am 1992;91:1085–98.
107. Cannito MP, Marquardt TP. Ataxic dysarthria. In: McNeil MR, editor. Clinical management of sensorimotorspeech disorders. New York: Thieme; 1997.
108. Kent RD, Kent JF, Duffy JR, et al. Ataxic dysarthria. J Speech Lang Hear Res 2000;43:1275–89.
109. Kent RD. Neurological bases of communication disorders. In: Minifie FD, editor. Introduction to communication sciences and disorders. San Diego (CA): Singular Publishing Group; 1994. p. 313–50.
110. Duffy JR. Motor speech disorders: substrates, differential diagnosis, and management. St. Louis: Mosby Year Book; 1995.
111. LaPointe LL, Katz RC. Neurogenic disorders of speech. In: Shames G, Wiig E, Secord WB, editors. Human communication disorders. 4th ed. Columbus (OH): Charles E. Merrill; 1994. p. 480–518.
112. Wingate M. A standard definition of stuttering. J Speech Hear Disord 1964;29:484.
113. Prins D. Fluency and stuttering. In: Minifie FD, editor. Introduction to communication sciences and disorders. San Diego (CA): Singular Publishing Group; 1994. p. 521–59.
114. Van Riper C. The nature of stuttering, 2nd ed. Englewood Cliffs (NJ): Prentice-Hall; 1982.
115. Starkweather CW. Fluency and stuttering. Englewood Cliffs (NJ): Prentice-Hall; 1987.
116. Hillman RE, Gilbert HR. Voice onset times for voiceless stop consonants in fluent reading of stutterers and nonstutterers. J Acoust Soc Am 1977;61:610–1.
117. Metz DE, Conture EG, Caruso A. Voice onset time, frication and aspiration during stutterers' fluent speech. J Speech Hear Res 1979;22:649–56.
118. Zimmerman G. Articulatory dynamics of fluent utterances of stutterers and nonstutterers. J Speech Hear Res 1980;23:95.
119. Ingham RJ. Stuttering and behavior therapy. San Diego (CA): College Hill Press; 1984.

120. Adams MR. A clinical strategy for differentiating the normal nonfluent child and the incipient stutterer. J Fluency Disord 1977;2:141.
121. Johnson W. People in quandaries. New York: Harper and Row; 1946.
122. Shames G. Disorders and fluency. In: Shames G, Wiig E, editors. Human communication disorders. 3rd ed. Columbus (OH): Charles E. Merrill; 1990.
123. Leonard L. Early language development and language disorders. In: Shames G, Wiig E, editors. Human communication disorders. 2nd ed. Columbus (OH): Charles E. Merrill; 1986.
124. Myklebust H. Childhood aphasia: an evolving concept. In: Travis LE, editor. Handbook of speech pathology and audiology. Englewood Cliffs (NJ): Prentice-Hall; 1971.
125. Chase R. Neurological aspects of language disorders in children. In: Irwin J, Marge M, editors. Principles of childhood language disabilities. Englewood Cliffs (NJ): Prentice-Hall; 1972.
126. Wood NE. Delayed speech and language development. Englewood Cliffs (NJ): Prentice Hall; 1964.
127. Holland A, Reinmuth O. Aphasia in adults. In: Shames G, Wiig E, editors. Human communication disorders. 2nd ed. Columbus (OH): Charles E. Merrill; 1986.
128. Chapey R. Language intervention strategies in adult aphasia. 3rd ed. Baltimore: Williams & Wilkins; 1994.
129. Goodglass H, Kaplan E. The asssessment of aphasia and related disorders. 2nd ed. Philadelphia: Lea & Febiger; 1983.

CHAPTER 50

Airway Control and Laryngotracheal Stenosis in Adults

Joseph C. Sniezek, MD, Brian B. Burkey, MD

The successful management of the airway is of paramount importance. Airway control requires a logical and systematic approach guided by the principles of basic and advanced life support techniques and a working knowledge of relevant pharmacology. Most importantly, the health professional must possess a mastery of the anatomy and physiology of the upper aerodigestive tract. The airway surgeon is also uniquely responsible for the diagnosis and treatment of airway lesions such as laryngotracheal stenosis and arytenoid fixation.

ANATOMY

Successful airway intervention necessitates an understanding of the anatomy of the laryngotracheal complex (Figure 50–1). The hyoid bone allows muscular attachment from the lingual muscles and mandible as well as the extrinsic muscles of the larynx and serves as a laryngeal anchor via the thyrohyoid membrane. This membrane is pierced bilaterally by the superior laryngeal neurovascular bundles. Embryologically, the hyoid bone is derived from mesodermal cells in the second and third branchial arches. The greater horn of the hyoid bone is derived from the third branchial arch, whereas the lesser horn is of second branchial arch origin.

The major cartilages of the larynx include the thyroid, cricoid, epiglottis, and paired arytenoids (Figure 50–2). These cartilages develop from mesenchymal tissue originating in the fourth, fifth, and sixth branchial arches. The thyroid cartilage is a shield-shaped structure whose central prominence tends to be more acutely angled in men than in women. The inferior horns of the thyroid cartilage articulate with the posterior facets of the cricoid cartilage and form the important cricothyroid joints, which are synovial and allow a hinge motion between the cartilages.

The cricoid cartilage is the only complete ring of the airway, which accounts for the association between prolonged or traumatic endotracheal intubation and subglottic stenosis occurring in this area. The cricoid is composed of hyaline cartilage that often calcifies later in adult life. It tends to be broader posteriorly and tapers to a smaller arch anteriorly, mimicking the shape of a signet ring. The thyroid and cricoid cartilages are attached anteriorly by the cricothyroid membrane.

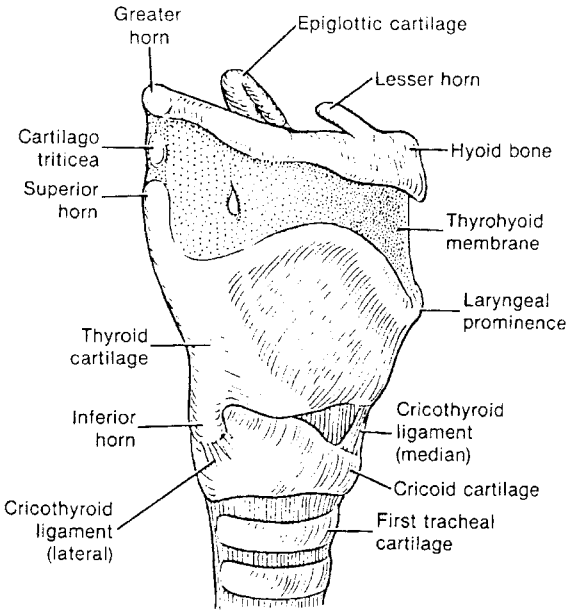


FIGURE 50–1. Lateral view of the larynx. Reproduced with permission from Graney DO, Flint PW. Anatomy. In: Cummings CW, et al, editors. Otolaryngology-head and neck surgery. 2nd ed. St. Louis: Mosby Year Book; 1993. p. 1693.

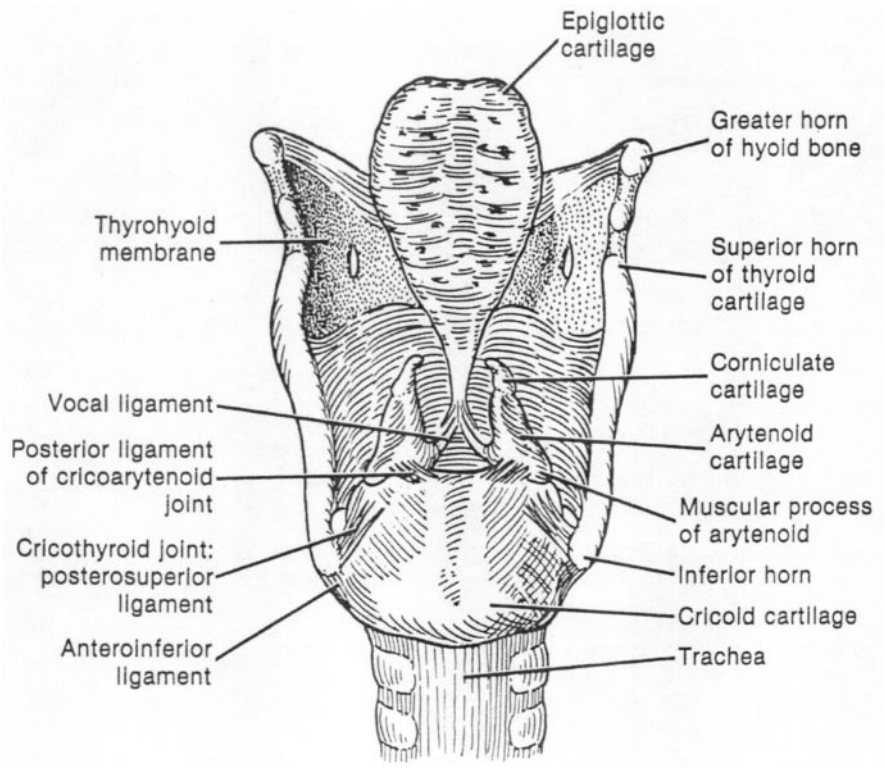


FIGURE 50–2. Posterior view of the larynx. Reproduced with permission from Graney DO, Flint PW. Anatomy. In: Cummings CW, et al, editors. Otolaryngology-head and neck surgery. 2nd ed. St. Louis: Mosby Year Book; 1993. p. 1694.

The epiglottis serves a protective function for the airway and is composed of elastic cartilage. It attaches to the thyroid cartilage via the thyroepiglottic ligament (Figure 50–3). The arytenoid cartilages provide the posterior attachments of the vocal cords, the anterior attachment being on the thyroid cartilage at Broyle's ligament. The arytenoids articulate with facets on the posterosuperior aspect of the cricoid cartilage and form synovial joints.

The trachea begins at the lower border of the cricoid cartilage and is composed of incomplete, C-shaped "rings" of hyaline cartilage. The rings are open posteriorly and connected on this aspect by fibroareolar connective tissue and smooth muscle fibers. The larynx and trachea are lined with a pseudostratified, ciliated columnar epithelium interspersed with mucus-secreting cells, except at the glottis, which is lined with a stratified squamous epithelium. The trachea bifurcates at approximately the sternal angle, forming the two bronchi. The right bronchus is wider, shorter, and more vertically oriented than the left bronchus. When using an open bronchoscope, one may most easily examine both bronchi by passing the instrument through the right side of the mouth, thus allowing the best view of the obliquely angled left bronchus while preserving visualization into the right bronchus.[1]

CLINICAL EVALUATION OF THE AIRWAY

The clinical evaluation of the airway must begin with a rapid assessment of the patient's ventilatory and respiratory status. If the patient is in impending airway distress, prophylactic measures to prevent deterioration of the airway must be employed. If the patient does not appear to be ventilating adequately, the airway must be definitively secured through either medical or surgical techniques. If medical control of the airway fails and intubation is impossible owing to an airway injury or upper airway obstruction, an urgent tracheostomy or cricothyroidotomy must be performed. The technique and advantages of each are discussed in this chapter. The symptoms and signs that indicate upper respiratory obstruction include hoarseness; dyspnea; stridor; intercostal, suprasternal, and supraclavicular retractions; restlessness; cyanosis; and drooling. The signs of trauma to the airway include bloody sputum, subcutaneous emphysema, and palpable laryngotracheal fractures.

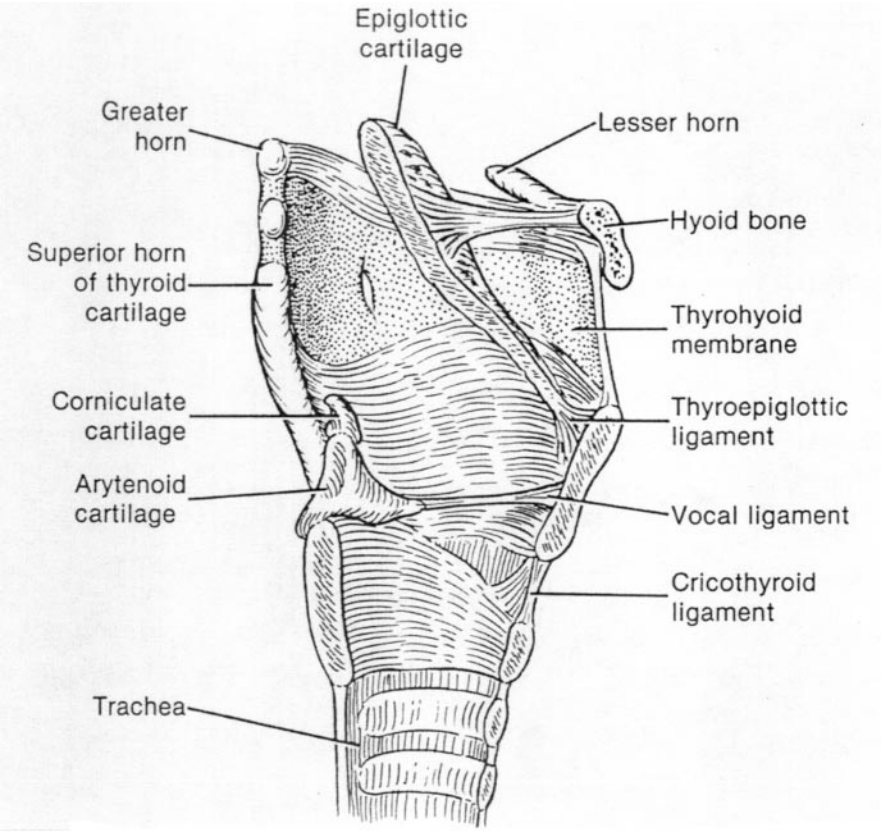


FIGURE 50–3. Sagittal section of the larynx. Reproduced with permission from Graney DO, Flint PW. Anatomy. In: Cummings CW, et al, editors. Otolaryngology-head and neck surgery. 2nd ed. St. Louis: Mosby Year Book; 1993. p. 1694.

If the patient does not require an immediate airway intervention, the caregiver may proceed with a diagnostic assessment. This assessment should include a complete history and a physical examination of the upper aerodigestive tract. If the patient is seen in the clinic setting, a mirror indirect laryngoscopy may be performed. Perhaps the most valuable study is a direct fiber-optic examination via nasopharyngoscopy. This technique allows unparalleled visualization of the upper airway and larynx and may allow for video documentation (Figure 50–4). The fiber-optic endoscopes and light sources are now portable and may be used in the emergency department, intensive care unit, and nursing unit. A blood gas determination, chest radiograph, and airway evaluation with radiography or computed tomography may also be indicated.

While remaining cognizant of the general status of the patient, the surgeon must accurately assess the nature and level of the airway obstruction and establish airway control below the lowest level of the obstruction. Once the airway has been controlled, a complete airway evaluation may be performed in the operating room through direct laryngoscopy and tracheoscopy. This mission must be accomplished while respecting the possibility of other injuries such as cervical spine dislocation, vascular injuries, or other associated medical conditions.

AIRWAY MANAGEMENT

Although the management of the airway in the controlled environment of the operating room is straightforward, patients presenting in the emergency department pose unique problems such as cervical spine injuries, closed head injuries with the possibility of increased intracranial pressure, laryngotracheal disruption, airway hemorrhage, and facial deformities. The intensive care unit setting also offers challenges such as confounding medical issues. These complicating factors can be handled only through an individualized approach that is guided by a stepwise progression of techniques aimed at controlling the increasingly difficult airway.

Supplemental oxygen is the most simple measure of airway support and should be given almost universally to patients with airway distress. After this intervention, the clinician should evaluate for airway

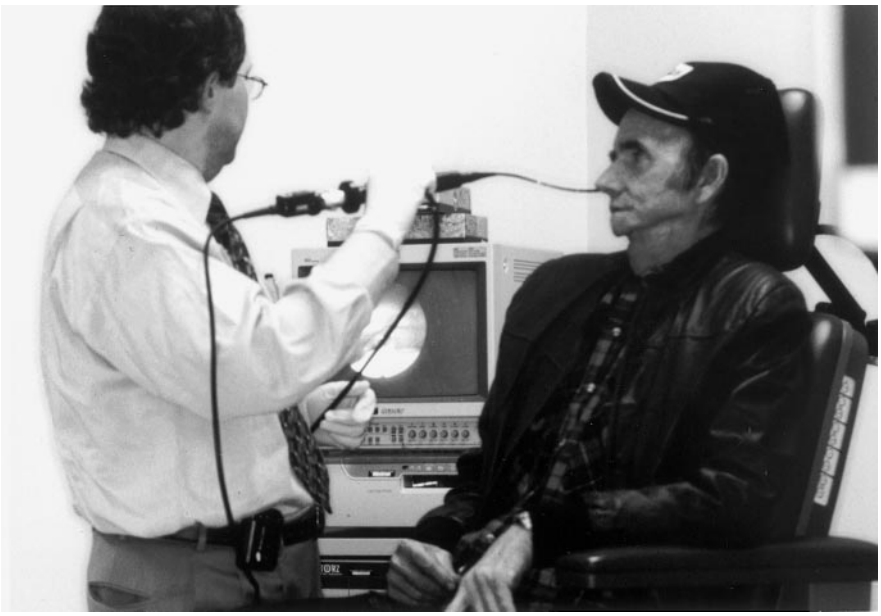


FIGURE 50–4. Flexible fiberoptic examination of the larynx with video documentation.

obstruction from foreign bodies such as displaced teeth or tongue and soft tissue collapse into the pharynx. Because of its efficacy and lack of cervical movement, the jaw thrust is the most appropriate positioning maneuver.[2] Lifting the chin and extending the neck may be successful in improving the airway but risks cervical spinal cord injury in a trauma patient. An oropharyngeal airway may relieve soft tissue obstruction of the airway (Figure 50–5). It serves to displace the tongue anteriorly, providing an unobstructed airway. A nasopharyngeal airway (trumpet) may be placed transnasally into the posterior hypopharynx, thus relieving soft tissue obstruction of the posterior oropharynx. Nasopharyngeal airways tend to be less stimulating than oral airways in awake patients but risk epistaxis. Nasopharyngeal trumpets should not be used in patients with known or suspected basilar skull fractures owing to the risk of inadvertent intracranial placement.[3]

A laryngeal mask airway is a nondefinitive technique sometimes employed by emergency med-

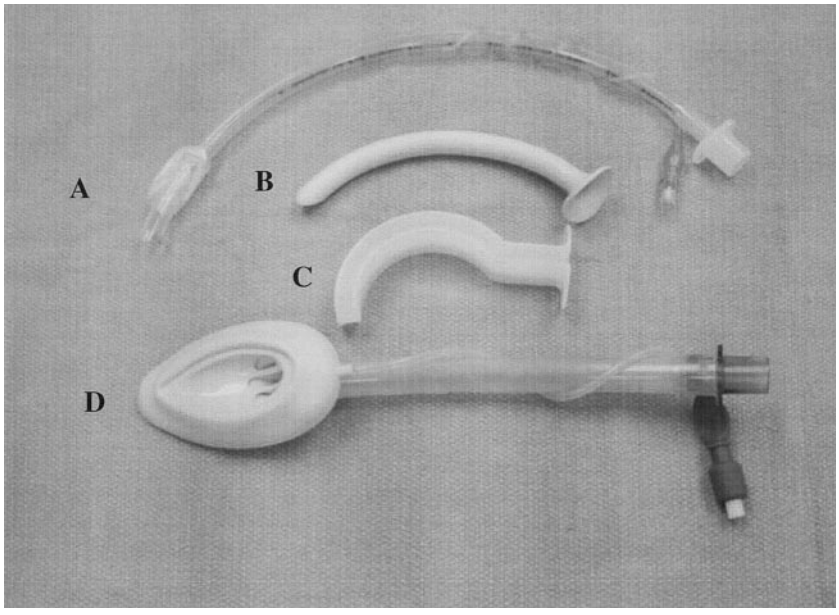


FIGURE 50–5. *A*, Endotracheal tube; *B*, nasopharyngeal trumpet; *C*, oral airway; *D*, laryngeal mask airway. Courtesy of Dr. Michael S. Higgins.

ical technicians or anesthesiologists to obtain a temporary airway for positive pressure ventilation. This device forms a "seal" over the larynx and provides a route for oxygen flow into the trachea. If a definitive airway is required, endotracheal intubation is the quickest and most successful option whenever possible. It allows ventilatory control with the ability to deliver high levels of oxygen and temporarily protects against aspiration. Endotracheal intubation may be achieved through either an oral or a transnasal route. In the case of head and neck trauma, transoral intubation may be achieved while holding in-line cervical traction and maintaining cervical stabilization. Manual in-line immobilization is the optimal method of stabilization and results in significantly less cervical spine movement than stabilization in a Philadelphia collar.[4]

The most common technique employed for oral intubation in the emergent situation is the rapid sequence induction. This uses preoxygenation and denitrogenation followed by a short-acting hypnotic agent such as thiopental or midazolam followed by a neuromuscular blocking agent. An assistant may hold cricoid pressure to prevent aspiration if the patient is unconscious. If a cervical spine injury is suspected, a second assistant maintains in-line stabilization during the intubation. A fiber-optic endoscope may also be placed inside an endotracheal tube to facilitate the localization of the airway. The tube can be passed over the endoscope after the airway has been identified and entered. One disadvantage of this endoscopic technique is that the presence of blood and secretions may obscure visualization through the endoscope.

A more invasive method of providing oxygenation to a patient who cannot be intubated or ventilated is the technique of percutaneous transtracheal ventilation. This procedure involves a caudally directed puncture of the cricothyroid membrane with a 14-gauge intravenous catheter. One must ensure that the catheter is within the trachea and not lodged in the subcutaneous tissue, in which case, subcutaneous emphysema may result and make further attempts at tracheal intubation more difficult. After the catheter has been placed into the trachea, the needle is removed and oxygen is provided via a jet injector, ventilating bag, or gas outlet of an anesthesia machine. This technique is merely a temporizing measure until a definitive surgical airway can be obtained.

Cricothyroidotomy

If the airway cannot be controlled through endotracheal intubation or other medical maneuvers, a surgical airway must be obtained. Cricothyroidotomy is a surgical method of airway management that was reintroduced to the medical community by Brantigan and Grow in 1976.[5] The technique simply involves the creation of an opening into the cricothyroid membrane followed by the placement of a stenting tube. Cricothyroidotomy is often considered the preferred method of obtaining an emergent surgical airway.[6–8] The procedure begins with the identification of the cricothyroid space between the thyroid and cricoid prominences. Local anesthesia may be infiltrated into the area if time permits. A horizontal incision is made over the middle third of the membrane and is carried directly into the airway. A dilating instrument is inserted next to the scalpel to secure the airway opening. Mayo scissors are used to enlarge the opening, and every effort is made not to damage the cricoid or thyroid cartilages. An endotracheal tube may be placed through the cricothyroidotomy and ventilation begun through the lumen of the tube. Once the airway has been secured, the patient may be taken to the operating room and the cricothyroidotomy may be repaired and converted to a tracheostomy.

Although cricothyroidotomy has been used for long-term airway management, it is most useful in the emergency setting, where a surgical airway must be rapidly gained.[9] Cricothyroidotomy is faster and usually easier to perform than a tracheostomy, especially in the hands of a nonsurgeon, requiring little surgical skill other than a knowledge of anatomy. Reported complications of cricothyroidotomy include bleeding, tube displacement, infection, true vocal cord damage, subcutaneous emphysema, and the development of subglottic or tracheal stenosis.[10,11] Hawkins et al confirmed the safety of emergency cricothyroidotomy in their series of 5,603 consecutive adult trauma patients, of whom 66 required cricothyroidotomy.[7] No significant morbidity or complications were noted in this series.

Tracheostomy

Tracheostomies have been performed for over 2,000 years. Until the early 1900s, however, the procedure was reserved for essentially moribund patients owing to high rates of morbidity and mortality. Attitudes

toward tracheostomy changed in 1909, when Chevalier Jackson described the modern tracheostomy.[12] Jackson then addressed the high rate of laryngeal and tracheal damage and stenosis that had previously been associated with tracheostomies in his landmark article of 1921, "High Tracheotomy and Other Errors: The Chief Cause of Chronic Laryngeal Stenosis."[13] In this treatise, the author hypothesized that the unacceptably high rates of laryngeal and tracheal stenosis associated with tracheostomies preceding his description were attributable to damage to the thyroid and cricoid cartilages incurred during the performance of "high" tracheostomies. Jackson implored that tracheostomies be performed below the second tracheal ring, thereby avoiding these dreaded complications. This dictum has been followed to the present day.

Technique A tracheostomy is generally begun with a horizontal skin incision approximately midway between the sternal notch and cricoid cartilage, although some surgeons prefer a vertical skin incision. The incision is carried down through the skin, subcutaneous tissue, and platysma to the level of the strap muscles. After identifying the strap muscles, the sternohyoid and sternothyroid muscles are separated in the midline by incising the fascia that connects them. Retractors are used to pull the muscles laterally, revealing the thyroid isthmus. The cricoid cartilage is palpated and identified at this point. The isthmus of the thyroid gland is dissected off of the trachea in a bloodless fascial plane just superficial to the trachea. The thyroid isthmus is transected and suture-ligated. A cricoid hook is used to pull the trachea superiorly to stabilize the airway and improve exposure of the anterior tracheal wall (Figure 50–6). The fine fascia overlying the trachea is carefully separated.

Although horizontal and vertical incisions into the tracheal wall have been proposed, an inferiorly based Bjork flap consisting of several tracheal rings is a superior option for adults. Following Jackson's principle, a horizontal incision is made between the second and third tracheal rings and carried inferiorly with Mayo scissors to fashion an inferiorly based flap. The space between the rings may be calcified in older patients. The Bjork flap is sewn to the subcutaneous tissue of the inferior skin margin with

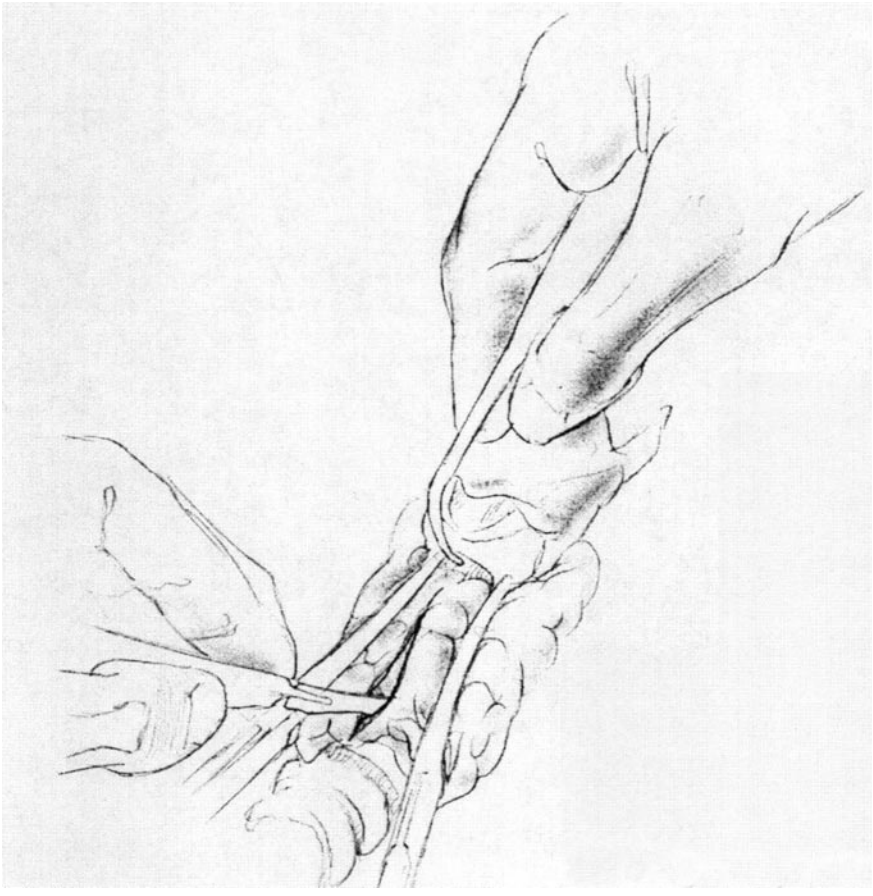


FIGURE 50–6. Stabilization of the trachea with a cricoid hook and division of the thyroid isthmus. Reproduced with permission from Bailey BJ, et al, editors. Head and neck surgery-otolaryngology. Philadelphia: JB Lippincott; 1993. Copyright Lippincott Williams & Wilkins. p. 716.

an absorbable suture (Figure 50–7). This tracheal flap makes reintubation safer in the event of accidental extubation but does lead to a slightly increased risk of tracheocutaneous fistula. For this reason, a horizontal H incision based on the third tracheal ring or the removal of an anterior section of a single tracheal ring may be preferable in patients who are expected to require a tracheostomy only for a short period of time.

When a tracheostomy is permanent or of long duration, the surrounding skin may be surgically defatted and sutured to the tracheostoma circumferentially. This is particularly useful in obese patients, in whom a semipermanent tracheostoma can be fashioned that is less prone to maceration or the formation of granulation tissue and stenosis because the skin directly abuts the respiratory mucosa of the trachea. The creation of a semipermanent tracheostoma in this fashion also serves to decrease the length of the tract of the tracheostomy, thereby facilitating removal and reinsertion of the tracheostomy tube.

The tracheostomy incision should never be closed tightly and must allow the passive egress of air from the wound to prevent subcutaneous emphysema, which could potentially lead to pneumomediastinum, pneumothorax, or infection.

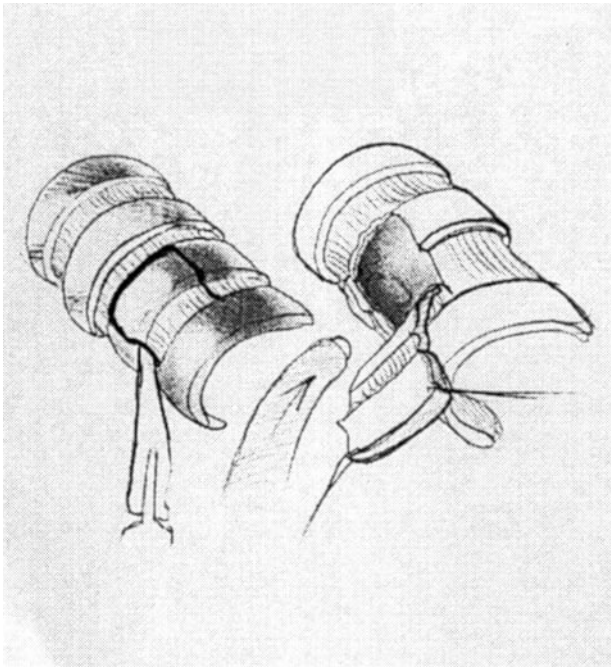


FIGURE 50–7. Creation of an inferiorly based Bjork flap. Reproduced with permission from Bailey BJ, et al, editors. Head and neck surgery-otolaryngology. Philadelphia: JB Lippincott; 1993. Copyright Lippincott Williams & Wilkins. p. 716.

Complications of Tracheostomy The complications of tracheostomy are often divided into two main categories: early and late. Early complications include hemorrhage, tube obstruction, tube displacement, subcutaneous emphysema, and pneumothorax. Late complications include innominate artery rupture, infection, aspiration, granuloma formation, and laryngeal or subglottic stenosis.[14]

Early Complications. Hemorrhage. Mild bleeding is the most common early complication of tracheostomy, occurring in about 5% of cases. The hemorrhage is usually from veins or the incised thyroid isthmus. Mild bleeding can generally be controlled by raising the head of the patient to decrease venous pressure or by placing hemostatic gauze over oozing areas. Major hemorrhage is unusual and most often involves a branch of the superior thyroid artery. Arterial bleeding is an indication for immediate surgical exploration for bleeding control and vessel ligation.

Tube Obstruction. Tube obstruction is the most common cause of ventilatory insufficiency in a posttracheostomy patient and most often results from a mucous plug. Dried mucus and secretions in and around the tracheostomy tube can form crusts that may become large enough to obstruct the lumen of the tube. Audible airflow and difficulty passing a suction catheter through the tube are signs of impending tube obstruction. Adequate humidification, irrigation, frequent suctioning, and cleaning of the inner cannula help to reduce the risk of plugging and obstruction. If the tracheostomy tube becomes obstructed and cannot be recannulized, the patient and care staff should be instructed to remove the entire tube.

Tracheostomy Tube Displacement. Displacement of the tracheostomy tube or dislocation of the tube from the tracheostomy tract poses a considerable risk to the patient who has recently undergone a tracheostomy. Displacement of the tube is most likely to occur in the early postoperative period before a stable tract has matured between the tracheal lumen and the skin. Displacement most often occurs owing to the tube being inadequately secured to the patient with sutures or ties. Excessive tension applied on the

tube from the ventilatory apparatus also poses a risk of tube dislodgment.

Tracheostomy tube replacement in the early postoperative period and during the first tube change should be done under direct vision with the surgeon in attendance. Blind reinsertion risks creating a false passage into the paratracheal tissues.

Subcutaneous Emphysema. Subcutaneous emphysema results from the forced entrance of air into the fascial planes of the neck and usually results from closing the tracheostomy incision too tightly. Tube obstruction or displacement can also lead to the escape of air into the fascial planes. Crepitus and soft tissue swelling can involve the neck, face, and upper torso. Minor amounts of subcutaneous emphysema usually resolve spontaneously without significant sequelae, whereas massive amounts may progress to pneumomediastinum and pose an increased risk of infection through bacterial contamination.

Pneumothorax. Pneumothorax occurs in less than 5% of tracheostomies and can result from damaging the pleural apices during dissection that veers away from the midline. Pneumothorax may also be caused by excessively forceful ventilation that prevents the passive exhalation of air. This air trapping can lead to alveolar rupture and the entrance of air into the pleural space. A misplaced tracheostomy tube can also introduce air into the pleural space or mediastinum. Decreased breath sounds may indicate a pneumothorax, although this complication is usually asymptomatic and noted incidentally on the postoperative chest radiograph. A small pneumothorax is generally observed with serial chest radiographs, whereas a large or symptomatic pneumothorax may require a chest tube for lung re-expansion.

Late Complications. Innominate Artery Rupture. Innominate artery rupture is perhaps the most dreaded tracheostomy complication and, although rare, results in only a 10 to 25% survival rate.[15] Innominate artery rupture is heralded by a "sentinel bleed," which is a brief episode of bleeding occurring minutes to hours prior to a massive hemorrhage. Rapid peristomal bleeding, even when it resolves, should trigger a fiber-optic examination of the trachea to rule out innominate artery involvement. When massive bleeding does occur, the tracheostomy tube should be rapidly replaced with a longer cuffed endotracheal tube to allow oxygenation while digital pressure is applied to compress the innominate artery against the manubrium. The patient should be taken immediately to the operating room for median sternotomy and vessel ligation. Many surgeons feel that the Bjork flap technique of tracheostomy lessens the risk of this potentially fatal complication.

Infection. Although bacterial colonization of the trachea and peristomal area is unavoidable owing to the moist, open nature of the tracheostomy wound, significant infections rarely occur after a tracheostomy. Local cellulitis can generally be managed with aggressive local cleaning measures and topical antibiotics. Uncontrolled local infection can rarely lead to serious complications such as mediastinitis or predispose a patient for innominate artery rupture or peristomal tissue loss. Local infection that does not respond to aggressive local wound care should be treated with systemic antibiotics.

Aspiration. Aspiration following tracheostomy is a phenomenon that is often overlooked. Although a tracheostomy is generally thought to "protect" the airway, tracheostomies have actually been shown to decrease the protective glottic closure reflexes of both the true and false vocal cords.[16] Loss of this protective mechanism does predispose the tracheostomized patient to aspiration and concomitant pneumonia. Prophylactic measures include a soft, solid diet, which poses a decreased risk of aspiration when compared to liquids, and maintaining the patient in a somewhat upright position. Although inflating the tracheostomy tube cuff can afford temporary airway protection, hyperinflation of the cuff actually exacerbates aspiration owing to compression of the esophagus. Consultation with a speech and swallowing therapist is advisable for any patient who requires a tracheostomy.

Granuloma Formation. Granulation tissue may form in the peristomal area owing to chronic mucosal and skin irritation by an inadequately fixed or improperly fitted tracheostomy tube. Granulation tissue may also form inside the tracheal lumen. Early and small granulomas may be managed with local débridement. Mature, fibrotic lesions generally require endoscopic laser excision.

Subglottic Stenosis. Tracheostomies may contribute to subglottic stenosis by causing chronic irritation to the tracheal mucosa while also providing a route of bacterial contamination that causes

increased inflammation in the damaged area. This inflammation results in the formation of scar tissue and subsequent subglottic stenosis. Gastroesophageal reflux is also felt to increase the risk of subglottic stenosis by providing additional mucosal irritation that contributes to granulation and scar tissue formation. The risk of this complication underscores the importance of minimizing extraneous motion of the tracheostomy tube, managing gastroesophageal reflux, and removing a tracheostomy tube when it is no longer required for airway support and protection.

Percutaneous Tracheostomy

Although percutaneous tracheotomy was first described in 1955,[17] it experienced a resurgence in popularity after its reintroduction by Ciaglia and colleagues in 1985.[18] The technique involves the transcervical insertion of a guidewire into the airway followed by blunt dilation of the guidewire tract. The tracheostomy tube is inserted over the guidewire using a modified Seldinger technique. The advantage of the procedure is that it can be performed at the bedside in the intensive care unit.[19] Visualization of the guidewire insertion, dilation, and tracheostomy tube placement may be achieved via a flexible fiber-optic laryngoscope inserted at the time of the procedure. This improved visualization decreases airway complications by preventing improper placement of the guidewire or esophageal perforation by the inserting device or guidewire itself.[20] The proper level of tracheostomy placement may also be confirmed with airway visualization at the time of the procedure.

One major disadvantage of percutaneous tracheostomy is that only a narrow, dilated tract, rather than a formal stoma, is created during the procedure. This results in an increased rate of tube displacement and difficult tube reinsertion with percutaneous tracheostomy as compared with a standard tracheostomy.[21] An increased rate of postoperative death owing to tube dislodgment has been described with the percutaneous technique.[22] The surgeon may wish to consider this factor in planning a tracheostomy procedure on a patient who is known to be difficult to intubate translaryngeally. Likewise, a patient with morbid obesity or cervical anatomic anomalies is not a good candidate for percutaneous tracheostomy and should likely undergo a standard tracheostomy performed in the controlled environment of the operating room.

An additional disadvantage of percutaneous tracheostomy is the risk of tracheal or subglottic stenosis. Unlike soft tissue, cartilage is rigid and incapable of changing shape and conforming to the shape of a dilator. Because of the progressive dilations performed during percutaneous tracheostomies, the tracheal cartilages are at high risk for crush injury during the procedure. Damaged cartilage poses a risk of scar tissue formation and subsequent tracheal or subglottic stenosis.

Since a percutaneous tracheostomy must sometimes be converted to a standard open tracheostomy, a surgical tracheostomy tray and someone with the ability to perform a standard tracheostomy should be present at the time of the percutaneous procedure. With these precautions and adequate training, percutaneous tracheostomy can be a safe and efficient procedure.

Open Bedside Tracheostomy

A tracheostomy can also be performed in the intubated patient as an open procedure at the bedside in the intensive care unit setting.[23] The only advantage of performing a tracheostomy at the bedside is that operating room costs are defrayed and the procedure is generally less expensive than a percutaneous tracheostomy. Obvious disadvantages are that lighting, surgical assistance, and equipment are inferior to that available in a standard operating room. Patients with unfavorable anatomy such as cervical mass lesions or morbid obesity are not good candidates for an open bedside tracheostomy and should proceed to the operating room for a standard tracheostomy.

LARYNGEAL AND TRACHEAL STENOSIS IN THE ADULT

The management of laryngeal and tracheal stenosis in adults is both challenging and intriguing. The diagnosis and evaluation of these lesions requires a complete mastery of the anatomy and physiology of the upper aerodigestive tract. The incredible variability of stenotic areas in the adult larynx and trachea requires the head and neck surgeon carefully to individualize treatment and provide unique strategies for the management of these diverse lesions.

Etiology

There are many factors that can lead to laryngotracheal stenosis (LTS) (Table 50–1). Most cases of adult LTS result from external trauma or prolonged endotracheal intubation. External trauma causes cartilage damage and mucosal disruption with hematoma formation. These hematomas eventually organize and result in collagen deposition and scar tissue formation. Endotracheal intubation can cause mucosal damage through pressure necrosis and direct injury resulting from the pressure of the endotracheal tube or cuff. Mucosal ulceration also leads to healing through collagen deposition, fibrosis, and scar tissue formation. Lesions from endotracheal intubation are usually located in the posterior part of the glottis, where the tube most often contacts mucosa, or in the trachea, where the cuff or tube tip causes mucosal damage. Low-pressure endotracheal tube cuffs have somewhat reduced the rate of cuff-induced damage. The length of intubation, tube movement, tube size, and gastroesophageal reflux can also contribute to LTS.

Classification

Since areas of LTS are so variable in their size, consistency, and location, a rigid classification scheme is essentially impossible, and the surgeon must describe and document the lesion in a way that is widely understandable and reproducible. The classification of LTS in adults begins with the anatomic location of the lesion as glottis, subglottis, trachea, or a combination of these. These stenotic segments may then be further described as anterior, posterior, or circumferential. The size and length of the stenotic area are critical in classifying the lesion as well.

Diagnostic Assessment

The evaluation of LTS must begin with a meticulous history and physical examination. Since most cases of LTS result from laryngotracheal trauma or endotracheal intubation, the timing of the predisposing incident should be recorded. Any previous airway evaluations or attempts at repair should also be noted. The patient should be questioned regarding the onset, duration, and severity of symptoms such as exercise intolerance, disruption of lifestyle, and tracheostomy dependence. Patients who do require a tracheostomy should be questioned as to how often

TABLE 50–1. Causes of Laryngeal and Tracheal Stenosis in Adults

- Trauma
  - External laryngotracheal injury
    - Blunt neck trauma
    - Penetrating wound of the larynx
  - Internal laryngotracheal injury
    - Prolonged endotracheal intubation
    - Post-tracheostomy
    - Postsurgical procedure
    - Postirradiation therapy
    - Endotracheal burn
      - Thermal
      - Chemical
- Chronic inflammatory disease
  - Bacterial: diphtheria
  - Syphilitic
  - Fungal: histoplasmosis
  - Tuberculosis
  - Leprosy
  - Sarcoidosis
  - Rhinoscleromosis
- Benign neoplasms
  - Intrinsic
    - Papillomas
    - Chondromas
    - Minor salivary gland
    - Neural
  - Extrinsic
    - Thyroid
    - Thymus
- Malignant neoplasms
  - Intrinsic
    - Squamous cell carcinoma
    - Minor salivary gland
    - Sarcomas
    - Lymphoma
  - Extrinsic
    - Thyroid carcinoma
- Collagen vascular disease
  - Wegener's granulomatosis
  - Relapsing polychondritis
  - Other

Reproduced with permission from Cummings C, et al, editors. Otolaryngology-head and neck surgery. 2nd ed. St. Louis: Mosby Year Book; 1993. p. 2002.

the tube may be plugged. Symptoms of aspiration, voice change, or dysphagia may indicate the degree of glottic involvement.

The entire upper aerodigestive tract must be carefully examined in a patient with suspected LTS. Indirect laryngoscopy and flexible fiber-optic laryngoscopy offer critical information regarding the supraglottic airway and mobility of the true vocal folds. In extreme abduction, areas of subglottic stenosis may be visible using these techniques. Video documentation of these procedures offers a valuable method of treatment planning and patient education.[24]

Although imaging studies such as airway radiographs, computed tomography, and magnetic resonance imaging occasionally provide useful information, the most valuable diagnostic assessment stems from the examination of the patient by endoscopy. After the patient has been examined by indirect laryngoscopy and flexible fiber-optic techniques, rigid endoscopic evaluation under general anesthesia should be performed in all patients with symptomatic airway abnormalities. Direct measurement and documentation of the diameter and length of stenotic areas are critical steps in the management of these lesions. Measurement of the diameter of stenotic segments is best evaluated by passing an endoscope, with a known diameter, that just fits through the stenotic area. Measurement of stenosis length may be performed by placing the endoscope at the distal end of the stenotic segment and marking the instrument at the incisors. The endoscope is withdrawn to the proximal aspect of the stenosis and remarked. The length of stenosis may be measured on the endoscope.

Treatment of Laryngotracheal Stenosis

Medical Most surgeons who treat LTS recommend the use of antibiotics, local and/or systemic corticosteroids, and histamine$_2$ antagonists or omeprazole when treating mild or acute subglottic injuries. In mature areas of laryngotracheal fibrosis and scarring, these medical treatments are unlikely to result in any clinical improvement.

Endoscopic Some areas of LTS are amenable to endoscopic treatment techniques such as laser vaporization and dilation, excision using a microtrapdoor technique, or serial dilation with radial incisions of the stenotic segment. Intralesional corticosteroids may also be injected under endoscopic guidance.[25] Lasers allow the precise treatment of tissue throughout the airway while avoiding external incisions and providing an excellent method of cutting, coagulating, or vaporizing tissue. Hemostasis may be achieved, and perioperative edema is often decreased with the use of lasers owing to smaller amounts of tissue sustaining thermal damage when compared to electrocautery.

The ideal laser for any particular case depends on the type of tissue and lesion that requires treatment. Owing to its precision (the small spot size) and availability, the carbon dioxide laser, which produces light in the mid-infrared region, remains the instrument of choice in the endoscopic management of LTS. Although the carbon dioxide laser is very precise, it is not a good instrument for coagulation and can only be used to coagulate vessels up to 0.5 mm in diameter. If the area of stenosis is thought to be very vascular, a laser with better hemoglobin absorption, such as a KTP or Nd:YAG laser, would be recommended. One further disadvantage of the carbon dioxide laser is that it lacks a good fiber-optic delivery system and must generally be controlled via a micromanipulator system mounted on a microscope.

Laser ablation of stenosis is a useful technique that may be combined with dilation of the stenotic segment or placement of an intraluminal stent. This procedure is most successful in the management of early lesions composed mostly of granulation tissue that have not yet evolved into a mature scar. Stenotic segments less than 1 cm in length may also be addressed with this method.

The microtrapdoor technique is used to debulk stenotic areas confined to a single quadrant of the airway. The procedure is somewhat more technically demanding than simple ablation of the stenosis. First, a laser incision is made on the superior aspect of the stenotic shelf (Figure 50–8). Once the incision has been made, a microelevator is used to elevate a mucosal flap over the area of scar tissue. The laser may be used to vaporize the scar tissue forming the stenosis. Once the scar tissue has been adequately vaporized, the flap is repositioned into its original location. The laser may be defocused and milliwatt power settings used to weld the edges of the incision together.

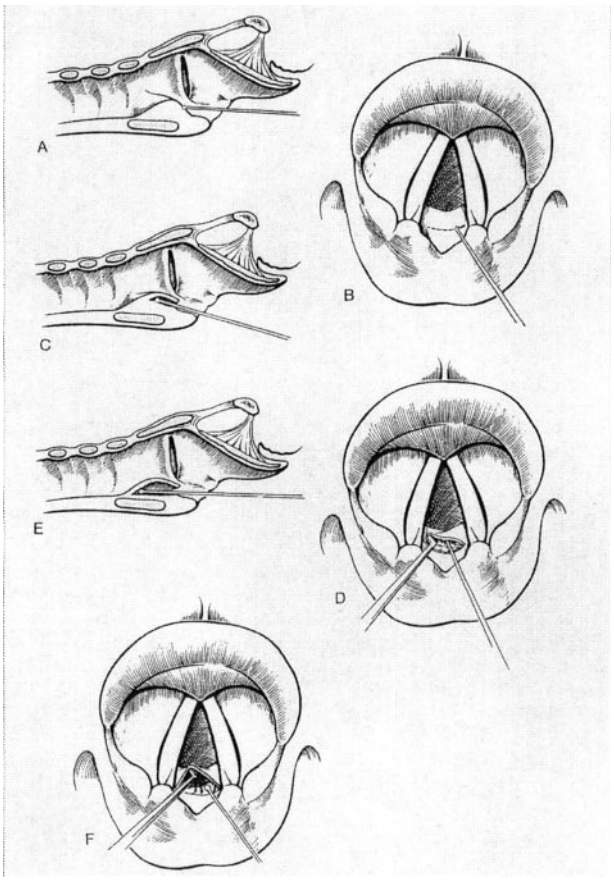


FIGURE 50–8. Microtrapdoor technique used to debulk areas of stenosis. *A* and *B*, The laser is used to make an incision on the superior aspect of the stenotic shelf. *C* and *D*, The flap is elevated distally. *E* and *F*, The scar tissue is vaporized in a submucosal plane. Reproduced with permission from Coleman JA et al.[25]

Areas of circumferential stenosis may be addressed endoscopically by making radial incisions in the scar tissue with a laser and dilating the treated area with a bronchoscope or dilating instrument. The laser is used to create radial incisions in the scar tissue in four to six areas (Figure 50–9). The laser may be coupled to a ventilating bronchoscope if the patient is not tracheostomy dependent. The incisions break up the circumferential scar band and leave islands of intact mucosa. Leaving areas of intact mucosa is critical to prevent circumferential areas denuded of mucosa that would ultimately reform scar tissue similar to the original lesion or possibly even make the scarring worse. The stenotic segment may be sequentially dilated. The procedure generally needs to be repeated at several 3- to 4-week intervals before an adequate airway is achieved. Local applications of corticosteroids or mitomycin-C may be used with any of the endoscopic techniques in an attempt to prevent the reformation of the cicatrix.

Open Surgical Techniques Severe areas of LTS that do not respond to endoscopic techniques require an open surgical procedure. Open techniques attempt either to excise the stenotic segment and reanastomose the airway or to augment the circumference of the stenotic segment with transplanted tissue. Stenoses that are longer than 1 cm, have glottic or extensive tracheal involvement, have failed endoscopic treatment, and have near-complete stenoses are candidates for an open technique. In patients with diabetes or severe systemic illnesses, open approaches must be considered with great care. Such patients suffer from poor wound healing, have a high risk of perioperative complications, and have a lower rate of successful outcome after open procedures. In these high-risk patients, a tracheostomy may be the most prudent choice of management.

Laryngotracheoplasty is often used to repair segments of LTS more than 1 cm in length.[26] The procedure begins with a transverse cervical incision in a natural skin crease at the level of the cricoid cartilage. Subplatysmal flaps are elevated superiorly to the thyroid notch and inferiorly to the sternum. The strap muscles are separated in the midline and retracted laterally. The thyroid isthmus is divided and ligated. The anterior cricoid ring is divided vertically with a knife or saw. The incision is extended superiorly and inferiorly based on the exact location and extent of the stenosis (Figure 50–10). The goal is to open the entire stenotic segment. To prevent voice damage, the incision is not extended to the anterior commissure unless this area is involved by the stenosis. If the cricoid cartilage cannot be opened sufficiently owing to calcification or extensive scarring, a posterior cricoid cartilage split may be performed to gain the needed mobility. If a posterior cricoid cartilage split is performed, the surgeon must be careful not to damage the underlying esophagus.

Although it is tempting to remove scar tissue at this point in the procedure, the surgeon must remember that the goal of the procedure is to open the area of stenosis widely rather than to remove tissue. Resection of scar tissue and mucosa risks scar regeneration and recurrent stenosis. Submucosal

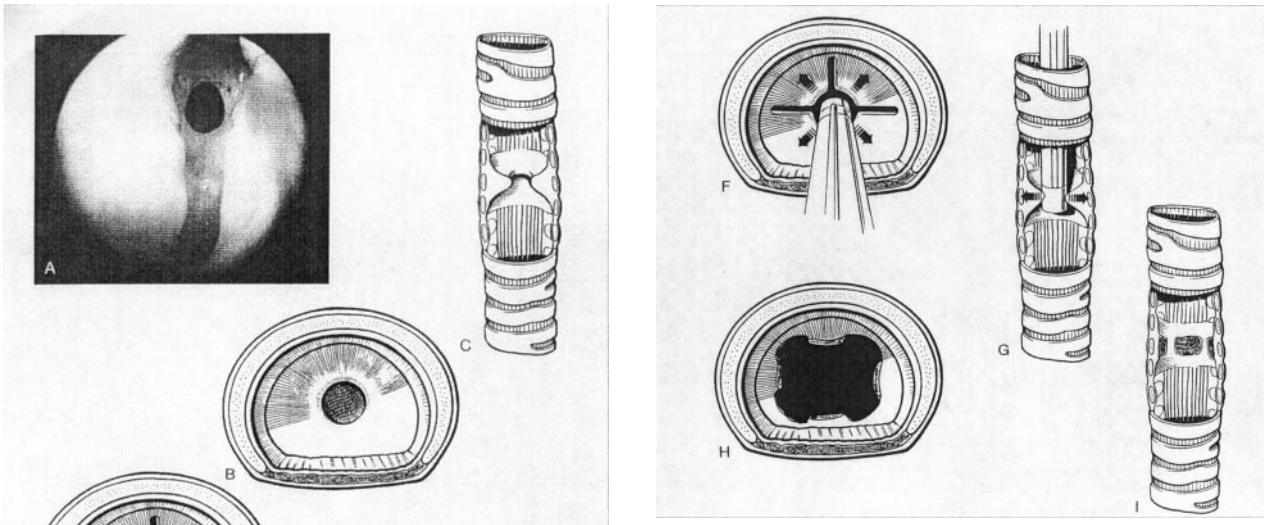


FIGURE 50–9. Laser-assisted radial incisions with dilation of tracheal stenosis. *A*, Clinical photograph of a proximal tracheal stenosis. *B* and *C*, Circumferential tracheal stenosis that would be amenable to radial incision and dilation. *D* and *E*, Radial incisions are made with the laser. *F* and *G*, The stenosis is dilated. *H* and *I*, Careful dilation results in an open airway with areas of intact mucosa. Reproduced with permission from Coleman JA et al.[25]

resection of scar tissue, however, is generally safe and does not increase the risk of restenosis.

A posterior graft, if it must be used, is placed next. Posterior grafts usually contain a flange that is tucked behind the posterior lamina of the cricoid cartilage bilaterally to prevent extrusion into the airway. An airway stent, if one is to be placed, is inserted after placement of the posterior cricoid graft. The anterior cartilage graft is generally cut in a fusiform or boat-shaped fashion with beveled edges to prevent the graft from slipping into the airway. The graft is sutured into place while taking great care to ensure that the sutures remain extraluminal. The strap muscles are reapproximated in the midline, and the platysma and skin are closed over a passive drain to allow air egress and prevent subcutaneous emphysema. The success rate of primary laryngotracheal reconstruction in adults is in the 60 to 70% range.[27]

Cartilage grafts for laryngotracheoplasty may be harvested from the costal cartilage or nasal septal cartilage. Costal cartilage is the most commonly used source of grafts. The fifth or sixth costal cartilage is harvested with the anterior perichondrium intact. The grafts are carved to the appropriate shape, with the perichondrium facing the airway lumen. The hyoid bone may also be harvested through the existing cervical incision and used as an interposition graft, which offers the advantage of avoiding the need to prepare a second operative area. The central portion of the hyoid bone between the lesser cornua may be used as a free graft and fixed into place. The hyoid bone has also been used as a pedicled graft based inferiorly on the sternohyoid muscle.[28]

Tracheal Resection and Reanastomosis Areas of cervical tracheal stenosis up to about 5 cm can generally be excised, and the proximal and distal tracheal segments reanastomosed primarily. A suprahyoid laryngeal release may be required to allow for closure under minimal tension. When performing this procedure, the surgeon must keep in mind the age and body habitus of the patient. Older patients

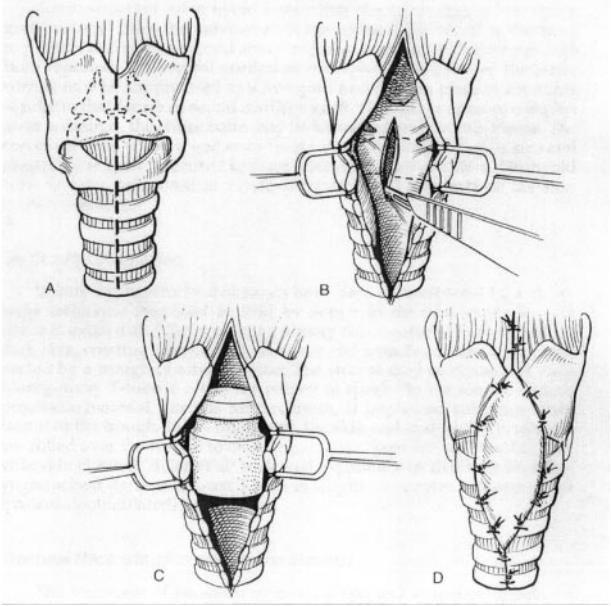


FIGURE 50–10. *A*, Vertical splitting of the cricoid cartilage during laryngotracheoplasty. *B*, Splitting of the posterior cricoid. *C*, Laryngeal stent in place. *D*, Graft sutured in place. Reproduced with permission from Duncavage JA and Koriwchak MJ.[24]

tend to have calcifications between the tracheal rings, resulting in decreased tracheal elasticity. Patients with large, thick necks and older patients with cervical kyphosis also tend to lack tracheal mobility.

Prior to the procedure, the patient is generally positioned with the neck extended on a shoulder roll. If a tracheostomy is present, the skin of the tracheostomy site is included in the incision planning and is excised. After a transverse skin incision has been made, subplatysmal flaps are raised superiorly to the thyroid cartilage and inferiorly to the manubrium. The strap muscles are separated in the midline and retracted laterally. The thyroid isthmus is divided and ligated, if necessary. The area of stenosis is then located and can be confirmed endoscopically. Care is taken to preserve the lateral blood supply to the trachea proximal and distal to the site of stenosis. The dissection around the area of stenosis must proceed directly on the trachea to prevent injury to the recurrent laryngeal nerves.

The esophagus is identified and carefully dissected away from the membranous trachea. The trachea may be incised above and below the stenotic segment. The endotracheal tube is placed transorally into the distal tracheal segment. The closure always begins posteriorly. Vicryl sutures are placed in an extraluminal fashion through the perichondrium of the proximal and distal tracheal segments (Figure 50–11). The neck is flexed prior to tying the sutures. Additional sutures may be placed around a tracheal ring above and below the site of anastomosis to decrease tension on the closure. The strap muscles are reapproximated in the midline. The platysma and skin are closed after placement of a passive Penrose drain that allows the egress of air and prevents subcutaneous emphysema.

It is essential to minimize the tension on the closure after tracheal reanastomosis. Surgical techniques that decrease tension include mobilization of the trachea through blunt anterior and posterior dissection and release of the suprahyoid muscular attachments. A cardiothoracic surgeon may also perform a division and reattachment of a bronchus, which allows further superior mobility of the distal trachea. Positioning maneuvers such as neck flexion also decrease tension on the site of anastomosis. Some surgeons place a stitch between the chin and manubrium (Grillo stitch) to prevent the patient from extending the neck during the postoperative period.

The patient may be extubated at the conclusion of the procedure or remain intubated overnight and extubated on the first postoperative day at the discretion of the surgeon. The success rate for this procedure generally exceeds 90%.[29]

Several new techniques have been described and applied in the treatment of segmental tracheal stenosis. Cryopreserved, irradiated tracheal homograft transplantation has been used to treat tracheal stenosis. This technique requires the harvest of tracheal tissue from donor cadavers within 24 hours of death. The grafts are placed in a –70°C chamber prior to irradiation to 25 kGy and are stored at –70°C until transplantation. Kunachak et al reported that three of four patients treated with this technique were successfully decannulated and display a near-normal tracheal lumen with histologic evidence of normal respiratory epithelium at the grafted site.[30]

The technique of microvascular free tissue transfer has also been applied to the challenging problem of LTS. Esclamado and Carroll reported the use of a fibular osseocutaneous free flap in the

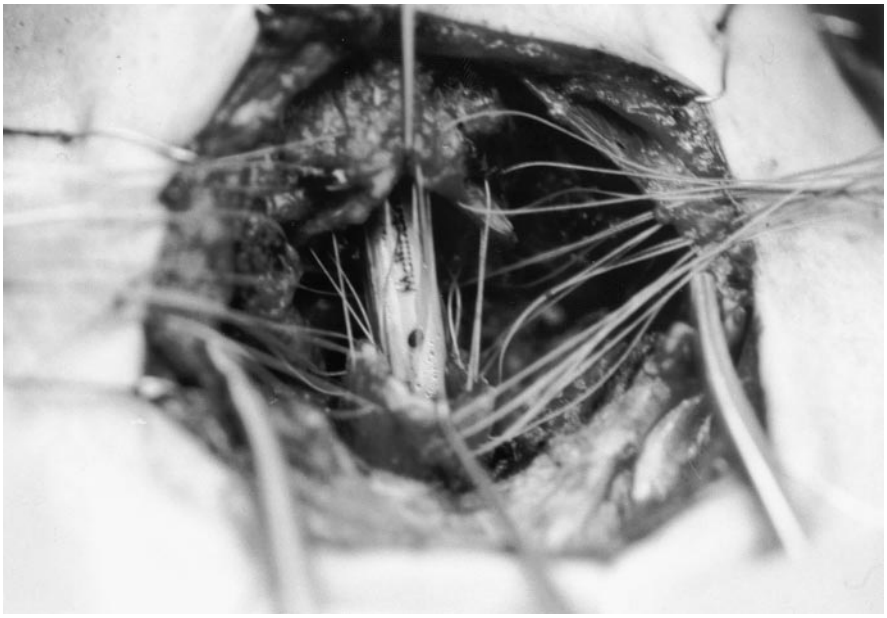


FIGURE 50–11. Tracheal anastomosis intraoperative photograph after the excision of a stenotic tracheal segment. Sutures have been placed but not tied. Courtesy of Dr. James L. Netterville.

reconstruction of a complex LTS.[31] At this point, the usefulness of microvascular free flaps appears to be limited to a subset of complex cases not amenable to more conventional surgical techniques.

CRICOARYTENOID JOINT FIXATION

The cricoarytenoid joint is a synovial joint formed by the articulation of the arytenoid cartilage with the posterosuperior aspect of the cricoid cartilage. The vocal process of the arytenoid cartilage is usually free to rotate in three dimensions to allow proper apposition of the true vocal folds. This normal mobility of the arytenoids can be impaired by several factors. Dislocation of the arytenoid cartilage may occur due to either external trauma or intubation. The arytenoid may be dislocated anteriorly or posteriorly, with anterior dislocation being slightly more common owing to the force vector exerted by the blade of a laryngoscope. Several inflammatory disorders such as rheumatoid arthritis and gout may also involve the cricoarytenoid joint and result in an abnormal "fixation" of the joint. Inflammatory disorders can result in unilateral or bilateral cricoarytenoid joint dysfunction. With cricoarytenoid arthritis, the patient generally presents with symptoms of stridor and dyspnea with a variable degree of dysphonia.[32] The dysfunction results from fixing of the vocal cords in the paramedian position and the inability to achieve normal apposition of the true vocal folds on phonation or normal abduction with inspiration. However, denervation of the larynx may also limit the normal mobility of the arytenoids.

In the patient with suspected cricoarytenoid joint dysfunction, the differential diagnosis is threefold. First, the posterior cricoarytenoid muscle may be denervated. Second, the arytenoid cartilage may be dislocated. A history of trauma should be carefully elicited in this group of patients. Finally, the joint may be fixed because of an inflammatory condition. The proper location and mobility of the cricoarytenoid joint are best assessed under general anesthesia by gently rocking the arytenoid cartilage back and forth. A "fixed" cricoarytenoid joint can be diagnosed with this method. An electromyogram of the intrinsic laryngeal musculature can determine if the immobility is caused by denervation or is secondary to fixation of the cricoarytenoid joint. A careful history aids in the diagnosis of a systemic inflammatory condition.

The treatment of cricoarytenoid joint dysfunction must be carefully individualized to the patient and the disease process. A patient with an inflammatory disease should be treated medically with the aid of a rheumatologist. Stable patients who are able to phonate, breathe, and swallow well without aspiration may be safely observed. Unstable patients or patients who do not improve on medical and rehabilitative strategies must be treated surgically. Patients with inadequate ventilation may undergo a tra-

cheostomy, which would likely result in maintenance of an excellent voice owing to the medial position of the vocal cords. Patients who wish to be decannulated can be offered a cordectomy or arytenoidectomy, but they must be carefully counseled that a more patent airway will result in an inferior voice. A successful treatment strategy requires excellent communication between the surgeon and the patient and intensive patient education and counseling.

Cordectomy and Arytenoidectomy

A carbon dioxide laser endoscopic posterior partial transverse cordotomy can sometimes provide an adequate airway and allow decannulation in a patient whose true vocal folds are in a medial position, cannot be abducted, and prevent adequate airflow through the glottis.[33] The cause of the medialized vocal folds may be either bilateral cricoarytenoid joint fixation or bilateral vocal fold paralysis. The procedure of laser cordotomy involves excising a transverse wedge of tissue from the posterior portion of one vocal fold just anterior to the vocal process of the arytenoid cartilage. The technique may also be performed bilaterally. The excision of this tissue in the posterior glottis allows a portal for airflow through an otherwise constricted airway. Complications of the procedure include granulation or scar tissue formation at the site of the cordotomy, requiring a repeat procedure, and decreased voice quality. The patient must be counseled that improving the airway necessarily decreases the quality of voice and that a balance between the two must be struck.

Excision or lateralization of the arytenoid cartilage is a second option for improving the patency of the airway when the vocal cords are adducted and airflow through the glottis is inadequate. The arytenoid cartilage may be removed surgically and the true vocal fold sutured laterally to improve the airway. Likewise, an endoscopic laser partial arytenoidectomy may be performed.[34] The goal of both techniques, as with laser posterior cordectomy, is to improve airflow through the posterior portion of the glottis while minimally affecting the anterior two-thirds of the glottis, which is largely responsible for phonation. Successful outcomes may result from each of the three procedures, and the technique of choice must be tailored to the experience of the surgeon and the individual patient.

CONCLUSION

The successful management of the airway requires a complete understanding of the structure and function of the upper aerodigestive tract. Likewise, a systematic and progressive algorithm must be employed in cases of airway distress. Several options exist for the management of patients who require a surgical airway, and the correct technique must be individually matched to the needs and situation of the patient.

Laryngotracheal stenosis poses a formidable challenge to the airway surgeon owing to the variability in the location and degree of stenotic segments. Areas of stenosis must be carefully evaluated prior to planning any treatment regimen. Once again, the treatment must be individualized to the patient, taking into account the etiology and anatomy of the lesion as well as the medical status of the patient.

REFERENCES

1. Graney DO, Marsh B. Anatomy of the trachea. In: Cummings CW, Fredrickson JM, Harker LA, editors. Otolaryngology-head and neck surgery. 2nd ed. St. Louis: Mosby Year Book; 1993. p. 2212–6.
2. Kharash M, Graff J. Emergency management of the airway. Crit Care Clin 1995;11:53–66.
3. Rodricks MB, Deutschman CS. Emergent airway management. Crit Care Clin 2000;16:389–409.
4. Walls RM. Airway management in the blunt trauma patient: how important is the cervical spine? Can J Surg 1992;35:27–30.
5. Brantigan CO, Grow JB. Cricothyroidotomy: elective use in respiratory problems requiring tracheotomy. J Thorac Cardiovasc Surg 1976;71:72–80.
6. Grande CM, Stene JK, Berhard WN. Airway management: considerations in the trauma patient. Crit Care Clin 1990;6:37–59.
7. Hawkins ML, Shapiro MB, Cue JI. Emergency cricothyroidotomy: a reassessment. Am Surg 1995;61:52–5.
8. Walls RM. Management of the difficult airway in the trauma patient. Emerg Med Clin North Am 1998; 16:45–61.
9. Burkey BB, Esclamado R, Morganroth M. The role of cricothyroidotomy in airway management. Clin Chest Med 1991;12:561–71.
10. Einarsson O, Rochester CC, Rosenbaum S. Airway management in respiratory emergencies. Clin Chest Med 1984;15:13–34.

11. American College of Surgeons. Advanced trauma life support for doctors 1997. 6th ed. Chicago: First Impression; 1997.
12. Jackson C. Tracheotomy. Laryngoscope 1909;18:285.
13. Jackson C. High tracheotomy and other errors: the chief causes of chronic laryngeal stenosis. Surg Gynecol Obstet 1921;32:382–98.
14. Reilly H, Sasaki CT. Tracheotomy complications. In: Complications in head and neck surgery. Philadelphia: WB Saunders; 1993. p. 257–74.
15. Jones JW, Reynolds M, Hewitt RL, Drapanas T. Tracheoinnominate artery erosion: successful surgical management of a devastating complication. Ann Surg 1977;184:194–204.
16. Sasaki CT, Suzuki M, Horiuchi M, Kirchner JM. The effect of tracheostomy on the laryngeal closure reflex. Laryngoscope 1977;87:1428–33.
17. Sheldon CH, Pudenz RH, Freshwater DB. A new method of tracheotomy. J Neurosurg 1955;12:428.
18. Ciaglia P, Firsching R, Syniec C. Elective percutaneous dilatational tracheostomy. A new simple bedside procedure; preliminary report. Chest 1985; 87:715–9.
19. Massick DD, Powell DM, Price PD, Chang SL. Quantification of the learning curve for percutaneous dilatational tracheotomy. Laryngoscope 2000;110: 222–8.
20. Donaldson DR, Emami AJ, Wax MK. Endoscopically monitored percutaneous dilational tracheotomy in a residency program. Laryngoscope 2000;110: 1142–6.
21. Graham JS, Mullory RH, Sutherland FR, Rose S. Percutaneous versus open tracheostomy: a retrospective cohort outcome study. J Trauma 1996;42:245–50.
22. Dulguerov P, Gysin C, Perneger TV, Chevrolet JC. Percutaneous or surgical tracheostomy: a meta-analysis. Crit Care Med 1999;27:1617–25.
23. Wang SJ, Sercarz JA, Blackwell KE, et al. Open bedside tracheotomy in the intensive care unit. Laryngoscope 1999;109:891–3.
24. Duncavage JA, Koriwchak MJ. Open surgical techniques for laryngotracheal stenosis. Otolaryngol Clin North Am 1995;28:785–95.
25. Coleman JA, Van Duyne JA, Ossoff RH. Laser treatment of lower airway stenosis. Otolaryngol Clin North Am 1995;28:771–83.
26. McCaffrey TV. Management of subglottic stenosis in the adult. Ann Otol Rhinol Laryngol 1991;100: 90–4.
27. Lano CF, Duncavage JA, Reinisch L, et al. Laryngotracheal reconstruction in the adult: a ten year experience. Ann Otol Rhinol Laryngol 1998;107: 92–7.
28. Ward PH, Canalis R, Fee W. Composite hyoid sternohyoid muscle grafts in humans. Arch Otolaryngol 1977;103:531–4.
29. Laccourreye O, Brasnu D, Cauchois R, et al. Tracheal resection with end-to-end anastomosis for isolated postintubation cervical trachea stenosis: long-term results. Ann Otol Rhinol Laryngol 1996; 105:944–8.
30. Kunachak S, Kulapaditharom B, Vajaradul Y, Rochanawutanon M. Cryopreserved, irradiated tracheal homograft transplantation for laryngotracheal reconstruction in human beings. Otolaryngol Head Neck Surg 2000;122:911–6.
31. Esclamado RM, Carroll WR. Repair of a complete glottic-subglottic stenosis with a fibular osseocutaneous free flap. Arch Otolaryngol Head Neck Surg 1997;123:877–9.
32. Gacek RR, Gacek MR, Montgomery WW. Evidence for laryngeal paralysis in cricoarytenoid joint arthritis. Laryngoscope 1999;109:279–83.
33. Laccourreye O, Paz Escovar MI, Gerhardt J, et al. CO2 laser endoscopic posterior partial transverse cordotomy for bilateral paralysis of the vocal fold. Laryngoscope 1999;109:415–8.
34. Ossoff RH, Duncavage JA, Shapshay SM, et al. Endoscopic laser arytenoidectomy revisited. Ann Otol Rhinol Laryngol 1990;99:764–71.

CHAPTER 51

Trauma to the Larynx

Peak Woo, MD, Philip Passalaqua, MD

ETIOLOGY

The position of the larynx in the neck, protected by the mandible superiorly and the clavicles and sternum below, shields it from falls and blows. The lateral mobility of the larynx within the neck also serves to prevent the severe effects of compression injuries. The ossification centers of the larynx do not develop until young adulthood and are more pronounced in the male. This makes laryngeal fractures a rare condition, one more often seen in men than in women.

Although the cartilaginous framework provides protection to the larynx, once this framework has been violated by trauma, the tight space defined by the laryngeal skeleton makes rapid airway compromise possible. The tight adherence of the soft tissue to the laryngeal framework and the potential spaces that can be easily expanded at the expense of the airway make rapid changes in the airway status possible in patients with laryngeal trauma.

External laryngeal trauma is a relatively uncommon injury. Jewett et al reported an incidence of blunt trauma to the larynx of 1 in 137,000 in an analysis of 54 million inpatients over a 5-year period in 11 US states.[1] Laryngeal trauma accounts for about 1 in every 25,000 emergency department visits. In addition to the obvious possibility of acute airway obstruction, the complications of late diagnosed or undiagnosed laryngeal injuries can be severe, including glottic insufficiency, aspiration, and laryngeal stenosis. Furthermore, studies have demonstrated that early diagnosis and surgical management result in improved functional outcomes.[1,2]

BIOMECHANICS OF LARYNGEAL TRAUMA

The larynx may be injured by blunt and penetrating wounds and may be traumatized by intubation and foreign bodies. In motor vehicle trauma, the head is often in extension. During deceleration, the larynx and neck are pulled forward while the thorax is held in restraint. The larynx is decelerated against blunt objects, such as the steering wheel and dashboard, and is crushed against the vertebral column. Similar mechanisms of injury may occur in sporting accidents (ie, karate, hockey, basketball, Frisbee throwing, water skiing). In laryngeal trauma, the severity of the injury is dependent on the velocity and site of injury. In mild cases, the patient may not report the injury for many days. Severe injuries may result in catastrophic asphyxia in the field.

The other major causes of blunt trauma to the larynx are of the "clothesline" type. The injury may be caused by a clothesline, chain, or tree branch. Injures of this type occur when the anterior part of the neck strikes one of these objects while bicycling, snowmobiling, or riding in all terrain vehicles. Clothesline injuries tend to be more severe because of the high impact of the force and therefore require special attention. In clothesline injuries, the possibility of cricotracheal separation is a special concern.

It appears that the larynx cannot be fractured by concussion unless it is supported to some extent by the vertebral column. Furthermore, a blow to the neck will do relatively little damage to the laryngeal skeleton if the laryngeal cartilage is soft and pliable. The cartilage may spring back into position without fracture. Such injury will result in soft tissue edema and hematoma. Soft tissue disruption may result in edema and hematoma severe enough to interrupt the airway or cause dysphonia and stridor. Rapid acceleration and deceleration of the soft tissue of the larynx may result in arytenoid cartilage dislocation and mucous membrane tears and disruption of the laryngeal ligaments.

TYPES OF LARYNGEAL TRAUMA

Soft tissue injuries to the larynx from blunt trauma include laryngeal edema, hematoma, and mucosal tears. Often edema or hematoma involves the

aryepiglottic folds and false vocal cords owing to the supraglottic submucosal distensibility.[3] Fluid or blood accumulations can occur rapidly, resulting in acute laryngeal obstruction.

During the course of a laryngeal trauma, the posterior ends of the thyroid cartilage alae may be compressed against the vertebral column. This pressure against the thyroid cartilage will splay the angle of the thyroid cartilage to a more obtuse angle. Therefore, most fractures of the thyroid cartilage occur in the midline with loss of the normal acute angulation of the thyroid alae (Figure 51–1). A similar force applied to the cricoid cartilage against the vertebral column will result in a comminuted fracture of the signet ring–shaped cartilage.

Most blunt injuries resulting in laryngeal trauma are caused by an anterior to posterior striking force. In strangulation injuries, the larynx may be injured by lateral to medial compressive forces. It is likely that because of the greater ossification of the cartilages and the greater development in both acute thyroid alae angulation and size, the male larynx is more brittle and more subject to fracture under the influence of violent impact. Because of the prominence of the thyroid cartilage and its thyroid notch, it is the cartilage most usually fractured. The cricoid cartilage is generally more protected. With the exception of isolated trauma to the lower part of the neck by a thin object, cricoid fractures are usually combined with thyroid fractures. When cricoid fractures are combined with thyroid cartilage fractures, they are usually more extensive and dangerous. A wide array of fractures, dislocations, and nerve injuries are also possible. Table 51–1 lists the various fractures and other trauma possible.

Gunshot and knife wounds account for the majority of penetrating injuries. Gunshot wounds produce a greater degree of soft tissue injury depending on the velocity and mass of the bullet. The higher-velocity bullets will result in greater injury to the surrounding soft tissues. Knife injuries

TABLE 51–1. Sites of Injury to the Larynx

Soft injury
Hematoma
Laceration
Cartilage injury
Thyroid cartilage fracture
Cricoid cartilage fracture
Combined fracture
Laryngotracheal separation
Hyoid bone fracture
Arytenoid cartilage dislocation

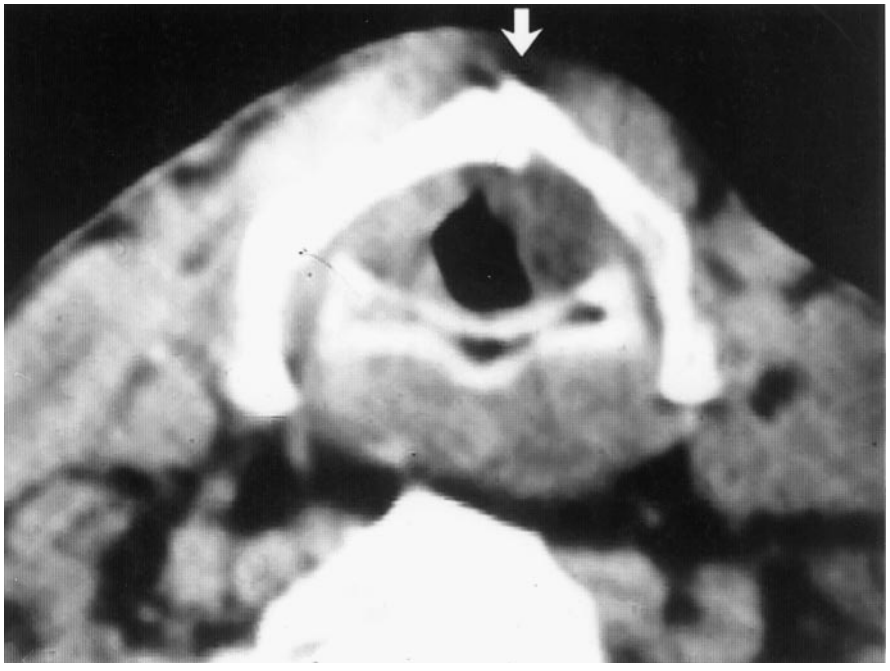


FIGURE 51–1. Axial computed tomographic scan of the larynx showing reduction in the anteroposterior dimension of the larynx and widening of the angle of the thyroid cartilage with a paramedian thyroid cartilage fracture (*arrow*).

usually result in less soft tissue damage but can be associated with injuries of vessels and nerves at some distance from the wound of entrance. During a knife injury, the larynx is often displaced without serious injury owing to the great lateral movement. A slash to the neck may result in lacerations of the cricothyroid membrane. Penetrating injuries of the larynx and trachea may occur in isolation but are usually part of more complex penetrating neck trauma that requires greater workup.

Laryngeal trauma in children tends to be less severe because of the higher position of the larynx in the neck in the pediatric population. Furthermore, owing to the pliable cartilage of a child's larynx, there is a lower incidence of laryngeal fractures. In addition, there is a lower incidence of laryngotracheal separation in the pediatric population because of the narrow cricothyroid membrane.[4]

Intubation injuries to the larynx usually result in soft tissue trauma to the posterior part of the larynx caused by an endotracheal tube or cuff. The most common occurrence is ulceration of the mucous membrane overlying the vocal process of the arytenoid cartilage. More severe injuries may be caused by prolonged intubations. These may take the form of intubation granuloma, posterior synechia, glottic stenosis (Figure 51–2), cricoarytenoid joint ankylosis, and subglottic stenosis. The factors contributing to poor outcomes in intubated patients include prolonged intubation, infection, gastroesophageal reflux, movement, pressure necrosis, and chondritis. The arytenoid cartilage can be dislocated by traction of the endotracheal tube that can cause a tear of the posterior cricoarytenoid ligament. In addition, an endotracheal tube pushed into the larynx without direct guidance through the glottis may result in a posterior vector against the cricoarytenoid joint, resulting in a posterior, lateral dislocation of the arytenoid cartilage. This may be associated with a tear of the anterior cricoarytenoid ligament.

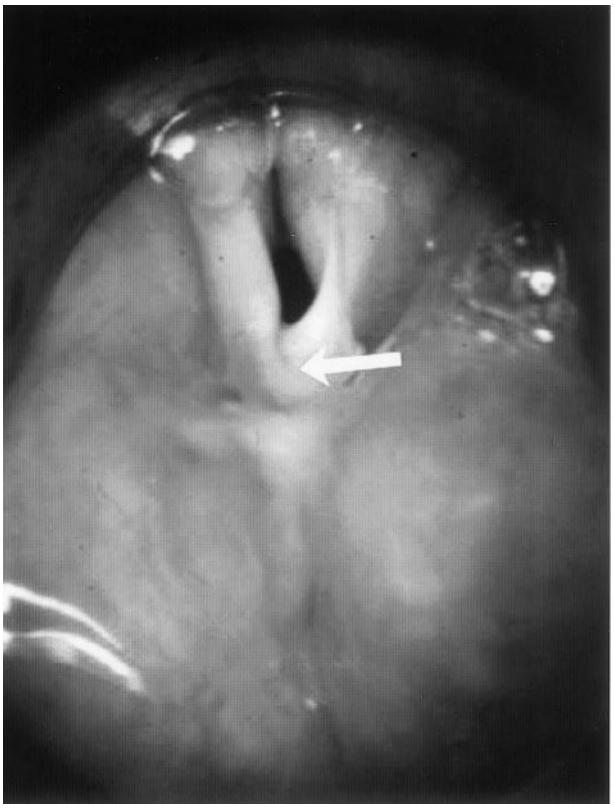


FIGURE 51–2. Endoscopic view of posterior laryngeal stenosis secondary to prolonged intubation. *Arrow* indicates a posterior stenotic scar.

DIAGNOSIS

The symptoms and signs depend on the cartilage involved and the extent of the soft tissue injury. Patients with laryngeal trauma may present to the emergency department with multiple injuries to other organ systems. Therefore, the laryngeal injury that has not caused acute airway obstruction may be missed. One should have a high index of suspicion for laryngeal injury in any patient presenting with blunt neck trauma. Diagnosis of subtle laryngeal injury in a timely manner is important for both treatment and prognosis.

All patients with injury to the anterior part of the neck should be carefully evaluated for the existence of laryngeal trauma. The symptoms suggestive of laryngeal injury include cough and expectoration. Pain is present in nearly every patient with laryngeal injury and is accentuated by phonation or deglutition. Tenderness and swelling may be quite marked. Swelling of the neck may be accompanied by loss of laryngeal landmarks in the anterior part of the neck. Dyspnea is often present in varying degrees owing to edema of the soft tissue or blood in the trachea. Hemoptysis may be present but is usually not severe. Stridor and voice change are other signs that suggest laryngeal trauma. Although dysphagia may be present with endolaryngeal trauma, it may also be suggestive of esophageal or hypopharyngeal injury. Emphysema of the neck suggests a perforation of a

viscus such as the larynx or hypopharynx. Soft tissue injuries may be accompanied by ecchymosis of the skin. In more severe injuries of the larynx in which the airway compromise is marked, asphyxiation and massive hemoptysis may occur. Crepitation of the neck secondary to subcutanous air may be elicited (Figure 51–3).

WORKUP FOR LARYNGEAL TRAUMA

The field management of the patient with suspected laryngeal injury consists of stabilization of the cervical spine and establishment of an airway. In the event of life-threatening airway compromise, intubation or tracheostomy in the field may be necessary.

After the patient has been stabilized and other potential life-threatening injuries are under control, the physical examination and assessment of the neck may commence for evaluation for laryngeal trauma. The workup for laryngeal trauma consists of physical examination of the neck, fiber-optic laryngoscopic examination, radiologic examination, and operative laryngoscopy.

Stridor is the most common sign in patients with upper airway compromise. The type of stridor may be indicative of the location of the injury. Combined inspiratory and expiratory stridor suggests some degree of obstruction at the level of the glottis. Expiratory stridor is more consistent with a lower airway injury. On the other hand, inspiratory stridor is indicative of supraglottic airway obstruction.[5] The presence of both stridor and hemoptysis has been associated with severe laryngeal trauma, displaced fractures of laryngeal cartilage, significant endolaryngeal or laryngopharyngeal edema or hematoma, or large mucosal tears exposing cartilage.[6]

A thorough physical examination must be performed with special attention to the presence of neck tenderness, crepitus owing to subcutaneous emphysema, soft tissue swelling, and loss of thyroid cartilage prominence. Fiber-optic laryngoscopy in the stable patient is an essential element of the physical examination and should focus on vocal cord mobility, tears, mucosal edema, hematoma, and dislocated or exposed cartilage. If limited range of motion of the vocal cords is noted, a structural deformity or arytenoid dislocation is likely. On the other hand, immobility of the vocal cords suggests recurrent laryngeal nerve injury.[5]

The patient presenting with soft tissue injury alone will often display edema, submucosal hemorrhage, and ecchymosis. Laceration of the mucosa, exposed cartilage, arytenoid cartilage dislocation, and disruption of the laryngeal architecture are highly suspicious of laryngeal framework injury. Despite the above generalizations, some patients

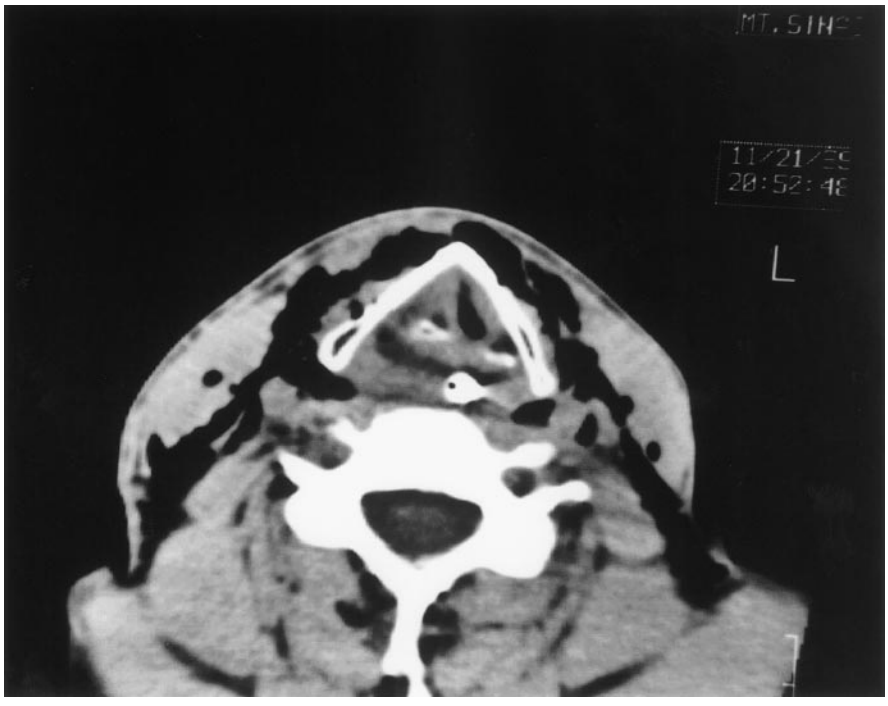


FIGURE 51–3. Axial computed tomographic scan showing massive subcutaneous emphysema in a patient with a laryngeal fracture.

with minimally displaced fractures of the larynx will present with a benign appearance of the larynx, and the fracture will be missed until the laryngeal edema has subsided, leaving the dislocation or laryngeal framework abnormality to be manifest as persistent hoarseness or dysphagia.

RADIOLOGIC EVALUATION

If cervical spine injury is suspected, it must be evaluated with cervical films prior to movement of the head. Chest radiography should be done to rule out a pneumothorax. Conventional neck radiographs can be used to screen for subcutaneous emphysema, whereas, historically, fluoroscopy was used in assessment of vocal function. With the advent of computed tomographic (CT) scanning, this study has supplanted laryngograms and laryngeal tomography. The laryngeal CT scan is the most useful radiologic examination for evaluating laryngeal injury. There is some debate regarding the utility of CT scans in cases at the ends of the disease spectrum. Stringer and Schaefer do not recommend CT scan in those individuals with minimal cervical trauma and a normal physical examination as the scan will not change the management. They also do not recommend CT scans in individuals who have penetrating injuries, obvious fractures, and large lacerations and in those who will require open exploration. They do note that there are exceptions for which the CT scan can serve as a roadmap for structural repair.[5] Bent et al routinely used CT scans to evaluate severe injuries that required operative management. They found that CT scans assisted in anticipating specific injuries and aided with the planning of the operative procedure.[7] In the vast majority of laryngeal trauma cases, CT scans offer better visualization of the subglottic and anterior commissure areas, can help identify clinically unapparent cartilage fractures, and can verify laryngoscopic findings.[8] Provided that the radiographic facilities are present and the patient is stable, a 2 mm CT scan or a spiral CT scan through the larynx offers the maximal detail as to the extent of laryngeal framework injury. More recently, with the use of spiral CT scans and three-dimensional reconstruction, virtual laryngoscopy and bronchoscopy have become possible. Their role in laryngeal fracture diagnosis and management has yet to be confirmed.

Carotid arteriography may be indicated in penetrating injuries to help reveal vascular injuries. If esophageal or pharyngeal tears are suspected, a Gastrografin (meglumine diatrizoate) swallow may be indicated. Gastrografin is the preferred agent over barium owing to the possibility of a fistula. In these cases, Gastrografin is less irritating to the soft tissues.

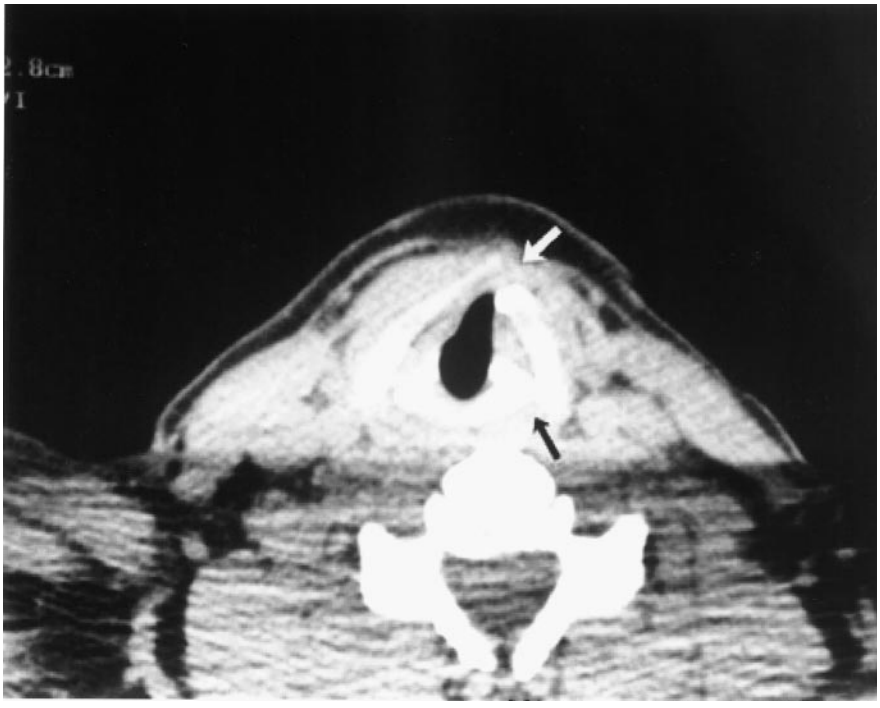


FIGURE 51–4. Axial computed tomographic scan showing a thyroid cartilage fracture (*white arrow*) and left arytenoid cartilage dislocation (*black arrow*) laterally and inferiorly.

OPERATIVE ENDOSCOPIC EVALUATION

Operative endoscopic examination by direct operative laryngoscopy, esophagoscopy, and bronchoscopy allows the treating surgeon to assess fully the extent of mucosal injury and disruption of the laryngeal framework. Direct visualization of laryngeal structures has advantages over flexible laryngoscopy or radiographic examination in that palpation of the structures may be performed. Subglottis and tracheal injuries may be missed by fiberoptic examination; injuries to these areas are better defined by operative laryngoscopy. Mucosal tears and cartilage exposure that appear as soft issue injuries or edema by CT scan can be mapped without difficulty. Provided that the operative risks are minimal, the managing physician should have a low threshold for operative endoscopic evaluation of the patients with suspected laryngeal trauma.

EMERGENCY MANAGEMENT

In the initial management of laryngeal trauma, securing and stabilizing the airway are of paramount importance. However, there is controversy regarding the optimal method of establishing a patent airway. Oral intubation in the setting of laryngeal trauma can be precarious owing to the possibility of further damaging the larynx, the creation of false passages, and precipitating loss of a tenuous airway. For these reasons, it is recommended that if intubation is necessary, it is to be done after cervical spine injuries have been ruled out and the laryngeal anatomy has been defined by prior fiber-optic laryngoscopy and that it be done in a setting with access to direct visualization with operative bronchoscopes and with equipment and personnel required for a tracheostomy. A clinician experienced in airway management should carry out the procedure of intubation.

Owing to the inherent risks of intubation in the setting of laryngeal trauma, many authors recommend a tracheostomy under local anesthesia.[7,9,10] A tracheostomy provides a method of stabilizing the airway with less risk of loss or further injury of the airway.[5] Optimally, tracheostomies for laryngeal trauma should be performed at the fourth to fifth ring of the trachea. By performing the tracheostomy lower than usual, further damage to existing injuries can be avoided. A vertical incision also affords better exposure so that, in the case of laryngotracheal separation, the trachea may be better accessed if it has retracted inferiorly.[11] Cricothyrotomy should be used only as an emergency method of airway establishment and should be converted to a standard tracheostomy as soon as possible. The possible sequela of not recognizing a cricothyrotomy in the setting of laryngeal injury include (a) the subsequent development of perichondritis and necrosis of the cricoid cartilage, (b) the development of subglottic stenosis, and (c) injury to the conus elasticus and the vocal cords. Therefore, any tracheostomy done in the field or in the emergency setting should be carefully evaluated for the possibility that it is a cricothyrotomy or that it has caused cricoid injury. In the event that a high tracheostomy or a cricothyrotomy is present, the airway access should be converted to a conventional tracheostomy below the second tracheal ring. The preferred time for revision tracheostomy after cricothyrotomy is within 24 hours after the initial operation.

CONSERVATIVE MANAGEMENT

After the initial evaluation and management of emergent airway issues, further treatment is divided into nonoperative and operative management. A conservative approach can be taken in patients with minimal soft tissue swelling or small hematomas who are stable and without evidence of respiratory compromise. Conservative management should be reserved for soft tissue injuries and nondisplaced laryngeal fractures. Hematoma and edema of the larynx often resolve spontaneously if there is no evidence of other injury. Conservative management is also appropriate in small mucosal lacerations or avulsions that do not involve the free edge of the vocal cord or anterior commissure. Observation is also appropriate when there is a nondisplaced single thyroid cartilage fracture without exposed cartilage and mucosal laceration.

Endoscopic evaluation of the larynx and trachea should be carried out if there is suspected laryngeal fracture, mucosal laceration, or emphysema of the neck. If there is extensive soft tissue swelling with airway compromise but no evidence of framework injury or mucosal laceration, the surgeon may elect to perform a tracheostomy without open surgical exploration of the larynx. Putting the larynx

at rest with conservative treatment will allow the optimum chance for edema to subside while safeguarding the airway.

Conservative management consists of observation, bed rest, voice rest, humidification of inspired air, antacids, and antibiotics. If a tracheostomy is not necessary, the patient is observed for 24 hours in an inpatient setting. During the 24 hours of observation, the patient should be followed for evidence of airway compromise from progressive swelling. In an effort to prevent further swelling and reduce the edema present, bed rest with elevation of the head is indicated. Similarly, the goal of voice rest is to reduce swelling. Corticosteroids have been used for their anti-inflammatory properties and anecdotally prevent the progression of edema. Humidification of inspired air is thought to prevent crust formation if mucosal damage is present.[12] Koufman has suggested that antacid use may be beneficial by decreasing the incidence of mucosal irritation and possible laryngeal stenosis from gastroesophageal reflux.[13] Antibiotics are indicated, especially in patients with multiple fractures in whom there is an increased risk of infection and perichondritis. Furthermore, a clear liquid diet may be beneficial initially. A nasogastric tube should be avoided if possible owing to the risk of further injury during placement. In addition, with extended placement of a nasogastric tube, there is risk of mucosal ulceration in the postcricoid region.[3]

SURGICAL MANAGEMENT

Injuries requiring surgical intervention include those with (1) exposed cartilage, (2) large mucosal lacerations, (3) lacerations involving the free edge of the vocal cord, (4) vocal cord immobility, (5) dislocated arytenoid cartilages, (6) displaced cartilage fractures, and (7) any neck injury with airway obstruction. The timing of surgical repair has been debated in the past, with some authors arguing that delaying surgery 3 to 5 days allowed resolution of edema and therefore better recognition of the anatomy. However, the current consensus is that outcomes are improved with early intervention. In one study, significantly better outcomes were noted in the group that was surgically repaired within 24 hours when compared with the group that was treated within 2 to 7 days.[2] A delay of greater than 24 hours to surgical repair has also been found to be associated with an increased incidence of wound infections, especially with supraglottic injuries.[14] Similar results have been reported in several studies with regard to the voice. Waiting longer than 48 hours for surgical repair is associated with increased laryngeal dysfunction.[3] Early intervention may help reduce "uncontrolled healing" of lacerations that result in scar formation.[7]

The immediate goals of surgical repair may be guided by the extent of the deficit and the needs of the patient. Clearly, intervention in a laryngeal fracture may have to be delayed in an unstable patient with multiple trauma. In contrast, the patient with a laryngeal fracture who is a voice professional will want to have everything possible done to restore voice function despite a stable airway. The early goals of surgical intervention should be to establish a safe airway, document the injury, repair cartilage injuries and restore the laryngeal framework, repair soft tissue lacerations, and promote early healing of the injury by prevention of infection and tissue necrosis. In the repair of laryngeal trauma, the long-term goals should be optimum functional restoration of voice, airway, and swallowing.

Although not a common injury, surgical management of laryngeal fractures has many possible variations. Some of these variations depend on the method of fracture reduction, whereas others depend on the method of stabilization and stenting technique used. Another variation depends on the site of fracture. Fractures of the thyroid and cricoid cartilages may be repaired by open reduction and rigid fixation, whereas arytenoid cartilage dislocation and laryngotracheal separation are more often treated by repair followed by endolaryngeal stenting. Table 51–2 outlines some of the types of injuries to be considered in the repair of laryngeal trauma.

In an effort to optimize the outcomes in laryngeal fractures, unique aspects of the laryngeal anatomy and function must be considered. The following factors are worth detailing: (a) laryngeal cartilages without great blood supply are in contact with mucosa; (b) the mucosal surface is in contact with contaminated upper airway secretions, making infection after injury possible; (c) the larynx is subject to great movement during normal deglutition and respiration, making stabilization difficult; and (d) the relatively small laryngeal cartilages have a great number of intrinsic and extrinsic muscles attached, making functional outcomes uncertain after laryngeal fracture. To counter these unique

TABLE 51–2. Types of Laryngeal Injuries: Causes, Manifestations, and Treatment

Type of Injury	*Causes*	*Symptoms and Signs*	*Treatment*
Fracture of the cricoid cartilage	MVA, closed injury Emergency tracheostomy	Dysphonia, odynophagia, +/– airway obstruction	ORIF with or without stenting
Fracture of the thyroid cartilage	MVA, closed injury Penetrating injury	Dysphonia, odynophagia, +/– airway obstruction	ORIF with or without stenting
Arytenoid cartilage dislocation	Closed injury Intubation	Hoarse voice, aspiration, dysphagia, +/– airway obstruction	Arytenoidpexy, stenting
Laryngotracheal separation	Vehicular trauma	Airway obstruction	Operative repair or resection, stenting
Complex fractures	High-speed MVA, GSW	Airway obstruction	Operative repair or resection, stenting

MVA = motor vehicle accident; GSW = gunshot wound; ORIF = open reduction internal fixation.

problems in the repair of laryngeal fractures, the laryngeal surgeon should (a) strive to reduce the amount of mucosal loss and hasten mucosal healing, (b) improve cartilage survival by rigid fixation and movement reduction, and (c) if necessary, prevent cicatrix formation during healing by the use of endolaryngeal stenting.

The first step in any planned surgical treatment of a laryngeal fracture is a rigid direct laryngoscopy to evaluate the extent of mucosal injury and to palpate the arytenoid cartilages for dislocation or immobility. Laryngeal exploration is performed through a horizontal incision at the level of the cricothyroid membrane. Subplatysmal flaps are elevated from the level of the hyoid bone to the sternal notch. The strap muscles are separated vertically at the midline raphe and retracted laterally. The larynx is entered, depending on the site of injury, through a midline thyrotomy or through the thyroid cartilage fracture if the fracture line runs within 2 to 3 mm of the midline of the thyroid cartilage. Care is taken to avoid injury to the anterior commissure. This can be accomplished by horizontally dividing the cricothyroid membrane prior to performing the midline thyrotomy and directly viewing the undersurface of the vocal folds from below. If there is disruption of the anterior commissure tendon, the surgeon must carefully search for and reattach it to the thyroid cartilage.

After exposure of the endolarynx, any mucosal lacerations are approximated with fine absorbable sutures. Fast-absorbing mild 4-0 chromic sutures are used. If possible, knots should be placed outside the lumen to prevent formation of granulation tissue. In most instances, mucosal injuries can be repaired with adjacent mucosa. When mucosal injuries are large, as may be present in an arytenoid cartilage avulsion, a piriform sinus mucosal flap may be used to cover any exposed cartilage. If there is extensive mucosal damage in the membranous portion of the vocal fold, re-establishment of mucosa may be accomplished by the use of free buccal mucosal or dermal grafts. A careful examination of the repair should show that all of the exposed cartilage is covered without tension. Pedicled mucosal flaps and mucosal or dermal grafts are used if there is not enough adjacent mucosa within the larynx to cover all exposed cartilage. These free grafts are of second choice because there is a greater likelihood of scar formation with their use compared with local mucosal flaps. In addition, these free grafts carry the morbidity associated with additional incisions.[5] The importance of restoring the lining of the larynx is attributable to the inherent tendency of exposed cartilage to facilitate the formation of granulation tissue, subsequently leading to fibrosis and laryngeal stenosis. There is also an increased risk of perichondritis when cartilage is left exposed. Care should be exercised to avoid aggressive débridement of soft tissue or mucosa. Because of the unique nature of the laryngeal framework and the attached mucosa, only devitalized tissue should be removed. Any removed tissue should be replaced by local adjacent tissue

transfer or soft tissue grafts. This is to prevent late glottal incompetence.

Another basic principle of laryngeal trauma surgery is to preserve as much cartilage as possible. Even free fragments of cartilage can be used to provide a scaffolding and stability, provided that they are covered with viable mucosa. If free cartilage fragments are used, they should be fixated with suture or miniplates.

Mildly displaced simple fractures can be repaired with nonabsorbable sutures through the outer perichondrium. Displaced fractures of laryngeal cartilage are reduced and immobilized. To prevent chondritis, very small pieces of cartilage without intact perichondrium are débrided. Fractures can be fixated with wire, nonabsorbable suture, or miniplates. If suture is used, fracture-site sutures should not be tightened until all fractures are reduced to avoid further damage to the mucosa by tearing. Internal fixation of laryngeal fractures may be performed if there are multiple fractures or fractures with displacement, or it may be used instead of internal laryngeal stenting. Titanium miniplates offer the advantage of superior alignment of displaced fractures because of the three-dimensional nature of miniplates (Figure 51–5).[15]

Reconstruction plates used in facial fracture repair may be used in laryngeal fracture repair. These plates are bioimplants made of titanium, stainless steel, or polyglycolic acid. Being malleable, these plates can be used as an internal scaffold for sutures or screws to fix the laryngeal fracture fragments into position. The three-dimensional contour of the laryngeal skeleton may be re-established with a single reconstruction plate bent to the shape of the thyroid cartilage. For the cricoid cartilage, it may be used to stabilize comminuted fragments and provide rigid fixation to the fragments. The screws and plates are best used in ossified cartilages but may be used with care in nonossified cartilage. Because of the soft nature of the cartilage, surgical tapping and drilling of the cartilage are not necessary. In general, four to six 5.0 mm by 2.0 mm screws serve to fix the laryngeal skeleton. Although we have also tried microplates, the smaller screws and the tendency to strip the smaller holes make the miniplate system more useful. In the rare adolescent or pediatric patient in whom the cartilage is thin, sutures alone are used to fixate the fragments onto the miniplate scaffold, thereby achieving stabilization without the need for screws.

In situations in which rigid internal fixation or immobilization cannot be achieved, laryngeal stenting will be necessary. Stents are also indicated when there is injury to the anterior commissure or when extensive soft tissue injury is present. If mucosal

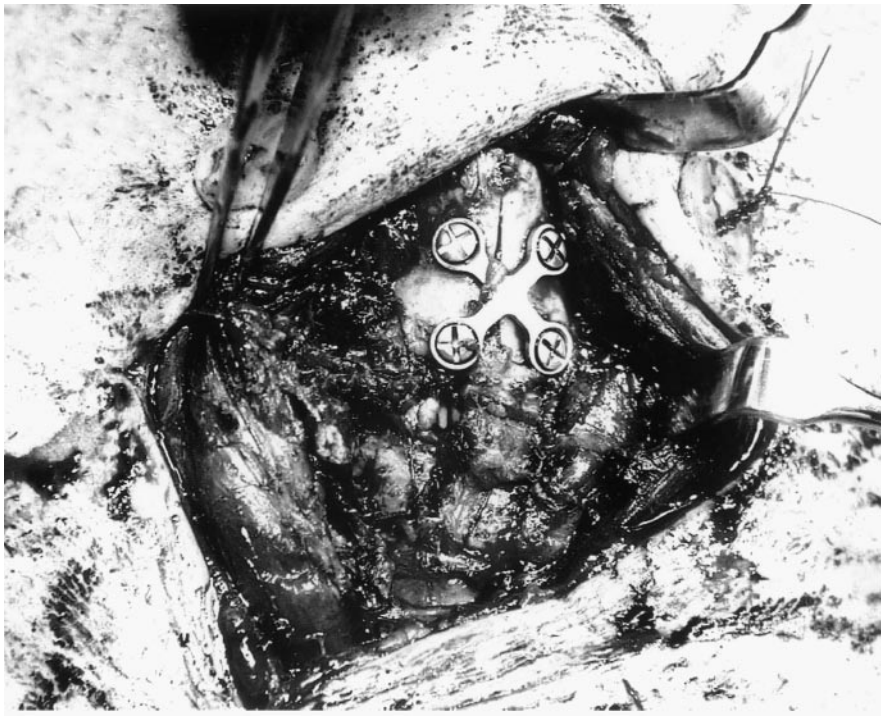


FIGURE 51–5. Miniplate used to fixate a thyroid cartilage fracture rigidly.

grafting is required, the internal stent serves to hold the graft in approximation to the raw surface to be grafted. The goal of stenting is to provide stability and prevent mucosal adhesions and subsequent laryngeal scarring. The type of stent to be used is controversial. Soft stents may be made of a finger cot packed with antibiotic-impregnated cotton gauze. Soft semirigid stents made of Silastic are commercially available in different sizes, from child to female to male. These can be inserted and modified according to need. Theoretically, "soft stents" result in less damage to surrounding tissue than firmer stents. However, many authors have found an increased risk of infection and formation of granulation tissue with "soft stents."[7,16] Firm stents are usually made of silicone, Silastic, or Portex. They are secured in a similar fashion as soft stents.

The stent is fixed with heavy nonabsorbable suture so that the superior end is at the aryepiglottic folds and the inferior end is above the tracheostomy site. The stent should be placed so that it moves with the larynx during swallowing. It can be secured with two mattress sutures through the skin, then through the thyroid cartilage below the level of the glottis, and then through the stent and out the cartilage and skin on the opposite side. The sutures can be fastened to buttons on the skin. Regardless of the method of suturing, the stent should be easily removed at endoscopy. Furthermore, the stent should be fixed in such a manner that if a suture breaks, the stent may still be recoverable.[5]

There has been controversy regarding the optimal time period that stents should remain in place. The beneficial effects of laryngeal stabilization and prevention of scar formation should be weighed against the risk of infection and irritation leading to granulation tissue and scar formation with subsequent laryngeal stenosis. Previously, longer periods of time, up to 6 weeks, were favored for stent placement. Currently, most authors recommend that stents be removed after about 2 to 3 weeks, providing that fractures have been stabilized and lacerations closed effectively.[2,5] After the stent is removed at direct laryngoscopy, the need for additional removal of granulation tissue can be assessed. If granulation tissue requires removal, a carbon dioxide laser can be used.[5]

Despite controversy regarding the management of severe laryngeal trauma involving comminuted fractures and significant endolaryngeal injury, there is consensus that neither grafts nor stents are substitutes for meticulous primary closure of mucosal wounds and diligent reduction and internal fixation of fractures.

Some specialized injuries from laryngeal trauma may occur as a consequence of closed or open injuries. These merit individual discussion below.

Dislocation of the Cricothyroid Joint

Usually with trauma, the inferior aspect of the thyroid cartilage is displaced posteriorly. As a result of this displacement, the recurrent laryngeal nerve may be compromised near the cricothyroid joint, resulting in vocal cord paralysis. Vocal fold paralysis owing to cricothyroid joint avulsion is often temporary, and function usually returns in 3 to 4 weeks. If no recovery of function is noted after 8 to 12 weeks, spontaneous function is unlikely to return. At this point, decompression of the nerve by resection of the inferior cornu of the thyroid cartilage may be considered since the cause of paralysis in many cases is neuropraxia. Electromyographic (EMG) studies of the vocal cord may be useful in assessing the utility of decompression. If the vocal cord is paralyzed longer than 9 months, the long-term prognosis is poor.[3]

Supraglottic Injuries

Supraglottic injuries require a more aggressive surgical approach owing to the high incidence of infection and fibrosis. In some isolated supraglottic injuries, the epiglottis may be avulsed from its anterior attachments, thereby allowing the epiglottis to prolapse posteriorly into the laryngeal lumen. In these instances, the epiglottis may be resected with no alteration in swallowing function. If the trauma is more extensive, a portion of the thyroid cartilage and false vocal cords can be resected in addition to the epiglottis as in a standard horizontal supraglottic laryngectomy.[3] Generally, supraglottic injuries result in horizontal fractures of the superior aspect of the thyroid alae, resulting in avulsion of the epiglottis superiorly and posteriorly. The epiglottis is then prone to fixation in the posterior position owing to scarring. In addition, fixation of the arytenoid cartilages can occur in untreated injuries, leading to eventual laryngeal stenosis. One exception to the above approach is when there is associated vocal cord paral-

ysis or glottic trauma. With vocal fold paralysis or concomitant glottic injuries, supraglottic resection should not be done owing to the increased incidence of aspiration. Rather, the supraglottic region should be reconstructed and the epiglottis removed.[3]

Glottic Injuries

With glottic injuries, fractures of the thyroid cartilage often occur at the level of attachment of the true vocal cords. Dislocation of one or both arytenoid cartilages with resultant vocal cord immobility is common (Figure 51–6). Dislocated arytenoid cartilages are realigned in their correct position with relation to the other laryngeal structures. These cartilages should not be fixated or removed. As with mucosal tears in other areas, primary reapproximation of the tissue should be attempted. If the mucosa cannot be closed primarily, advancement flaps of mucosa of the piriform sinus area can be used. If vocal cord immobility is present and is secondary to nerve injury, no corrective action should be taken since spontaneous return of vocal cord function is common.

Cricotracheal Separation

Cricotracheal separation is a devastating injury with unique characteristics that require close attention. Unfortunately, most of these patients do not survive the acute airway obstruction that occurs. If they do make it to the hospital alive, the airway is often tenuous owing to the loss of cricoid support. There is also a high likelihood of injury to the recurrent laryngeal nerves. Once an airway is secured through tracheostomy, definitive repair can be planned. These patients are at high risk of development of subglottic stenosis.

Cricoid Cartilage Fracture

Often the cricoid cartilage is fractured in two sites laterally, resulting in a telescoping of the cricoid cartilage anteriorly to posteriorly. If this is the case, the cartilage fractures should be exposed, reduced, and fixated, often requiring a laryngeal stent (Figure 51–7). Care should be taken to avoid injury to the cricothyroid muscle when repairing cricoid cartilage fractures. When the cricoid cartilage is intact, mucosa should be approximated with absorbable suture. To lessen the tension on this repair, sutures can be placed from the upper border of the cricoid cartilage around the first or second tracheal ring.[5] In situations in which the cricoid cartilage fracture is comminuted, internal fixation must be performed to offer stabilization during restructuring of the cricoid. As much as one-third of the anterior part of

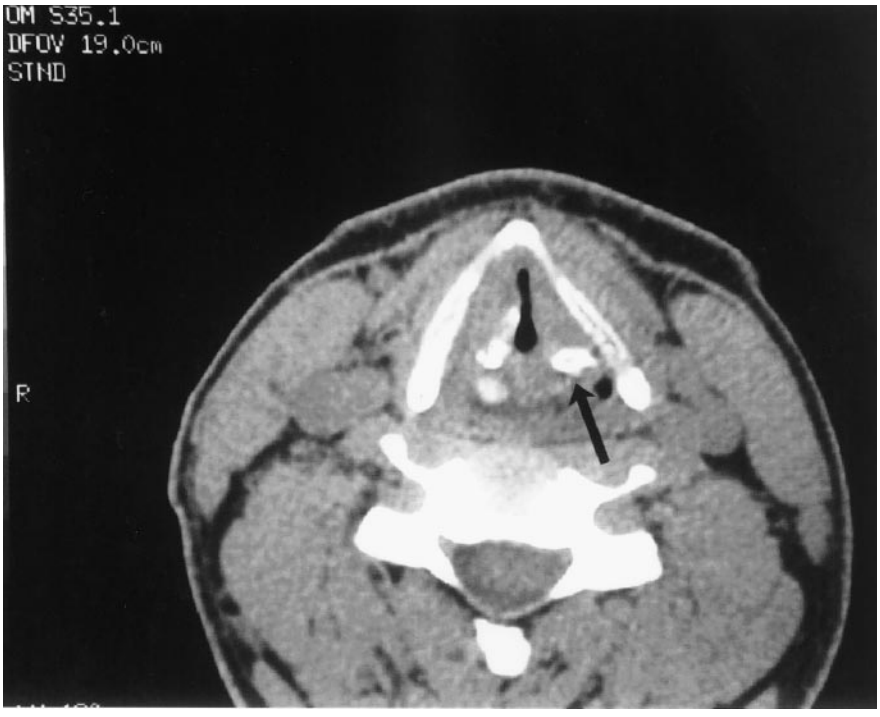


FIGURE 51–6. Axial computed tomographic scan showing arytenoid cartilage dislocation (*arrow*).

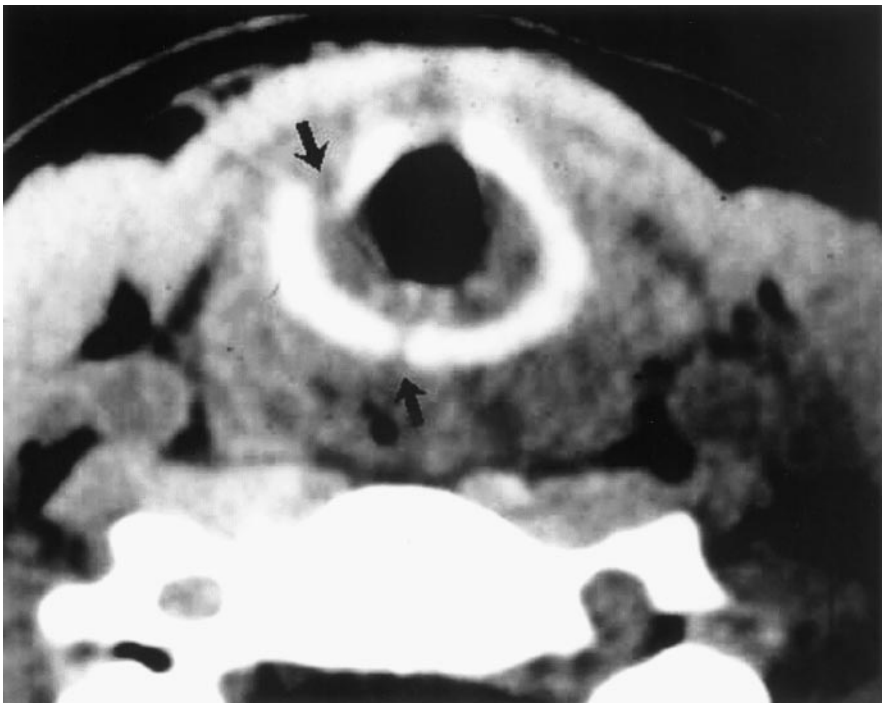


FIGURE 51–7. Axial computed tomographic scan showing cricoid cartilage fracture. *Arrows* indicate fracture sites.

the cricoid cartilage can be repaired using the sternohyoid muscle and its overlying fascia.[17] In patients with combined compromise of the trachea and the cricoid cartilage, sleeve resection and primary repair may be performed.[18]

Recurrent Laryngeal Nerve Injury

If the recurrent laryngeal nerve is severed and this is recognized at the time of exploration, the best approach is to reanastomose the transected nerve. Although vocal fold mobility may not be regained, muscle atrophy can be minimized through nerve regeneration, thereby improving the quality of the voice.[5]

Penetrating Laryngeal Injuries

The basic principles of management of blunt laryngeal injuries apply to penetrating injuries to the larynx. Owing to the especially destructive nature of high-velocity gunshot wounds, these injuries deserve special comment. With gunshot wounds, tissue viability is difficult to assess initially. The high energy that the bullet imparts to the soft tissue leads to a wider area of significant injury than simply the path of the missile. Furthermore, the course of the bullet is often erratic. Some authors recommend delaying definitive surgical repair so that the extent of nonviable tissue becomes apparent.[3,6] Other authors do not suggest delaying surgical intervention but rather emphasize that a wide area must be excised because the extent of devitalized tissue is not evident initially.[2,9,17] In either situation, all viable structures should be conserved. However, owing to the severe structural disruption of this type of injury, stenting is often necessary. At times, the injury is so massive that partial or total laryngectomy is indicated.

Unlike gunshot wounds, knife injuries do not impart a significant force to soft tissue distant from the course of the penetrating object. Therefore, the risk of nonviable soft tissue distant from the penetrating course of the knife is not present.

Pediatric Laryngeal Trauma

The same basic management principles discussed for the adult population also apply to the pediatric population, with a few additional caveats regarding securing a child's airway. Owing to the rapid oxygen desaturation of a child compared with an adult, efficiency in obtaining and stabilizing the airway is paramount. Furthermore, tracheostomy with local anesthesia is not a reasonable option in a scared, injured child. The airway should be stabilized under direct visualization with bronchoscopy; if needed, a

tracheostomy may then be performed over the bronchoscope.[4] During endoscopy following the tracheostomy, the laryngeal injuries can be assessed.

COMPLICATIONS

Even with the prompt recognition of laryngeal injuries and appropriate management, complications including granulation tissue, laryngeal stenosis, vocal cord paralysis, and aspiration are possible. Untreated displaced fractures may heal with dystrophic chondrification (Figure 51–8). Chondronecrosis with prolonged granulation tissue and progressive cicatrix formation is a feared complication of laryngeal fracture. Factors that increase the formation of granulation tissue include the presence of cartilage without mucosal covering and the presence of a stent. Granulation tissue increases the risk of fibrosis, which subsequently leads to stenosis. The primary closure of lacerations and the careful covering of all exposed cartilage with mucosa are of paramount importance. Other techniques have been attempted to reduce the formation of granulation tissue. These methods include intralesional and systemic corticosteroids, splinting, and irradiation. However, none of these treatment options have proven to be beneficial. Therefore, the optimal methods of preventing granulation tissue are the initial complete mucosal closure over cartilaginous structures and maintaining a stent in place the shortest period of time required for fixation.

Laryngeal stenosis may develop despite meticulous care after injury. In the case of supraglottic stenosis, the scar tissue may be excised with the carbon dioxide laser. If needed for wound coverage, local mucosal flaps or buccal mucosal grafts can be used. Similarly with glottic stenosis, posterior or interarytenoid scarring can be excised and local advancement flaps used if necessary. In the case of thin webs, a laser may be used to divide the web. If the glottic stenosis is severe, a laryngofissure with direct excision of the stenotic area is likely necessary. In these severe cases, a stent with a tissue graft is often placed to facilitate re-epithelialization.[12] In late or repeat reconstructions of the larynx in which laryngeal stenosis is an ongoing issue, the use of a conforming laryngeal prosthesis is invaluable. This application is described by Montgomery and Montgomery.[19]

Laryngeal stenosis in the subglottic area is more difficult to treat effectively. With limited subglottic stenosis, excision of fibrotic tissue with a carbon dioxide laser and repeated dilations of the stenotic area may be adequate. In cases of more severe subglottic stenosis, cricoid splits with cartilage grafting are required. To stabilize the restructuring postoperatively, a stent is often needed.

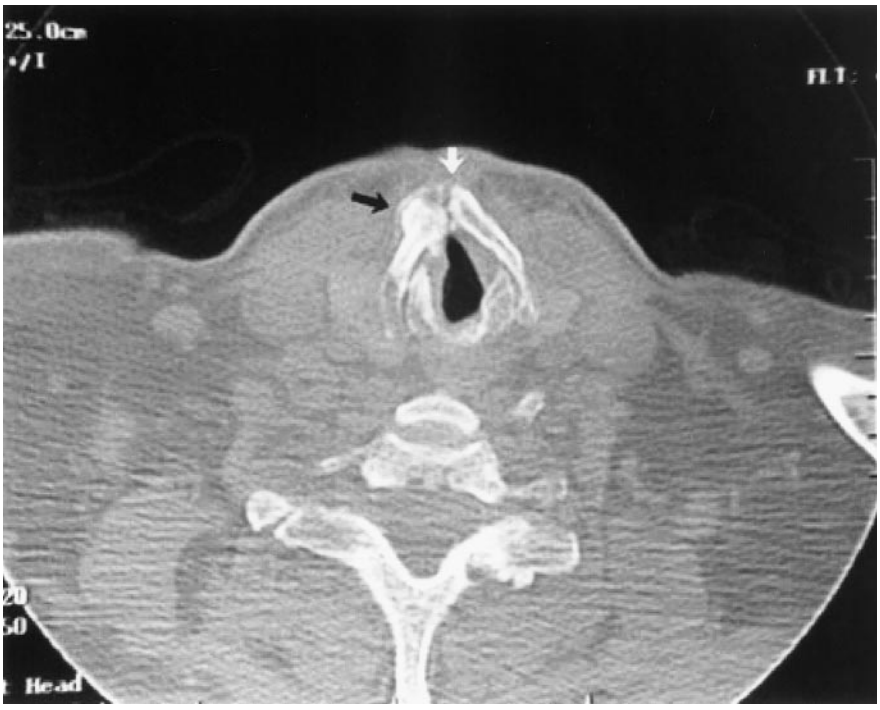


FIGURE 51–8. Axial computed tomographic scan showing dystrophic calcification several months after a missed thyroid cartilage fracture. *Black arrow* indicates dystrophic calcification. *White arrow* indicates thyroid cartilage fracture.

Tracheal stenosis of up to 4 cm can be resected with subsequent end-to-end tracheal anastomosis.[12] Late reconstructions of the larynx using pedicled and autogenous grafts have been used with good results in adults suffering from laryngeal stenosis after laryngeal trauma.[20–23]

In addition to laryngeal stenosis, another possible complication is paralysis of the vocal cords. The cause of an immobile vocal cord must be determined to make appropriate management decisions. To evaluate whether an injured recurrent laryngeal nerve or a cricoarytenoid joint fixation is responsible for vocal fold dysfunction, direct laryngoscopy is performed. During the laryngoscopy, the arytenoid cartilage is palpated to evaluate mobility. If the arytenoid cartilage is mobile, injury of the recurrent laryngeal nerve is the probable cause. If any question still exists, laryngeal EMG can also be performed. Electromyography is a valuable way to evaluate if the laryngeal muscles are denervated or the recurrent laryngeal nerve is intact. If the recurrent laryngeal nerve is severed, the EMG will show total denervation of the muscles. If the nerve is injured but recovering, the EMG will show signs of electrical recovery. If the EMG shows that the muscle activity is intact, the vocal cord immobility is likely secondary to cricoarytenoid joint fixation.

Bilateral vocal fold fixation or bilateral laryngeal nerve paralysis is especially challenging in the management of late effects of laryngeal trauma. Reinnervation procedure using the ansa cervicalis to the recurrent laryngeal nerve has been reported with success by Crumley. In all reinnervation procedures, the duration of reinnervation process and aberrant reinnervation with synkinesis continue to be problems in the optimal restoration of function in patients with bilateral vocal fold paralysis.[24,25]

Patients with bilateral cricoarytenoid joint ankylosis owing to intubation or laryngeal fracture are tracheostomy dependent unless additional surgery can be done. The surgery for bilateral cricoarytenoid ankylosis varies from open surgical arytenoidectomy to arytenoidpexy to endoscopic arytenoidectomy or transverse cordectomy. All have their advantages and disadvantages.

Beside laryngeal stenosis, laryngeal incompetence may result in disabling aspiration or dysphonia. Glottic incompetence is most obvious after the laryngeal edema and hematoma have fully resolved. Unfortunately, this is often when a previously unrecognized fracture is found with misalignment of the fracture, resulting in glottal incompetence. The cause of late glottic incompetence may be attributable to loss of soft tissue, vocal fold paralysis, and laryngeal ankylosis as well as inadequate repair and reconstruction of the normal endolaryngeal volume or inadequate approximation of the laryngeal fragments. These problems may result in inadequate soft tissue needed for glottic closure during deglutition or phonation. Since spontaneous recovery of vocal cord function with injury of the recurrent laryngeal nerve is likely unless the nerve is severed, observation for up to 8 months is indicated. However, if dysphonia or aspiration is caused by laryngeal ankylosis or anatomic deficits, repair of laryngeal incompetence by laryngoplasty is indicated. Using modern phonosurgery techniques, the vocal folds can be medialized as a temporizing measure with Gelfoam or fat injection.[26] In situations in which the paralysis does not resolve, vocal cord medialization by laryngoplasty with or without arytenoid cartilage adduction may be performed.[27]

RADIATION INJURY

External beam radiation has been used successfully for the treatment of laryngeal carcinoma for many decades. Full-course therapeutic radiation always causes temporary laryngeal edema. In addition, radiation therapy can cause permanent morbidity of the larynx. The spectrum of injury ranges from edema to perichondritis and cartilage necrosis. The incidence of radiation-induced soft tissue injuries has decreased from 5% in 1970 to 1% in 1990.[28] However, owing to the recent trend in treating advanced-stage laryngeal carcinoma with organ-preservation protocols, more patients are receiving high-dose radiation and are therefore at risk for developing radiation-induced complications. Although most cases of radiation-induced chrondronecrosis occur within the first year after treatment, the changes in the vasculature are permanent, and necrosis has been reported to occur 50 years after treatment.[29]

Pathology

The first effects of radiation therapy are in the epithelium. There is loss of glandular secretions, and ciliary function is damaged. This leads to dry, irritated mucosa that is prone to infection. Radiation induces

fibrosis of the submucosal layer, thereby reducing venous outflow that, in turn, causes acute edema. With higher doses, the perichondrium degenerates, and a lymphocytic infiltration occurs. When the dose exceeds 1,000 to 1,200 cGy, irreversible vascular and lymphatic injury occurs. Subintimal fibrosis and proliferation of the endothelium cause obstruction of small arterioles. These changes in the endothelium are permanent and are the main cause of subsequent perichondritis and chondronecrosis.

Clinical Presentation and Management

Clinical symptoms and signs of radiation injury to the larynx include dysphonia, dysphagia, cough, pain, odynophagia, weight loss, aspiration, upper airway obstruction, malodorous breath, and sepsis. Chandler described a clinical grading system for radiation-induced changes in the larynx.[30] Grades I and II include moderate hoarseness, moderate dryness, moderate edema and erythema, and mild impairment of vocal fold mobility. These symptoms and signs are expected in all patients receiving radiation therapy. These patients can often be successfully treated with humidification, antireflux medication, and sialagogues.

Grade III consists of severe hoarseness, dyspnea, moderate dysphagia, and odynophagia. On examination, there is fixation of the vocal folds with marked edema and erythema. Treatment consists of humidification, antireflux medication, antibiotics, and corticosteroids. A 2- to 3-week course of this treatment often results in marked improvement in symptoms.

If there is no response or recurrent episodes, the patient is likely progressing to frank chondronecrosis. Grade IV involves respiratory distress, severe pain, dehydration, and fever. There is evidence of airway obstruction and fetid odor with possible skin necrosis. Treatment is the same as for grade III reactions except that tracheostomy is often needed for airway control and gastrostomy is often needed for adequate nutritional support.

Often patients with radiation-induced changes in the larynx become a diagnostic dilemma for the clinician. It is often impossible to distinguish between recurrent carcinoma and radiation reaction, particularly in the face of increasing edema and impairment of vocal fold mobility. If a patient fails to respond to conservative management (corticosteroids, antibiotics, antireflux medication, and humidification), direct laryngoscopy and biopsy are necessary. If tumor is present at biopsy, total laryngectomy is performed. If no tumor is demonstrated, conservative measures can be continued for several weeks, with frequent follow-up to detect recurrent carcinoma. In addition, hyperbaric oxygen has been used successfully in preventing the need for laryngectomy and tracheostomy in patients with grade IV laryngeal radionecrosis. If after a few months the larynx remains nonfunctional (tracheostomy dependence, severe pain, aphonia, and aspiration), total laryngectomy should be considered even if no tumor is present at biopsy.[31]

Many clinicians are reluctant to biopsy a radiated larynx unless their index of suspicion for recurrent tumor is very high. The already compromised laryngeal cartilage will suffer further ischemia and potential salivary contamination at biopsy, thus increasing the risk of irreversible necrosis. The problem arises in distinguishing who is at high risk for recurrent carcinoma. Computed tomographic scanning and magnetic resonance imaging have not consistently been useful in distinguishing between radiation changes and tumor. Positron emission tomographic scanning has shown some promise in this area by appropriately distinguishing tumor from radiation changes.[32]

LARYNGEAL BURNS

The most common type of thermal injury to the larynx occurs with inhalation of hot gases or smoke in burning buildings. Rarely are thermal injuries secondary to ingestion of hot liquids or foods. When ingestion of a hot liquid or solid is the cause, the burn is usually confined to the supraglottis and posterior part of the larynx. Often there are oral cavity burns as well. Owing to reflexive glottic closure, there is almost never subglottic injury. Isolated burns to the supraglottis can mimic acute infectious epiglottitis. If this occurs, the airway should be controlled by either tracheostomy or intubation in the operating room. The edema secondary to hot liquid or solid burns usually resolves rapidly, and decannulation or extubation can be accomplished in a short time. Rarely do hot liquids or solids cause long-term sequelae in the larynx.

In contrast, inhalation injuries frequently cause severe and potentially life-threatening acute and long-term complications. Both heat and irritant inhalation produce severe tracheobronchitis with

sloughing of mucosa. If the basal cell layer of the epithelium is intact, early repair with minimal complications is expected. If the basal cell layer is lost owing to the original insult or subsequent local trauma, delayed repair will result in granulation tissue formation and potential airway stenosis.[33]

Thermal inhalation injury most commonly occurs in patients with concomitant cutaneous burns. The incidence of inhalation burns increases as the extent of the cutaneous burn increases. The location of maximal injury within the airway depends on the size of the inhaled particles. Large particles mainly affect the trachea and larynx, whereas small particles mainly affect the alveoli. Thermal injury induces rapid edema that can cause airway obstruction.

Patients with inhalation injury have a wide variability of presenting symptoms and signs. These range from cough, to stridor, to respiratory arrest. Some authors have advocated elective intubation of all patients with suspected inhalation injury. Others argue that this leads to many unnecessary intubations, and expectant management of the airway can be safely performed for stable patients.[34] Most authors agree that if a patient presents with stridor, loss of consciousness, or massive burns, immediate intubation should be carried out.

If the patient is stable, the decision regarding intubation can be deferred until fiber-optic laryngoscopy and bronchoscopy have been performed to assess the airway. Some centers use flow-volume loops and ventilation-perfusion nuclear medicine scans to assess the damage to the lower airways and predict who requires intubation. If on fiber-optic laryngoscopy and bronchoscopy soot or mucosal injury is found, intubation should be performed. Patients who require intubation are at risk for late laryngeal and tracheal complications because their mucosas have already been damaged by the burns. No definitive improvement in outcome has been demonstrated with empiric treatment with antibiotics and corticosteroids in inhalation burns.

There is some controversy regarding the indications for tracheostomy in the patient with inhalation injury. Some argue that tracheostomy should be avoided to prevent bacterial contamination of the lower airway. Others state that early tracheostomy will "put the larynx at rest" and prevent possible laryngeal stenosis. Overall, the management of the airway in the patient with inhalation injury should be individualized based on the patient's clinical situation.

CONCLUSION

External laryngeal trauma is a rare injury in which morbidity is significantly decreased by early and accurate diagnosis and management. Stabilization and maintenance of the airway are paramount concerns. Tracheostomy is the preferred method of airway stabilization. If endotracheal intubation is to be performed, it should be done under direct visualization by an experienced clinician and with a tracheostomy kit at hand. The history, physical examination, and diagnostic imaging provide information regarding the severity of the injury. If displaced cartilaginous fractures and mucosal lacerations are present, the goal is realignment and fixation of the fractures and primary closure of all lacerations so that all exposed cartilaginous structures are covered. In some cases, stents and miniplates must be used to provide adequate stabilization of structures. With appropriate management, complications such as laryngeal stenosis can be minimized.

REFERENCES

1. Jewett BS, Schockley WW, Rutledge R. External laryngeal trauma analysis of 392 patients. Arch Otolayngol Head Neck Surg 1999;125:877–80.
2. Leopold DA. Laryngeal trauma. A historical comparison of treatment methods. Arch Otolaryngol Head Neck Surg 1983;109:106–12.
3. Biller HF, Moscoso J, Sanders I. Laryngeal trauma. In: Ballenger JJ, Snow JB, editors. Otorhinolaryngology head and neck surgery. Baltimore: Williams and Wilkins; 1996. p. 518–31.
4. Meyer CM, Orobello P, Cotton RT, Bratcher GO. Blunt laryngeal trauma in children. Laryngoscope 1987;97:1043–8.
5. Schaefer SD, Stringer SP. Laryngeal trauma. In: Bailey BJ, et al, editors. Head and neck surgery-otolaryngology. Philadelphia: JB Lippincott; 1998. p. 947–57.
6. Cherian TA, Rupa V, Raman R. External laryngeal trauma: analysis of thirty cases. J Laryngol Otol 1993; 107:920–3.
7. Bent JP, Silver JR, Porubsky ES. Acute laryngeal trauma: a review of 77 patients. Otolaryngol Head Neck Surg 1993;109:441–9.

8. Schaefer SD, Brown OE. Selective applicaton of CT in the management of laryngeal trauma. Laryngoscope 1983;93:1473–5.
9. Schaefer DS. Primary management of laryngeal trauma. Ann Otol Rhinol Laryngol 1982;91:399–402.
10. Olson NR, Miles WK. Treatment of acute blunt laryngeal injuries. Ann Otol Rhinol Laryngol 1971; 80:710–20.
11. Camnitz PS, Shepherd SM, Henderson RA. Acute blunt laryngeal and tracheal trauma. Am J Emerg Med 1986;5:157–62.
12. Schaefer SD. Acute laryngeal trauma. In: Gates GA, editor. Current therapy in otolaryngology-head and neck surgery. St Louis: Mosby; 1998. p. 456–8.
13. Koufman JA. The otolarynologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. Laryngoscope 1991;101 Suppl:1–78.
14. Harris HH, Tobin HA. Acute injuries of the larynx and trachea in 49 patients (observations over a 15-year period). Laryngoscope 1970;80:1376–84.
15. Woo P. Laryngeal framework reconstruction with miniplates. Ann Otol Rhinol Laryngol 1990;99(10 Pt 1):772–7.
16. Schaefer SD, Close LG. Acute management of laryngeal trauma. Update. Ann Otol Rhinol Laryngol 1989;98:98–104.
17. Schaefer SD. Laryngeal and esophageal trauma. In: Cummings CW, Frederickson JM, Harker LA, et al, editors. Otolaryngology head and neck surgery. St Louis: Mosby; 1998. p. 2001–22.
18. Peskind SP, Stanley RB, Thangathurai D. Treatment of the compromised trachea with sleeve resection and primary repair. Laryngoscope 1993;103:203–11.
19. Montgomery WW, Montgomery SK. Manual for use of Montgomery laryngeal, tracheal and esophageal prostheses. Ann Otol Rhinol Laryngol 1986;95 Suppl 125:1–16.
20. Schuller DS, Parrish R. Reconstruction of the larynx and trachea. Arch Otolaryngol Head Neck Surg 1988;114:278–86.
21. Toohill RJ. Autogenous graft reconstruction of the larynx and upper trachea. Otolaryngol Clin North Am 1979;12:909–17.
22. Friedman WH, Biller HF, Som ML. Repair of extended laryngotracheal stenosis. Arch Otolaryngol Head Neck Surg 1975;101:151–6.
23. Alonso WA, Druck NS, Ogura JH. Clinical experiences in hyoid arch transposition. Laryngoscope 1976;86:617–24.
24. Crumley RL. Teflon versus thyroplasty versus nerve transfer: a comparison. Ann Otol Rhinol Laryngol 1990;99:759–63.
25. Crumley RL. Laryngeal synkinesis revisited. Ann Otol Rhinol Laryngol 2000;109:365–71.
26. Hsiung MW, Woo P, Minasian A, Mojica JS. Fat augmentation for glottic insufficiency. Laryngoscope 2000;110:1026–33.
27. Netterville JL, Stone R, Luken ES, et al. Silastic medialization and arytenoid adduction: the Vanderbilt experience. A review of 116 phonosurgical procedures. Ann Otol Rhinol Laryngol 1993;102: 413–24.
28. Oppenheimer RW, Krespi YP, Einhorn RK. Management of laryngeal radionecrosis: animal and clinical experience. Head Neck 1989;11:252–6.
29. Fitgerald PJ, Koch RJ. Delayed radionecrosis of the larynx. Am J Otolaryngol 1999;20;245–9.
30. Chandler JR. Radiation fibrosis and necrosis of the larynx. Ann Otol Rhinol Laryngol 1979;88:509–14.
31. McGuirt WF. Laryngeal radionecrosis versus recurrent cancer. Otolaryngol Clin North Am 1997;30: 243–50.
32. Greven KM, Williams DW, Keyes JW, et al. Distinguishing tumor recurrence from irradiation sequelae with positron emission tomography in patients treated for larynx cancer. Int J Radiat Oncol Biol Phys 1994;29:841–5.
33. Gaissert HA, Lofgren RH, Grillo HC. Upper airway compromise after inhalation injury. Ann Surg 1993;218:672–8.
34. Pruitt BA, Cioffi WG. Diagnosis and treatment of smoke inhalation. J Intens Care Med 1995;10: 117–27.

CHAPTER 52

Infectious and Inflammatory Diseases of the Larynx

James A. Koufman, MD, Peter C. Belafsky, MD, PhD

The majority of patients with disorders of the larynx and voice suffer from infectious and noninfectious inflammatory conditions. It is important to remember that the term *inflammation* implies a local response to tissue injury, characterized by capillary dilation and leukocyte infiltration. The typical symptoms and signs of inflammation are swelling, redness, and, sometimes, discomfort or pain. The term *laryngitis* is synonymous with laryngeal inflammation, although not with hoarseness. Laryngitis (laryngeal inflammation) may result from infection by an invading microorganism or from irritative, traumatic, metabolic, allergic, autoimmune, or idiopathic causes.

Acute and chronic laryngitis are very common in otolaryngologic practice, and the causes (differential diagnoses) in pediatric and in adult patients are different. In infants and children, for example, the most common cause of laryngitis is acute infection, whereas in adults, laryngitis generally tends to have a chronic, noninfectious cause.

Within the last 20 years, laryngopharyngeal (gastroesophageal) reflux has been discovered to be a far more important cause of laryngeal inflammation than was previously recognized. In addition, primarily because of acquired immune deficiency syndrome (AIDS), infecting microorganisms that were rarely encountered in the practice of otolaryngology just a few years ago are re-emerging. Thus, the otolaryngologist today must again become familiar with the clinical manifestations of infection by a wide spectrum of microorganisms.

Laryngeal inflammatory disorders are unusual in that often more than one causative factor or condition can be identified. For example, patients with Reinke's edema are frequently smokers who also misuse their voices and have reflux. Such patients may also develop laryngeal carcinoma.

Each of the underlying causes must be identified and corrected if treatment is to be successful. As more has been learned about the larynx, environmental influences, and the effects of systemic disorders on the larynx, imprecise diagnostic terms, such as *nonspecific laryngitis*, have appropriately begun to disappear from the otolaryngologic literature.

The differential diagnosis of the (infectious and noninfectious) causes of laryngitis is shown in Table 52–1.

LARYNGOPHARYNGEAL (GASTROESOPHAGEAL) REFLUX DISEASE

It has been estimated that 10% of Americans have heartburn on a daily basis, and an additional 30 to 50% have it less frequently.[1] Of all of the causes of laryngeal inflammation, gastroesophageal reflux disease (GER, GERD) is the most common cause, and as many as 10 to 50% of patients with laryngeal complaints have a GER-related underlying cause.[2,3]

The term *reflux* literally means *back flow*. Reflux of stomach contents into the esophagus is common, and many patients with GERD have symptoms such as heartburn and regurgitation related to inflammation of the esophagus by acid and digestive enzymes. When refluxed material escapes the esophagus and enters the laryngopharynx above, the event is termed *laryngopharyngeal reflux* (LPR). Although the terms *gastroesophageal reflux* and *laryngopharyngeal reflux* are often used interchangeably, the latter is more specific. Laryngopharyngeal reflux is the preferred term for use in otolaryngology because the patterns, mechanisms, and manifestations of LPR differ from classic GERD.

Laryngopharyngeal reflux affects both children and adults and may be associated with an acute,

TABLE 52–1. Inflammatory Disorders of the Larynx

I. Gastroesophageal (laryngopharyngeal) reflux disease
II. Pediatric laryngitis
  A. Acute (viral or bacterial) infections
    1. Laryngotracheitis (croup)
    2. Supraglottitis (epiglottitis)
    3. Diphtheria
  B. Noninfectious causes
    1. Spasmodic croup
    2. Traumatic laryngitis
III. Acute laryngeal infections of adults
  A. Viral laryngitis
    1. Common upper respiratory infection
    2. Laryngotracheitis
    3. Herpes simplex
  B. Bacterial laryngitis
    1. Supraglottitis
    2. Laryngeal abscess
    3. Gonorrhea
IV. Chronic (granulomatous) diseases
  A. Bacterial
    1. Tuberculosis
    2. Leprosy
    3. Scleroma
    4. Actinomycosis
    5. Tularemia
    6. Glanders
    7. Syphilis
  B. Mycotic (fungal)
    1. Candidiasis
    2. Blastomycosis
    3. Histoplasmosis
    4. Coccidiomycosis
    5. Aspergillosis
    6. Sporotrichosis
  C. Idiopathic
    1. Sarcoidosis
    2. Wegener's granulomatosis
V. Allergic, immune, and idiopathic disorders
  A. Hypersensitivity reactions
    1. Angioedema
    2. Stevens-Johnson syndrome
  B. Immune and idiopathic disorders
    1. Infections of the immunocompromised host
    2. Rheumatoid arthritis
    3. Systemic lupus erythematosus
    4. Cicatricial pemphigoid
    5. Relapsing polychondritis
    6. Sjögren's syndrome
    7. Amyloidosis
VI. Miscellaneous inflammatory conditions
  A. Parasitic infections
    1. Trichinosis
    2. Leishmaniasis
    3. Schistosomiasis
    4. Syngamosis
  B. Inhalation laryngitis
    1. Acute (thermal) injury
    2. Pollution and inhalant allergy
    3. Carcinogens
  C. Radiation injury
    1. Radiation laryngitis
    2. Radionecrosis
  D. Vocal abuse and misuse syndromes
    1. Vocal fold hemorrhage
    2. Muscle tension dysphonias
    3. Contact ulcer and granuloma

chronic, or intermittent pattern of laryngitis, with or without granuloma formation. Indeed, LPR has also been implicated in the development of laryngeal carcinoma and stenosis, recurrent laryngospasm, and cricoarytenoid joint fixation, as well as with many other otolaryngology-related conditions, including globus pharyngeus, cervical dysphagia, and subglottic stenosis.[3–10]

The symptoms of LPR are quite different from those of classic GERD as seen in the gastroenterology patient, who characteristically has heartburn, regurgitation, and esophagitis. Patients with "reflux laryn-

gitis" (LPR) present with hoarseness, but almost two-thirds deny ever having heartburn.[3,6] Other throat symptoms, such as globus pharyngeus (a sensation of a lump in the throat), dysphagia, chronic throat clearing, and cough, are often associated with LPR.[6] Gastroenterologists call reflux patients who deny gastrointestinal symptoms *atypical refluxers*, but these patients are quite typical of those encountered in otolaryngologic practice.[11] The symptoms and laryngeal conditions that have been reported to be associated with LPR are summarized in Table 52–2, and the differences between the LPR patient and the typical GERD patient are summarized in Table 52–3. We have developed and validated a nine-item reflux symptom index (RSI) to quantify patient symptoms of LPR and evaluate treatment efficacy (Table 52–4). This outcome instrument has displayed excellent reproducibility and criterion-based validity.[12]

Laryngeal Findings

Physical findings of LPR can range from mild, isolated edema and/or erythema of the area of the arytenoid cartilages to diffuse laryngeal edema and hyperemia with granuloma formation and airway obstruction. We have validated an eight-item reflux finding score (RFS) to document the severity of the clinical findings of LPR (Table 52–5).[12,13] Use of the RFS not only helps physicians identify subtle findings of reflux, it also assists in evaluating the severity of laryngeal tissue injury, as well as documenting treatment efficacy. Figure 52–1 displays the physical findings of LPR, which are discussed in detail below.

Pseudosulcus vocalis refers to a pattern of subglottic edema that extends from the anterior commissure to the posterior part of the larynx; it appears like a groove or sulcus. It can easily be differentiated from a true sulcus (*sulcus vergeture*), which is the adherence of the vocal fold epithelium to the vocal ligament secondary to the absence of the superficial layer of lamina propria. True sulcus is related to scarring of the vocal fold(s) in the phonatory striking zone. Whereas true sulcus stops at the vocal process and is in the midportion of the striking zone, pseudosulcus vocalis extends all the way to

TABLE 52–2. Symptoms and Laryngeal Conditions Reported to be Associated with Laryngopharyngeal Reflux

Symptoms	*Conditions*
Chronic dysphonia	Reflux laryngitis
Intermittent dysphonia	Subglottic stenosis
Vocal fatigue	Carcinoma of the larynx
Voice breaks	Endotracheal intubation injury
Chronic throat clearing	Contact ulcers and granulomas
Excessive throat mucus	Posterior glottic stenosis
"Postnasal drip"	Arytenoid cartilage fixation
Chronic cough	Paroxysmal laryngospasm
Dysphagia	Paradoxical vocal fold movement
Globus	Globus pharyngeus
Intermittent airway obstruction	Vocal nodules
Chronic airway obstruction	Polypoid degeneration
	Laryngomalacia
	Pachydermia laryngis
	Recurrent leukoplakia
	Sudden infant death syndrome

TABLE 52–3. Differences between the Typical Gastroenterology Patient with GERD and the Otolaryngology Patient with LPR

	GERD	*LPR*
Symptoms		
Heartburn and/or regurgitation	+++++	
Hoarseness, dysphagia, globus, etc	+	++++
Findings		
Endoscopic esophagitis	+++++	
Laryngeal inflammation	+	++++
Diagnostic yield (abnormality)		
Esophageal biopsy (inflammation)	+++++	
Abnormal esophageal radiography	++	—
Abnormal esophageal pH monitoring	+++++	
Abnormal pharyngeal pH monitoring	—	+++
Pattern of reflux		
Supine (nocturnal)	+++++	
Upright (daytime)	+	++++
Response to treatment		
Dietary/lifestyle modification	++	+
$Histamine_2$ antagonists*	+++	++
Proton pump inhibitors*	+++++++	

*Assuming adequate dosage and duration of therapy.
GERD = gastroesophageal reflux disease; LPR = laryngopharyngeal reflux.

the back of the larynx. Its presence contributes 2 points to the RFS.

Ventricular obliteration is a frequent finding that may be identified in up to 80% of patients with LPR. Swelling of the true and false vocal folds causes this space to become obliterated and thus poorly visualized. With partial ventricular obliteration, the space is reduced, and the false vocal fold edge is indistinct. With complete ventricular obliteration, the true and false folds appear to touch, and there is no true ventricular space. Partial obliteration contributes 2 points and complete obliteration contributes 4 points to the RFS.

Laryngeal erythema/hyperemia is a nonspecific finding that is dependent on the available endoscopic equipment. Subtle changes in erythema are difficult to quantify and vary depending on the quality of the fiberscope, video monitor, and light source. Nonetheless, isolated erythema of the area of the arytenoid cartilages contributes 2 points and diffuse laryngeal erythema contributes 4 points to the RFS.

True vocal fold edema is graded as mild (1 point) if only slight swelling exists and moderate (2 points) when it becomes more perceptible. Edema is graded as severe (3 points) when swelling of the cord becomes sessile. Finally, polypoid degeneration of the true vocal fold contributes 4 points to the RFS.

Diffuse laryngeal edema is judged by the size of the airway relative to the size of the larynx. It is graded as mild (1 point) to obstructing (4 points). Hypertrophy of the posterior commissure is another frequent finding of LPR. It is graded as mild (1 point) when there is a mustache-like appearance of the posterior commissure mucosa and moderate (2 points) when the posterior commissure mucosa is swollen

TABLE 52–4. Reflux Symptom Index

Within the last month, how did the following problems affect you?	*Score**					
1. Hoarseness or a problem with your voice	0	1	2	3	4	5
2. Clearing your throat	0	1	2	3	4	5
3. Excess throat mucus or postnasal drip	0	1	2	3	4	5
4. Difficulty swallowing food, liquids, or pills	0	1	2	3	4	5
5. Coughing after you ate or after lying down	0	1	2	3	4	5
6. Breathing difficulties or choking episodes	0	1	2	3	4	5
7. Troublesome or annoying cough	0	1	2	3	4	5
8. Sensations of something sticking in your throat or a lump in your throat	0	1	2	3	4	5
9. Heartburn, chest pain, indigestion, or stomach acid coming up	0	1	2	3	4	5
			Total			

*0 = no problem; 5 = severe problem.

enough to create a straight line across the back of the larynx. Posterior commissure hypertrophy (PCH) is graded as severe (3 points) when there is bulging of the posterior part of the larynx into the airway and obstructing (4 points) when a significant portion of the airway is obliterated. The final two items on the RFS are granuloma/granulation and thick endolaryngeal mucus. Patients get 2 points when these entities are present and 0 points otherwise.

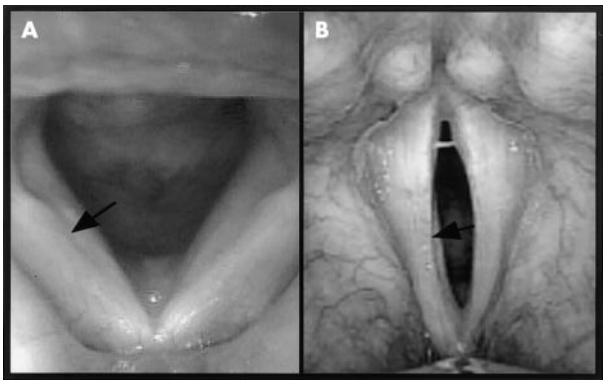
Diagnosis

Ambulatory 24-hour double-probe (simultaneous esophageal and pharyngeal) pH monitoring (pH-metry) is the current gold standard for the diagnosis of LPR.[3,14,15] The distal probe is placed 5 cm above the lower esophageal sphincter (LES), and the proximal probe is placed in the hypopharynx 1 cm above the upper esophageal sphincter (UES), just behind the laryngeal inlet (Figure 52–2). The traditional technique of probe placement is to place both the proximal and distal pH probes under manometric guidance. A manometer is inserted through the nasal cavity and advanced through the esophagus into the stomach. It is then slowly withdrawn, and the locations of the LES and UES are determined. The correct pH catheter is then chosen based on these measurements. An alternative technique involves placing the proximal probe just above the UES under direct fiber-optic visualization. The distance between the proximal and distal

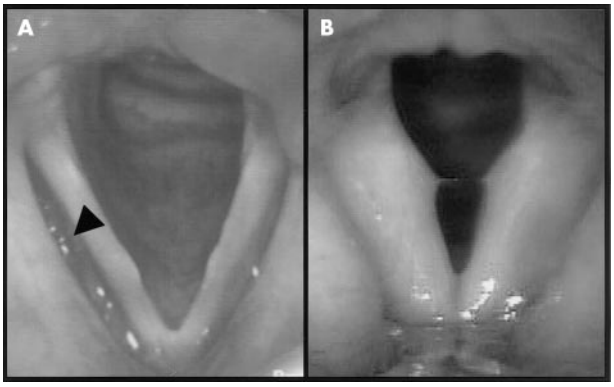
TABLE 52–5. Reflux Finding Score (RFS)

Subglottic edema	2 = present
	0 = absent
Ventricular obliteration	2 = partial
	4 = complete
Erythema/hyperemia	2 = arytenoids only
	4 = diffuse
Vocal fold edema	1 = mild
	2 = moderate
	3 = severe
	4 = polypoid
Diffuse laryngeal edema	1 = mild
	2 = moderate
	3 = severe
	4 = obstructing
Posterior commissure hypertrophy	1 = mild
	2 = moderate
	3 = severe
	4 = obstructing
Granuloma/granulation	2 = present
	0 = absent
Thick endolaryngeal mucus	2 = present
	0 = absent
	Total

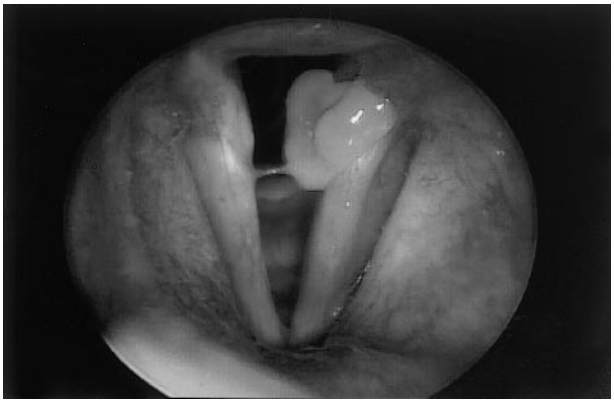
I. Subglottic Edema (Pseudosulcus Vocalis)



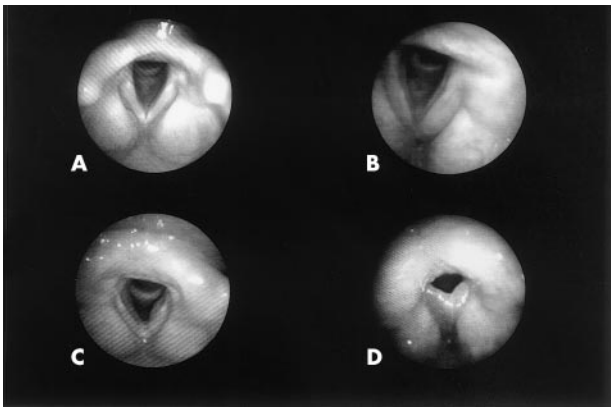
II. Ventricular Obliteration



III. Granuloma/Granulation Tissue



IV. Vocal Fold Edema



V. Posterior Commissure Hypertrophy

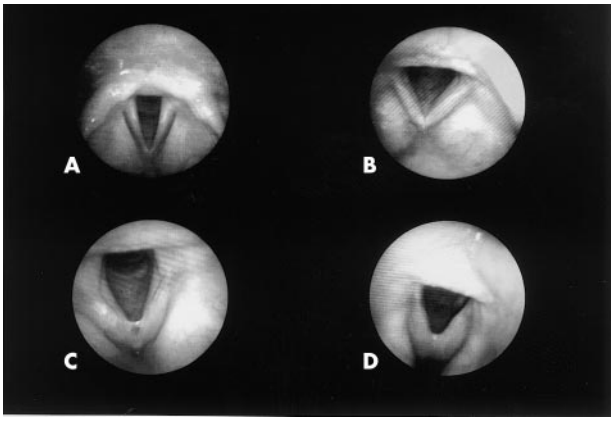


FIGURE 52–1. Findings of laryngopharyngeal reflux. I: *A*, Bilateral pseudosulcus vocalis. Notice that the subglottic edema extends past the vocal process (*arrow*). Also present are mild posterior commissure hypertrophy and vocal fold edema. *B*, True sulcus vocalis (adhesion of overlying vocal fold epithelium to underlying vocal ligament) (*arrow*). II: *A*, Open laryngeal ventricles (*arrowhead*). Also present are small prenodules. *B*, Near-complete ventricular obliteration. Also present are mild posterior commissure hypertrophy and moderate vocal fold edema. III: Granuloma of left vocal process. IV: *A*, Mild vocal fold edema. *B*, Moderate vocal fold edema. Pseudosulcus, mild posterior commissure hypertrophy, and partial ventricular obliteration are also present. *C*, Severe vocal fold edema. Sessile changes are noted. *D*, Polypoid degeneration of the vocal folds. V: *A*, Normal posterior commissure. Cuneiform cartilages can be visualized. *B*, Mild posterior commissure hypertrophy. Mustache-like configuration of posterior larynx. Cuneiform no longer visualized. *C*, Moderate posterior commissure hypertrophy. Straight bar across posterior larynx. *D*, Severe posterior commissure hypertrophy. Protrusion of posterior commissure into airway.

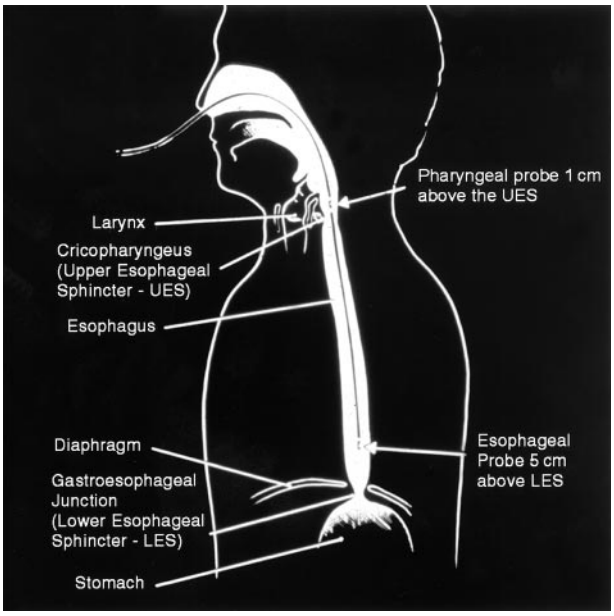


FIGURE 52–2. Technique of ambulatory 24-hour double-probe pH monitoring. The distal probe is placed 5 cm above the lower esophageal sphincter (LES), and the proximal (pharyngeal) probe is placed 1 cm above the upper esophageal sphincter (UES).

pH probes is fixed at 15 cm. This technique is easier, less time consuming, and less costly than using manometric guidance. Placement of the proximal probe above the UES can be performed accurately using this method. This technique, however, is unable to estimate precisely interprobe distances.[16] Thus, the exact location of the distal probe is uncertain, and the esophageal data are often grossly inaccurate using this technique.

pH-metry has been available for many years, and standards (normal values) have been established in many laboratories.[2,17] In general, the most important parameter used to evaluate the presence of GERD at the distal probe is the percentage of time that the pH is less than 4. This measurement is usually recorded for time in the *upright position*, time in the *supine position*, and the *total time* of the study. For the upright period, the upper limit of normal is approximately 8.0%, and for the supine period, approximately 2.5%.[17]

The proximal pharyngeal probe is invaluable in patients with LPR because it is placed behind the larynx just above the cricopharyngeus; thus, reflux recorded by it is diagnostic of LPR (Figure 52–3). In addition, it has been found that without the proximal (pharyngeal) probe, the diagnosis of LPR will be missed in approximately 30 to 50% of patients.[18]

It is also advisable to evaluate the esophagus of patients with LPR. This may be accomplished by performing a barium esophagram[19] or, more recently, transnasal esophagoscopy. Although the prevalence of esophagitis in patients with LPR is only 20%, the percentage of other esophageal

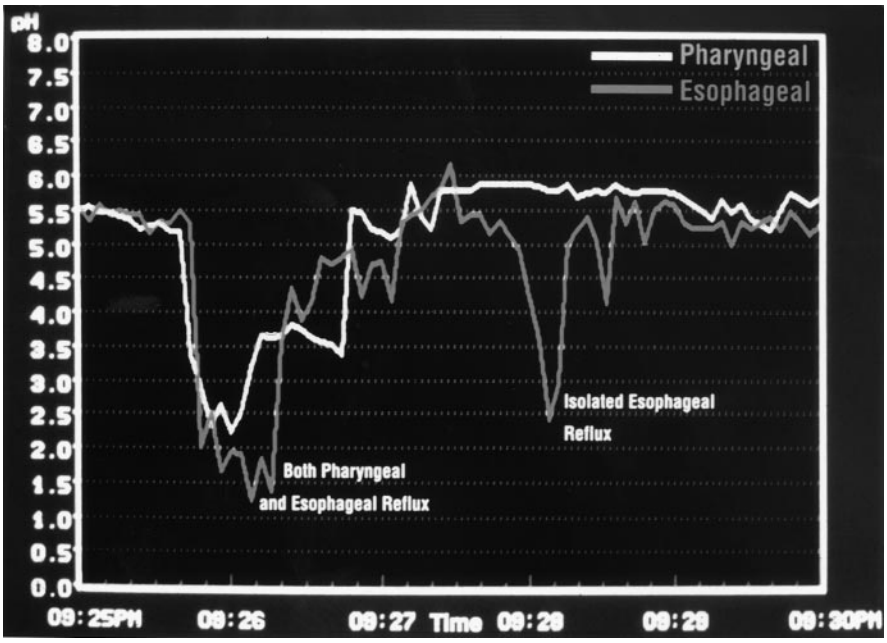


FIGURE 52–3. Example of an abnormal double-probe pH study. The *darker tracing* is the esophageal pH, and the lighter tracing is the pharyngeal pH. A combined esophageal and pharyngeal reflux event (pH < 4) is demonstrated at 9:26. An isolated esophageal reflux event is demonstrated at 9:28.

abnormalities such as Barrett's metaplasia (7%) may be high (JA Koufman and PC Belafsky, unpublished data).

Treatment

There are three levels of antireflux treatment: *level I*—dietary and lifestyle modification plus antacids, *level II*—level I plus use of a histamine H_2-receptor antagonist (such as cimetidine, ranitidine, or famotidine), and *level III*—antireflux surgery (eg, fundoplication) or proton pump inhibitor (PPI) therapy (eg, omeprazole, esomeprazole, lansoprazole, pantoprazole, or rabeprazole). The details of each of the three levels are listed in Table 52–6.

Clinical experience with LPR suggests that treatment must be individualized, with the level of treatment depending on the severity of the patient's condition. For many patients with LPR, level I or level II treatment is appropriate initial therapy, but both forms will fail in up to 35% of patients.[2,15,20] If level II treatment fails, level III treatment with a PPI is indicated. In patients who present with severe LPR or complications of LPR (laryngospasm, obstructing

TABLE 52–6. Treatment ("Levels") for Reflux Laryngitis

Level I—Antireflux therapy (ART)

- A. Dietary modification
  1. No eating or drinking within 3 hours of bedtime
  2. Avoid overeating or reclining right after meals
  3. No fried food; low-fat diet
  4. Avoid coffee, tea, chocolate, mints, and soda pop
  5. Avoid other caffeine-containing foods and beverages
  6. Avoid alcohol, especially in the evening
  7. Avoid other foods that cause you problems
- B. Lifestyle modification
  1. Elevate the head of the bed 4 to 6 inches
  2. Avoid wearing tight-fitting clothing or belts
  3. If you use tobacco, quit!
- C. Liquid antacids qid (1 tablespoon 1 hour after each meal and at bedtime)

Level II—Medication plus ART

- A. Level I (above), plus B or C (below)
- B. Initial treatment
  1. Histamine H_2-receptor antagonists bid
- C. Escalation for treatment failures
  1. Increase histamine H_2-receptor antagonist dose up to double-dose qid or
  2. Use a proton pump inhibitor (PPI)

Level III—Proton pump inhibitors or antireflux surgery

- A. Level I (above) minus antacids, plus B or C (below)
- B. Proton pump inhibitor bid (first dose in the morning, second at 5 pm; the duration of initial treatment should be 6 months; large patients may require larger doses. Patient symptoms should improve by 2 months.)
- C. Antireflux surgery (fundoplication)

granuloma, subglottic stenosis, etc), initial treatment with twice-daily PPI is indicated.

At present, the PPIs are the most effective antireflux medicine available because, unlike the H_2-antagonists, they can achieve total acid suppression.[20,21] Unlike treatment for GERD that usually responds to once-a-day dosing, complete resolution of tissue injury to the larynx may require higher doses of medication. Our approach to the patient with severe or complicated LPR is to begin therapy with twice-daily PPI. Research suggests that patient symptoms will achieve maximal relief by 2 months. Laryngeal reflux findings resolve more slowly, however, and may take 6 months or longer to reverse.[12] If patients are still symptomatic at 2 months, or if the laryngeal findings have not significantly improved by 6 months, repeat pH testing is indicated to ensure reasonably complete acid suppression (efficacy study). The failure of adequate acid suppression on twice-daily PPIs is significant, and a low threshold for obtaining an efficacy study on medication should be maintained.[22,23] In some patients who face a lifetime of antireflux treatment, or in patients who fail medical therapy, referral for a fundoplication is warranted.

Potentially Life-Threatening Manifestations of Laryngopharyngeal Reflux

Paroxysmal Laryngospasm Laryngospasm is an uncommon complaint, but patients who experience it usually are able to describe the event in vivid detail. If the clinician mimics severe inspiratory stridor, the patient will confirm that his or her breathing pattern during attacks was indeed similar. Laryngospasm is often paroxysmal and usually occurs without warning. For some patients, the attacks awake them from sleep; for others, the attacks occur during the day. The attacks may have a predictable pattern (eg, occurring postcibal or during exercise), and some patients are aware of a relationship between reflux and the attacks (others are not). Loughlin and Koufman have studied 12 patients with recurrent paroxysmal laryngospasm, 11 of whom had documented LPR by pH-metry and all of whom responded to treatment with PPIs by a cessation of laryngospastic episodes.[24]

Laryngeal Stenosis Excluding trauma, LPR is the primary cause of laryngeal stenosis, including subglottic stenosis and posterior laryngeal stenosis;[2,4,5,10] pH-metry-documented LPR has been found in 92% of patients.[2] Treatment with PPIs or fundoplication will result in subsequent decannulation of the vast majority of patients with such stenosis, in some cases without surgery.[2,10]

The traditional dichotomy between mature and immature stenoses probably represents an oversimplification. *Immature* implies that massive edema and granulation tissue are present and that the inflammatory process is ongoing; *mature* implies that acute inflammation has resolved and that the stenosis is composed of mature fibrotic tissue with thin (normal) overlying epithelium. Surgical attempts to correct immature stenoses usually fail unless the LPR is controlled. Conversely, mature stenoses usually can be corrected surgically. In reality, however, many cases are neither mature nor immature but somewhere in between. The same may also be true of many acquired laryngeal webs.

pH-metry and aggressive antireflux therapy should be used in all patients with laryngeal stenosis. As the database grows, this admonition will probably become dogma. Meanwhile, treatment with PPIs or fundoplication should be considered adjuncts to surgical therapy.

Laryngeal Carcinoma The risk factors for the development of laryngeal carcinoma include LPR.[7,8] Koufman reported a series of 31 consecutive patients with laryngeal carcinoma in whom abnormal reflux was documented in 84%, but only 58% were active smokers.[2] The relationship between LPR and malignant degeneration remains unproven, but the available pH-metry data suggest that most patients who develop laryngeal malignancy both smoke and have LPR. In addition, apparently premalignant lesions may resolve with appropriate antireflux therapy.[2]

The previously identified risk factors for laryngeal carcinoma, tobacco and ethanol, also strongly predispose one to reflux.[2] Tobacco and alcohol adversely influence almost all of the body's antireflux mechanisms; they delay gastric emptying, decrease LES pressure and esophageal motility, decrease mucosal resistance, and increase secretion of gastric acid. Routine pH-metry, followed by PPI treatment, is recommended for all patients with laryngeal neo-

plasia, with or without other risk factors for the development of carcinoma of the larynx.

Laryngopharyngeal Reflux Disease in Infants and Children

Laryngopharyngeal reflux is ubiquitous and pernicious in pediatric otolaryngology patients. The diagnosis may be particularly difficult to make because infants and children almost never complain of heartburn or other reflux symptoms. Laryngopharyngeal reflux has been clearly established to be associated with the development of childhood laryngeal and tracheal stenosis, as well as with reactive airway disorders (laryngospasm, asthma, sudden infant death syndrome, laryngomalacia, bronchopulmonary dysplasia, and aspiration pneumonia).[9,10,25,26]

A high index of suspicion is necessary for the diagnosis of LPR to be considered; currently, only prolonged pH-metry is diagnostic. Other tests, such as barium esophagography, radionuclide scanning, and the lipid-laden macrophage test, lack sufficient sensitivity and/or specificity to be of value in the majority of cases.

PEDIATRIC LARYNGITIS

Pediatric patients with laryngeal inflammation and edema present with one or more of the following symptoms: dysphonia, odynophonia, cough, dysphagia, odynophagia, stridor, and dyspnea. Airway obstruction from inflammatory laryngeal edema is more common in children than in adults owing to the small size of the pediatric larynx. Equivalent amounts of mucosal swelling may result in critical narrowing and obstruction in a child, while causing only minimal symptoms in an adult. Table 52–7 shows the effects of 1 mm of edema on the cross-sectional (subglottic) area of a small neonate, an average child, and an adult male. The magnitude of the difference between the effect in each explains why laryngeal inflammation is more often a life-threatening illness requiring airway management in infants and children than it is in adults.

TABLE 52–7. Effect of 1 mm of Edema on the Cross-sectional Area of the Subglottic Larynx in the Neonate, Child, and Adult (Area = πr^{2*})

	Neonate	*Child*	*Adult*
Normal			
Subglottic diameter (mm)	4	8	14
Subglottic radius (mm)	2	4	7
Subglottic area (mm^2)	12	48	147
Effect of 1 mm of edema			
Subglottic diameter (mm)	2	6	12
Subglottic radius (mm)	1	3	6
Subglottic area (mm^2)	3	27	108
Percent reduction of airway area	75	44	27

*For the sake of simplicity, for these calculations $\pi = 3$.

Laryngotracheitis (Croup)

Viral laryngotracheitis is the most common laryngeal inflammatory disorder of childhood. It is responsible for 15% of respiratory disease seen in pediatric practice. Usually, this condition is self-limited, occurs in children under the age of 3 years, and has a seasonal peak, with most cases occurring during the winter. Typically, for several days prior to presentation, the child's history is compatible with that of a viral upper respiratory infection with rhinitis, cough, and low-grade fever. When symptoms of hoarseness, dyspnea, stridor, and a barking cough develop, the diagnosis of laryngotracheitis is made. The characteristic cough gives laryngotracheitis its common name, croup. Parainfluenza viruses (types 1, 2, and 3) account for more than half of croup cases. Other viruses frequently implicated in the disorder include rhinovirus, respiratory syncytial virus, and adenovirus. Less common causes of laryngotracheitis are influenza, measles, mumps, pertussis, and chickenpox.[26]

When airway obstruction is caused by laryngotracheitis, the stridor is characteristically inspiratory, or biphasic. Although the diagnosis of laryngotracheitis is generally based on the history, examination of the larynx, although not necessary, shows erythematous and edematous mucosa with normal vocal fold mobility. Radiographs, which reveal a narrowing of the subglottic lumen, the "steeple sign," may be used to differentiate this condition from supraglottitis.

The need for inpatient hospitalization depends on the degree of airway obstruction. Treatment is

aimed at decreasing laryngeal edema and preventing stasis and crusting of secretions within the airway. Therapy usually includes hydration, humidification of inspired air, and treatments with nebulized racemic epinephrine. Antipyretics, decongestants, and parenteral corticosteroids are often administered. Although the value of corticosteroids remains controversial, many experienced clinicians use them because they believe that the anti-inflammatory effect can obviate the need for intubation when the airway is marginally adequate. Artificial airway support (eg, intubation) is necessary in a relatively small proportion of patients with laryngotracheitis. When needed, however, intubation should be carried out by experienced personnel, preferably in the operating room, where maximum airway control can be achieved.

Endoscopy with cultures of the airway surfaces and secretions should be considered for patients with atypical or severe laryngotracheitis, particularly those who cannot be extubated successfully. Secondary bacterial infection of the airway (membranous croup) is more serious and is usually suspected when the patient experiences high temperature spikes and exudative, purulent drainage. Bacterial tracheitis may ensue. Radiographically, the lumen of the upper airway will appear narrowed, shaggy, and irregular. The organisms most commonly involved are *Haemophilus influenzae*, *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Moraxella catarrhalis*, and hemolytic streptococci. Antibiotic therapy is indicated and is directed at the causative microorganisms.

Supraglottitis (Epiglottitis)

Acute supraglottitis is a life-threatening infection of the supraglottic larynx traditionally caused by *H. influenzae* type B. Since childhood immunization against type B *H. influenzae* has become commonplace, however, the disease is significantly less prevalent, and epidemiologic data suggest that non–type B *H. influenzae* is more frequent among those vaccinated children who are affected.[27] Other organisms causing supraglottitis include *Streptococcus pyogenes, S. pneumoniae*, and *S. aureus*. The infection is a true medical emergency. Children aged 2 to 4 years are the most frequently affected group, and cases are more frequent in the winter and spring months. The illness begins rapidly over 2 to 6 hours with the onset of fever, sore throat, and inspiratory stridor. The voice tends to be muffled, and there is no barking cough, as in croup. As the supraglottic structures become more edematous, airway obstruction develops.

The child is generally ill-appearing, stridulous, sitting upright, and drooling because swallowing is painful. The diagnosis is usually based purely on the history and clinical findings. Examination of the epiglottis (in the emergency room) may precipitate airway obstruction and thus is not recommended. Lateral soft tissue radiographs may reveal the classic "thumb sign" of the edematous epiglottis with a dilated hypopharynx. Occasionally, the epiglottis itself is not enlarged, but the supraglottic region still appears hazy and indistinct owing to edema of other supraglottic structures. In severe cases, treatment should not be delayed to obtain radiographs. If radiographs are deemed necessary, the study should be carried out in the presence of personnel capable of immediately intubating the patient should airway obstruction occur.

The child with suspected epiglottitis should be taken to the operating room immediately to establish the diagnosis and secure an airway. The child is transported to the operating room by the parents, an otolaryngologist, and an anesthesiologist. Direct laryngoscopy usually will show the epiglottis to be very swollen and cherry red, as are the aryepiglottic folds and the false vocal folds. The true vocal folds and subglottis typically appear to be normal or to be only minimally involved.

Treatment is directed at airway maintenance and then toward providing antimicrobial and supportive care. Drawing blood, starting intravenous lines, obtaining a rectal temperature, or otherwise disturbing the patient should be postponed until the airway is secured.

After the child is anesthetized and the diagnosis is made, the child should be orally intubated to secure the airway. The oral endotracheal tube can then be converted to a nasotracheal tube, with direct visualization of the glottis as the orotracheal tube is removed. Instruments necessary for rigid bronchoscopy and tracheostomy should be available and ready in the operating room in case the airway is lost before intubation.[28] In institutions without a highly skilled pediatric intensive care staff, a tracheostomy may be preferable to endotracheal intubation. It is easier to secure and to re-establish the airway if a tracheostomy tube is in place.[29]

Cultures of the epiglottis and blood are obtained after the airway is secured. Empirically,

antimicrobial therapy is initiated against *H. influenzae* and group A streptococcus and classically consists of a second- or third-generation cephalosporin (cefuroxime, cefotaxime, or ceftriaxone). Ampicillin/sulbactam and trimethoprim/sulfamethoxazole are alternatives. Chloramphenicol is highly effective but should not be used because of the availability of less toxic alternative agents. Extubation is usually possible after 48 to 72 hours, at which time the edema has subsided sufficiently to allow an air leak around the endotracheal tube. Transnasal fiberoptic laryngoscopy is the most reliable and most commonly employed technique to ensure resolution of the edema before extubation. Differentiating acute epiglottitis from laryngotracheitis is not always easy, but it is of paramount importance. Some of the differentiating signs and symptoms are shown in Table 52–8.

Diphtheria

Worldwide, diphtheria is an uncommon laryngeal infection, although a report from India suggests that outbreaks still occur there.[30] Laryngeal diphtheria is caused by *Corynebacterium diphtheriae* and generally affects children over the age of 5 years. A febrile illness of slow onset associated with sore throat and hoarseness is then followed by progressive airway obstruction.

The microorganism causes an inflammatory response in the mucous membranes, which results in a thick, grayish-green, plaque-like membranous exudate over the tonsils, pharynx, and laryngeal structures. Characteristically, the exudate is difficult to dislodge and bleeds when removed. The diagnosis requires that the clinician have a high index of suspicion confirmed by observing the typical oropharyngeal findings. Cultures and smears are obtained for confirmation.

Treatment consists of establishing a safe airway via a tracheostomy (intubation is contraindicated because it may dislodge a portion of the plaque and cause airway obstruction) and administering diphtheria antitoxin and penicillin or erythromycin to eradicate the microorganisms. Mortality results largely from the neuropathies that develop secondary to the diphtheria toxin. Even in the patient immunized against diphtheria, the disease may still occur but will tend to be mild.

Spasmodic Croup

Spasmodic croup, or "false croup," is a noninfectious form of laryngeal inflammation associated with a mild, chronic-intermittent, croup-like pattern. Spasmodic croup generally affects children 1 to 4 years of age and may be associated with a mild respiratory tract infection but not with a febrile illness. An

TABLE 52–8. Some of the Distinguishing Characteristics of Laryngotracheitis (Croup) and Supraglottitis (Epiglottitis)

Feature	*Croup*	*Epiglottitis*
Age	Less than 3 y	Over 3 y
Onset	Gradual (d)	Rapid (h)
Cough	Barky	None
Posture	Supine	Sitting
Drooling	No	Yes
Radiograph	Steeple sign, narrowed subglottis	Thumb sign, enlarged epiglottis, dilated hypopharynx
Etiology	Viral	Bacterial
Treatment	Supportive (croup tent, corticosteroids)	Airway management (intubation or tracheostomy) and antibiotics

affected child generally awakes at night with a barking cough, stridor, and mild dyspnea of sudden onset. Nocturnal attacks may occur as isolated events or recur over two to three nights, but generally the child is asymptomatic during the day.

Typically, each episode subsides spontaneously, within a few hours. Examination may show mildly erythematous laryngeal mucosa, with boggy mucosal edema present in the subglottis. Humidified oxygen is generally recommended as the principal treatment. Awareness of spasmodic croup is important in helping to differentiate it from viral laryngotracheitis.

Owing to the pattern of the affliction, allergy was once assumed to be the underlying cause of spasmodic croup. Currently, although the etiology of spasmodic croup remains uncertain, recent evidence suggests that extraesophageal reflux may frequently be the cause.[9] When spasmodic croup is suspected, 24-hour pH monitoring may be diagnostic. If, after a positive pH study, the patient's condition is alleviated by antireflux therapy, it may be concluded appropriately that reflux was the cause.[25]

Traumatic Laryngitis

Traumatic laryngitis is most commonly caused by vocal abuse, such as excessive shouting or yelling, but it also can result from persistent coughing, inhalation of toxic fumes, or direct endolaryngeal injury. Such patients present with varying degrees of hoarseness and odynophonia. The mucosa of the true vocal folds is hyperemic from dilated vessels present on the superior and free surfaces. Edema within Reinke's space develops and submucosal hemorrhage may occur. This form of laryngitis is self-limited and subsides within a few days when treated with voice conservation and humidification.

Vocal nodules may become chronic in children who continually abuse their voices. Nevertheless, for two reasons, surgical treatment is rarely indicated: first, it is difficult to modify the vocal behaviors of such children, and when such nodules are removed, they usually promptly recur. Second, the nodules resolve spontaneously in most children before puberty. One notable exception is female cheerleaders, who may still have nodules in adulthood.

ACUTE LARYNGEAL INFECTIONS OF ADULTS

Viral Laryngitis

Laryngitis Owing to Viral Upper Respiratory Infection Acute viral laryngitis in adults is common and is generally less serious than in children because of the larger adult airway, which is able to accommodate more swelling without compromise of the airway. The typical type of acute laryngitis seen in adults is almost always viral. Influenza and parainfluenza viruses, rhinoviruses, and adenoviruses are the most common causative agents, although many other viruses have been implicated. Adult patients with viral laryngitis do not usually seek medical attention unless they are professional voice users, in which case, the laryngitis may be of great significance to the patient's ability to earn a living. Such patients present with symptoms of a generalized viral syndrome (low-grade fever, malaise, rhinitis) and hoarseness. The dysphonia is characterized by voice breaks, episodic aphonia, and a lowering of pitch. Often a hoarse cough is present.

Characteristically, the laryngeal mucosa is diffusely erythematous and edematous, especially over the true vocal folds. The disease is self-limited and is best treated symptomatically with humidification, voice rest, hydration, cough suppressants, and expectorants. Antibiotic treatment is not usually necessary. In the professional vocalist, corticosteroids are sometimes used to reduce the vocal fold edema, particularly during the recovery phase. However, the physician's administration of such anti-inflammatory medications does not imply that the vocalist with laryngitis can or should perform.

Laryngotracheitis Adult infectious croup is uncommon; when it does occur, it is quite similar to the laryngotracheitis seen in children. The viral prodrome lasts 1 to 7 days, followed by the development of a barking cough and sometimes inspiratory stridor. Throat cultures are usually negative. On fiber-optic laryngoscopy, adults with laryngotracheitis demonstrate subglottic edema and erythema. It has been reported that a relatively large proportion of adults with this syndrome require airway intervention.[31]

Herpes Simplex Herpes simplex infection is ubiquitous, may affect any age group, and, uncom-

monly, may infect the larynx. Most cases of herpetic laryngitis have been reported in very young or debilitated patients. At the time of delivery, a neonate passing through the birth canal may contract genital herpes from a mother with active disease. Subsequent herpes infection in the infant may involve the upper airway and, if the larynx is involved, may cause acute airway obstruction.[32] Therefore, when a child develops stridor in the neonatal period, herpetic infection should be one of the diagnoses considered. Today, laryngeal herpes is most commonly seen in the immunocompromised patient, although herpetic epiglottic infection, causing airway obstruction in otherwise healthy adults, has been reported.[33]

Herpes infection should be suspected whenever a patient presents with a painful vesicular mucosal eruption. After the vesicles rupture, ulceration and tissue necrosis may occur, and the surrounding inflammatory response may be intense. The diagnosis of a primary herpetic infection may be made by serologic testing; however, swab, culture, and polymerase chain reaction detection of viral deoxyribonucleic acid (DNA) are the most reliable diagnostic tests. Symptomatic treatment depends on the site of involvement; topical or systemic acyclovir, or another specific antiherpetic medication, may hasten recovery.

Bacterial Laryngeal Infections

Bacterial laryngitis may develop secondary to purulent rhinosinusitis or tracheobronchitis and generally is less commonly diagnosed in adults than in children. Supraglottic involvement, as in pediatric patients, is the most common form of bacterial laryngitis.

Supraglottitis in the Adult In adults, supraglottitis is manifest by fever, sore throat, a muffled voice, dysphagia, and odynophagia. The onset of symptoms prior to presentation is typically longer than that seen in children (usually more than 24 hours). The diagnosis of supraglottitis is made by observing the swollen, bright-red epiglottis and/or supraglottic structures by fiber-optic laryngoscopy or a swollen epiglottis and dilated hypopharynx on a lateral neck radiograph.

Supraglottitis in adults has a slightly different clinical picture. In adults, the infectious etiology is less likely *Haemophilus* and more likely group A streptococcus. The clinical course appears less severe, with less seasonal variation and airway compromise. Although one must never be complacent with management of the airway, conservative airway management in an intensive care setting is often successful, and tracheostomy is rarely required.[34,35] Conservative measures include oxygenation, humidification, hydration, corticosteroids, and intravenous antibiotics.

In adults, there appear to be two relatively useful clinical predictors of the need to establish an airway. Patients who present to the emergency department less than 8 hours after the onset of a sore throat and patients who are drooling at presentation (in preference to swallowing because of severe odynophagia) almost always require airway intervention.[36]

Laryngeal Abscess Laryngeal abscess is a complication of perichondritis and a recognized sequela of bacterial laryngeal infection. It is more common in adults than in children, even though epiglottic abscess occurs most often as a complication of supraglottitis. Often the adult patient who develops a laryngeal abscess has a preexisting, predisposing laryngeal condition, such as prior irradiation for cancer.

In the preantibiotic era, typhoid fever was a frequent cause and was usually fatal. Other less frequently associated infections include measles, scarlet fever, erysipelas, gonorrhea, syphilis, tuberculosis, and diphtheria. Laryngeal abscess may also be a complication of prolonged nasogastric or endotracheal intubation.[37]

The symptoms of laryngeal abscess are similar to those of supraglottitis. Localizing tenderness to palpation over the laryngeal framework is its hallmark. Fluctuance in the anterior neck, although uncommon, denotes necrosis of the thyroid cartilage. The diagnosis of laryngeal abscess may be made by computed tomographic scanning of the larynx and by fiber-optic or direct laryngoscopy. No current radiographic method accurately predicts the condition of the laryngeal cartilages unless there is obvious widespread destruction.

When abscess is suspected and the patient's airway is marginal, a tracheostomy should be performed under local anesthesia prior to direct laryngoscopy. If the airway is not compromised, direct laryngoscopy, with laser-assisted incision and drainage of the abscess, may suffice as initial treat-

ment. Granular mucosa should be removed, and an attempt should be made to determine if necrotic cartilage is exposed within the endolarynx. Infection of the laryngeal framework can lead to severe laryngeal stenosis.

Gram stain and bacteriologic cultures of the abscess contents or necrotic debris should be obtained routinely, and the results should influence the choice of antibiotic. In cases not complicated by necrotic cartilage, conservative management is usually successful. When necrotic cartilage is present, the combination of surgical débridement, prolonged parenteral antibiotic therapy, and hyperbaric oxygen therapy may enhance recovery and preserve laryngeal function. When all other alternatives have failed, laryngectomy may be considered.

Gonorrhea Gonorrhea is a sexually transmitted genital and sometimes oropharyngeal infection caused by the bacteria *Neisseria gonorrhoeae*. The genital infection may be asymptomatic, so carriers can unknowingly infect their sexual partners. Pharyngeal gonorrhea is transmitted by orogenital contact and is manifested as a diffuse and severe exudative pharyngitis that may directly or indirectly involve the larynx. The infection may produce a pseudomembranous inflammation, which may be confused with diphtheria or streptococcal pharyngitis. Diagnosis is made by culture and identification of the *N. gonorrhoeae* microorganism from swabs of the pharynx (and of the genitalia). Today, most cases of laryngopharyngeal gonorrhea are treated with a single intramuscular dose of ceftriaxone or a course of oral cefixime.

CHRONIC INFECTIONS (GRANULOMATOUS DISEASES)

Granulomatous infections (eg, tuberculosis and syphilis) were recognized long ago and continued to be common afflictions throughout the world in the preantibiotic era. During the twentieth century, these conditions appeared to be on the decline, and some of them had all but disappeared until relatively recently. With the advent of effective anticancer chemotherapy, organ transplantation, and human immunodeficiency virus (HIV) infection, it was discovered that the microorganisms causing many of the granulomatous diseases thrive in the immunocompromised host.

Granulomas are nodular histopathologic lesions characterized by a central mass of epithelioid cells and giant cells surrounded by lymphocytes and other inflammatory cells. Central necrosis (caseation) is seen in many granulomatous conditions and is conspicuously absent in others.

Unlike children, adults have many chronic forms of granulomatous laryngitis, and these may go unrecognized for many years. Chronic granulomatous conditions involving the larynx may be attributable to numerous different types of microorganisms (bacteria, fungi), to an autoimmune process, or to an idiopathic cause. Although not included in this section, some parasitic infections may also cause laryngeal granulomas.

Granulomatous lesions of the larynx may appear as smooth, diffuse swellings of the affected tissues; diffuse cobblestone mucosa; well-defined, discrete nodules; or an ulcerating inflammatory mass. Sometimes, granulomatous diseases mimic laryngeal carcinoma, so the necessity of biopsy is obvious. However, even when the suspicion of malignancy is low, biopsy is frequently necessary to make a diagnosis. Some of the distinguishing features of the most frequent laryngeal granulomatous conditions are shown in Table 52–9.

Granulomatous Diseases Caused by Bacteria

Tuberculosis In 1900, laryngeal tuberculosis was very common, occurring in approximately half of patients with advanced pulmonary disease. With the discovery of effective antituberculous drugs, the incidence of both pulmonary and laryngeal tuberculosis rapidly declined. Nevertheless, tuberculous laryngitis remains one of the most common granulomatous diseases of the larynx, and, today, it is frequently unassociated with advanced active pulmonary disease. In 1993, Ramadan et al reported 16 patients with laryngeal tuberculosis, only 5 of whom had active cavitary lung disease.[38] Because persons with laryngeal tuberculosis often have no simultaneous or previous pulmonary involvement, the clinical presentation may often be similar to that of a neoplastic process.[39] A biopsy is necessary to obtain an accurate diagnosis.

Patients with laryngeal tuberculosis commonly present in their third to fourth decade of life with varying symptoms, including hoarseness, odynopha-

TABLE 52–9. Some Distinguishing Characteristics of Granulomatous Conditions That Commonly Affect the Larynx

Tuberculosis	Posterior third of larynx involved
Leprosy	Supraglottic involvement
Scleroma	Catarrhal stage; Mikulicz's cells
Actinomycosis	Draining sinuses; sulfur granules
Syphilis	Painless ulcers; positive syphilis serology
Candidiasis	Leukoplakia-like lesions; identifiable on Gram stain
Blastomycosis	Painless ulcers; microabscesses
Histoplasmosis	Anterior laryngeal involvement; pseudocarcinoma
Coccidiomycosis	Painless abscesses; spores seen on histology
Sarcoidosis	Supraglottic swelling, nodules, granulomas
Wegener's granulomatosis	Subglottic involvement; necrotizing vasculitis; pulmonary and/or renal involvement

gia, and otalgia. Respiratory obstruction may develop in later stages of the disease, and approximately one-quarter of patients with laryngeal tuberculosis have airway obstruction at the time of their initial presentation.

Laryngeal examination may reveal diffuse edema and hyperemic, hypertrophic mucosa involving the posterior third of the larynx, or the process may be diffuse, nodular, and ulcerative. In the nodular, ulcerative (later) stages, tuberculosis may easily be confused with laryngeal carcinoma.

It is a misconception that interarytenoid (posterior commissure) involvement, alone, is the most commonly observed pattern.[40] In the modern era, the true vocal folds appear to be the most commonly involved site.[41]

The diagnosis is made by demonstrating typical caseating granulomas on histology and acid-fast staining microorganisms by smear and/or culture. Treatment with antituberculous drugs usually resolves both the pulmonary and the laryngeal disease. If tuberculous laryngitis is left untreated, cicatricial laryngeal stenosis with vocal fold fixation may develop, necessitating surgical correction or tracheostomy.

Leprosy (Hansen's Disease) Leprosy is caused by infection with the bacillus *Mycobacterium leprae*. The disease usually affects the skin and peripheral nerves but may affect other organ systems. Cutaneous nodules may become disfiguring and will eventually cause peripheral neurologic damage and muscular weakness. In 2000, 738,284 cases of leprosy were identified worldwide. In 1999, 108 cases were reported in the United States. Over one-third of the global cases are limited to India.[42]

The larynx is involved in approximately one half of patients with leprosy, and it most consistently involves the epiglottis. As the disease progresses to involve the glottis, hoarseness develops, and later destruction of the cartilaginous laryngeal framework may occur, as well as lepromatous nerve involvement, which may cause laryngeal paralysis.[43,44] With lepromatous laryngitis, odynophagia, odynophonia, and referred otalgia are uncommon, but they do occur.

In patients with lepromatous leprosy, direct laryngoscopy typically reveals a nodular, edematous supraglottis with ulceration. Diagnosis is made by biopsy, which reveals a chronic inflammatory cell infiltrate with foamy leprous cells that contain the bacillus *M. leprae* (Hansen's bacillus). Nasal smears for the intracellular microorganisms may be diagnostic. Treatment consists of the long-term use of diaminodiphenylsulfone (dapsone). Tracheostomy may be required if laryngeal stenosis develops.

Scleroma Scleroma, previously called "rhinoscleroma" because of its predilection for the nose, is a chronic infection caused by *Klebsiella rhinoscleromatis*. Scleroma remains endemic today in Europe, Mexico, Central America, South America, and Egypt. Most cases found in the United States occur in recent immigrants.[45–47]

The disease primarily involves the mucosa and submucosa of the nasal cavity but may also involve the larynx, especially the subglottic region, and the trachea. Laryngeal involvement has been reported in approximately 15 to 80% of cases.[48]

The disease has three distinct stages: (1) *the catarrhal stage,* characterized by persistent puru-

lent rhinorrhea, with nasal crusting and obstruction; (2) *the granulomatous stage,* characterized by small, painless granulomatous nodules within the upper respiratory tract, including the larynx; and (3) *the sclerotic stage,* in which the glottis and subglottis are usually involved and hoarseness and respiratory obstruction may develop. The progression through these three stages usually takes many years.

The diagnosis is made by isolating the microorganism from the tissues, although positive complement fixation and agglutination tests are highly suggestive. Foamy vacuolated histiocytes (Mikulicz's cells) and bloated plasma cells with red birefringent inclusions (Russell bodies) are seen histologically.

Treatment consists of intravenous aminoglycosides, tetracycline, or cephalosporins. Laryngeal dilatation, endoscopic resection, and tracheostomy may be required during the sclerotic phase. Untreated rhinoscleroma may cause death owing to airway obstruction. Progression of laryngeal scleroma to laryngeal carcinoma has been suggested in some affected persons.[49]

Actinomycosis Cervicofacial actinomycosis is a chronic suppurative disease caused by the anaerobic bacterium *Actinomyces bovis* or *Actinomyces israelii.* Initial involvement of the cervical or mandibular region leads first to paralaryngeal and then to laryngeal disease. Pain is the most common initial manifestation, followed by hoarseness, cough, and, eventually, airway obstruction.

The larynx appears diffusely erythematous and swollen with draining sinuses; its consistency is firm and woody.[50] Diagnosis is made by identifying the typical "sulfur granules" in biopsy material and by culturing the microorganism. Long-term therapy with penicillin or tetracycline is effective. Stenosis and laryngeal fixation secondary to deep ulcerations and chondritis may develop if the disease is left untreated. Laryngeal dilatation, arytenoidectomy, or tracheostomy may be required.

Nocardia species of the *Actinomyces* family are soil saprophytes that are widely distributed throughout the world. Like other *Actinomyces*, *Nocardia* may infect humans by inhalation or through a break in the skin. Nocardiosis is characterized by diffuse microabscesses that, on histologic examination, show a neutrophilic predominance; sulfur granules are atypical. Aerodigestive tract involvement is common. The diagnosis is made by culture and isolation of the *Nocardia* microorganism, and treatment is with systemic sulfisoxazole.

Tularemia Tularemia, also called "rabbit fever" and "deer fly fever," is caused by the bacterium *Francisella tularensis.* It is found only in the northern hemisphere (Europe, Asia, and North America). Most cases occur from contact with rabbits or squirrels, but many other wild and domesticated animals have been reported to carry the disease. Transmission to humans can also occur by a bite from a tick or a deer fly, which are the intermediate hosts and insect vectors.

In humans, the most common portal of entry is through the skin or mucous membranes. Headache, myalgia, and malaise are common symptoms. Oropharyngeal tularemia, which occurs in approximately 1% of cases, produces an intense exudative pharyngitis associated with lymphadenopathy.[51] Diagnosis is made by serologic tests since the microorganisms are difficult to identify on culture or histologic examination. Treatment is with streptomycin or gentamicin for 7 to 10 days.

Glanders At one time in history, glanders was "the plague of horses," and, secondarily, it affected man. Today, glanders is rare around the world but still occurs in Asia, Africa, and South America. The disease is caused by infection by the bacterial organism *Pseudomonas mallei* (*Burkholderia mallei*). Transmission is still by contact with an infected horse or by inhalation or inoculation of contaminated material, and infected humans are almost exclusively horse handlers.[52]

Infection by inoculation of broken skin causes systemic infection characterized by fever, malaise, prostration, pneumonia, and lymphadenopathy. Truly systemic (septicemic) involvement may be fatal. Infection by inhalation produces an intense, ulcerative mucopurulent granulomatous reaction in the mucous membranes of the aerodigestive tract and pneumonia. Treatment is with sulfonamides.

Syphilis Syphilis is a sexually transmitted spirochetal infection caused by *Treponema pallidum.* Syphilitic chancres do not usually involve the larynx, and, most commonly, the larynx becomes involved during the secondary and tertiary stages. However, tertiary syphilis may not develop until years after the

initial infection. Diffuse erythematous papules, painless superficial ulcers, and cervical lymphadenopathy are seen during the secondary stage and generally clear without treatment within several weeks. Gumma formation during the tertiary stage can lead to laryngeal fibrosis, chondritis, and stenosis.

The typical laryngeal findings are diffuse erythema and edema, with necrotic ulcers that mimic carcinoma or tuberculous laryngitis. Serologic tests for syphilis are diagnostic, and a lumbar puncture should be performed to rule out central nervous system involvement. Penicillin is still the treatment of choice.[53]

Although very rare, congenital syphilitic laryngitis can occur in infants born to mothers with syphilis. This diagnosis should be considered in the differential diagnosis of neonates with laryngeal stenosis. Treatment should be instituted if there is serologic evidence of active disease.[54]

Mycotic (Fungal) Granulomatous Diseases

Candidiasis Although laryngeal candidiasis (usually infection by species, *Candida albicans*) is infrequently reported, its occurrence is not rare in laryngologic practice. Laryngeal candidiasis virtually never occurs in the absence of identifiable (local and/or systemic) predisposing factors, but it does sometimes occur in reasonably healthy patients who are not severely immunocompromised.

Among the risk factors for development of laryngeal candidiasis are the use of corticosteroids and broad-spectrum antibiotics, diabetes, burns, alcoholism, endotracheal intubation, and previous (viral or bacterial) laryngeal infection. *Candida* laryngotracheitis, for example, may occur as a complication of treatment in previously healthy children hospitalized for laryngotracheitis and treated with systemic corticosteroids and broad-spectrum antibiotics.[55]

Cases of laryngeal *Candida* infection may occur in ambulatory patients as well. A relatively benign, isolated form of laryngeal candidiasis (without the involvement of any other contiguous anatomic structures) occurs in some patients who use corticosteroid inhalers. Typically, such patients have used them on a daily basis for years. (The most common indication for the chronic use of corticosteroid inhalers is asthma.) This form of isolated laryngeal candidiasis has a distinctive appearance: intense, diffuse laryngeal erythema with an irregular, friable, white exudate (occurring most notably on the true vocal folds). The infection is superficial, and no ulceration or necrosis is seen (Figure 52–4). Patients with this particular condition usually complain of hoarseness but no other symptoms. The diagnosis is usually made on the basis of the history and clinical findings. A 2-week course of treatment with nystatin (swish and swallow) or an oral antifungal agent, such as ketoconazole or fluconazole, is effective treatment for most patients with this type of *Candida* laryngitis; however, in some cases, treatment must be continued for several weeks to eradicate the disease. Candidal esophagitis should be considered in a patient with laryngeal involvement and dysphagia.

Severe immunosuppression (owing to chemotherapy or AIDS) is more often associated with the development of invasive laryngeal candidiasis, and, as expected in such patients, the infection is often more serious and sometimes can be life-

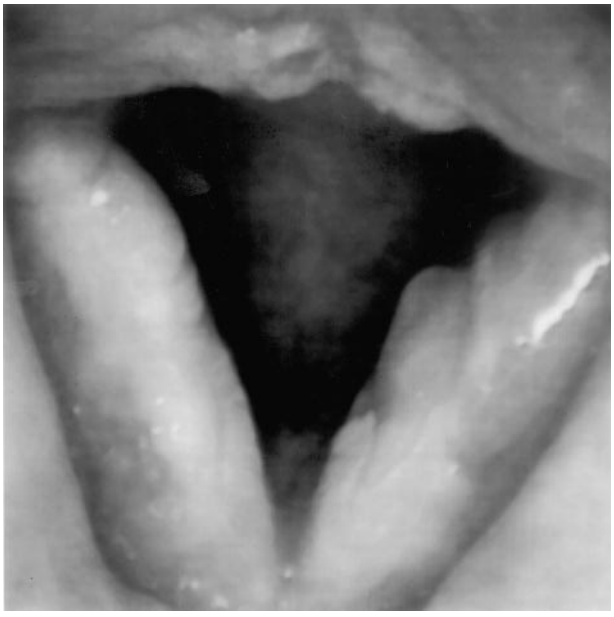


FIGURE 52–4. Candidiasis of the larynx. This is the appearance of the larynx of a 45-year-old woman with lifelong asthma who had been using corticosteroid inhalers for many years. This appearance is different from that usually seen in immunocompromised patients, the latter group having a more invasive, necrotizing form of candidiasis.

threatening owing to bleeding or airway obstruction. In the immunocompromised patient, laryngeal candidiasis may be caused by local aerodigestive tract infection, which can subsequently give rise to widespread systemic candidiasis. The converse, secondary involvement of the larynx by disseminated candidiasis, is uncommon. However, infection is rarely confined to the larynx but usually involves adjacent areas of the airway and esophagus as well.

Sometimes the esophagus is involved but, initially, not the larynx. When a patient presents with dysphagia and barium esophagography or transnasal esophagoscopy reveals *Candida* esophagitis, the clinician should suspect that the patient has AIDS. Indeed, aerodigestive tract candidiasis may be the first presenting manifestation of AIDS.

Invasive *Candida* laryngitis in the immunocompromised host produces painful, ulcerative lesions and deep tissue necrosis and may progress rapidly. In addition to hoarseness, patients with this type of infection complain of sore throat, dysphagia, and odynophagia. For the diagnosis to be made, it must be suspected. Unlike other fungi, *Candida* species can easily be identified on Gram stain.[56] Confirmation of the diagnosis is made by the histopathology and culture of biopsied tissue. Invasive laryngeal candidiasis is treated with parenteral amphotericin B and, when necessary, airway support.[57]

Blastomycosis North American blastomycosis is a chronic pulmonary infection caused by the fungus *Blastomyces dermatitidis*. Laryngeal involvement occurs in 2 to 5% of cases.[56,58] The microorganism is generally found in damp areas where decaying wood is present. Its distribution in the United States and Canada is concentrated around the Great Lakes and along the Mississippi, Ohio, and St. Lawrence rivers. Primary laryngeal blastomycosis has been reported.[59]

Patients typically present with multiorgan systemic involvement and severe hoarseness and cough when the larynx is involved. The microorganism produces erythematous, granular, mucosal lesions in the larynx, which progress to small, painless abscesses and ulcerations. Histologically, caseous necrosis with abundant acute inflammatory cells and microabscesses are seen, as well as giant cells in the surrounding tissue. Pseudoepitheliomatous hyperplasia is a characteristic change seen in the epithelial layer. The fungus in yeast form may be seen in the region of the microabscesses and is periodic acid–Schiff positive. Treatment is with long-term oral itraconazole, with amphotericin B being reserved for very severe or recalcitrant cases. In the absence of treatment, progressive fibrosis with vocal fold fixation develops, as do pharyngocutaneous fistulaes.

Histoplasmosis Histoplasmosis is a systemic mycotic disease caused by *Histoplasma capsulatum* and may involve the larynx (Figure 52–5). Nodular superficial granulomas that may ulcerate and become painful involve the anterior portions of the larynx and epiglottis. Esophageal involvement may occur (Figure 52–6). Histologic examination shows granulation tissue composed of plasma cells, microorganism-laden macrophages, lymphocytes, and giant cells, which may be confused with the granulation tissue of carcinoma or tuberculosis. Diagnosis is made by culture and the complement fixation test.

Amphotericin B is the treatment of choice. Laryngeal stenosis may develop when extensive ulceration leads to chondritis. In this instance, laryngeal dilatation, arytenoidectomy, or tracheostomy may be required to provide a safe airway.

Coccidioidomycosis Coccidioidomycosis, also called "desert fever" (or "San Joaquin Valley fever" in the United States), is caused by the microorganism

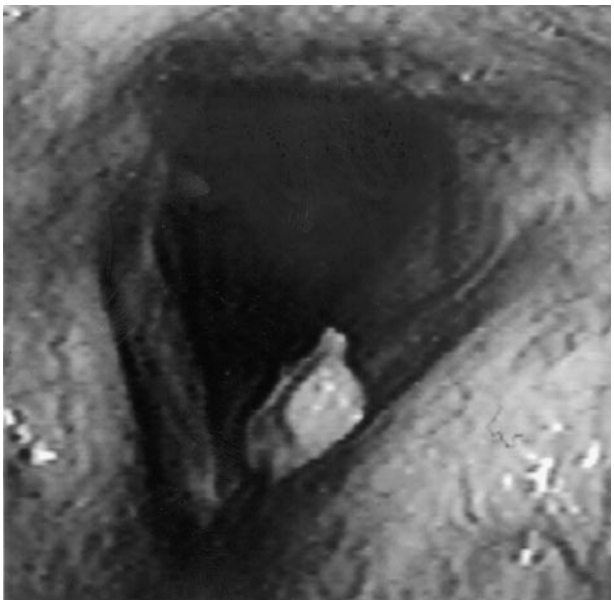


FIGURE 52–5. Histoplasmosis of the larynx. Note pedunculated lesion of the left true vocal fold.

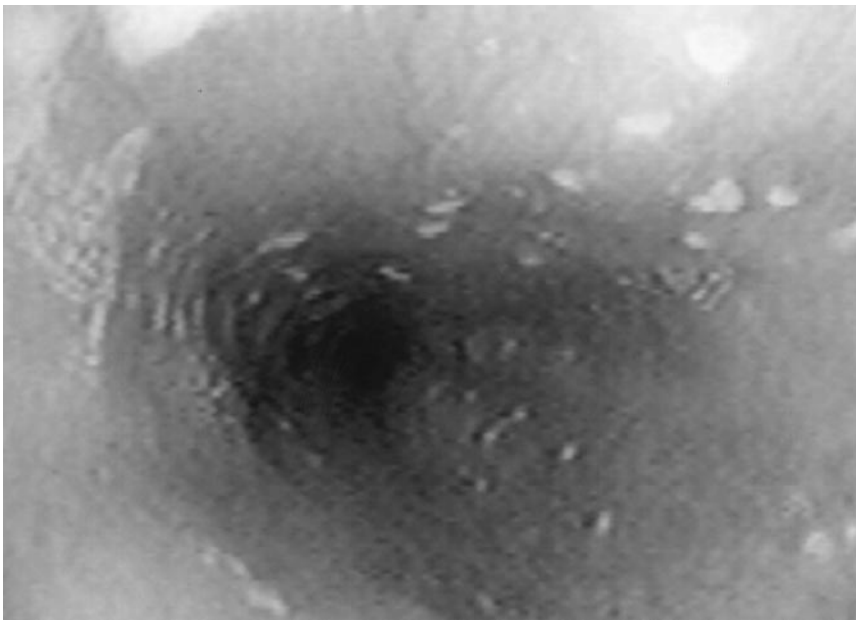


FIGURE 52–6. Diffuse esophageal histoplasmosis.

Coccidioides immitis, which is found in desert soil. It is primarily a pulmonary fungal infection that is endemic to the southwestern United States, Mexico, and central South America. Reportedly, 60% of people with this infection are asymptomatic; 40% develop a flu-like illness, and among those, 0.5% develop a systemic, disseminated, more severe form of the disease. Patients with the disseminated form may develop hoarseness, cough, and airway obstruction owing to laryngeal coccidioidomycosis.[60] Disseminated extrapulmonary disease is far more common in males and in blacks.[61]

Laryngeal disease usually develops during the acute phase of the primary infection. Laryngeal involvement can, however, develop in patients who have had granulomatous lung disease for months or years. In addition to the laryngeal findings of intense, diffuse laryngeal erythema (with or without focal ulceration), most patients with *C. immitis* have cervical lymphadenopathy. Histology reveals caseating granulomas with multinucleated giant cells and pathognomonic, double-walled endospores. The diagnosis is made by serologic testing and biopsy of affected tissue. Treatment is with amphotericin B. When death occurs, it is usually owing to meningeal involvement.

Aspergillosis Aspergillosis is generally an infection of immunocompromised patients, and respiratory tract involvement is common. When the larynx is involved, patients complain of hoarseness, dysphagia, and, sometimes, symptoms of airway obstruction.

In the immunocompromised patient, *Aspergillus* infection is usually necrotizing, invasive, and associated with a poor prognosis. Despite aggressive antifungal treatment with amphotericin B and attempted wide surgical excision (including laryngectomy), most such patients with this infection die of progressive disease.[62]

Sporotrichosis Sporotrichosis, an uncommon fungal infection of the skin or airway, is caused by *Sporothrix schenckii* and occurs worldwide. The causative fungus is most commonly found in sphagnum moss and wood. People who work with wood usually get the cutaneous form of sporotrichosis, whereas most cases of laryngeal sporotrichosis occur in people working with the moss.[63]

The more common cutaneous form of sporotrichosis causes granulomas in the subcutaneous layer of the skin and in regional lymph nodes. If the mucous membranes of the upper airway are damaged or abraded for any reason, inhalation of the fungus may result in laryngopharyngeal infection. Hoarseness and cough are the most common symptoms, and the lesions appear granulomatous. Diagnosis is made by biopsy and by culturing the microorganism. Oral potassium iodide is sufficient

treatment for cases with superficial involvement; deep tissue involvement requires a course of amphotericin B therapy.[64]

Idiopathic Granulomatous Diseases

Sarcoidosis Sarcoidosis is a slowly progressive, rarely fatal, systemic granulomatous disease of unknown cause. The disease afflicts blacks 10 times more commonly than whites. The lungs and skin are most commonly involved, and laryngeal sarcoidosis occurs in 1 to 5% of cases.[65,66] When the skin of the nasal rim is affected, upper respiratory sarcoidosis involvement is seen in approximately 75% of cases[67] (Figure 52–7). Rarely, the larynx may be involved without clinical or radiographic evidence of lung involvement.[68]

Laryngeal sarcoid usually involves the supraglottic larynx and sometimes the subglottis but typically spares the true vocal folds. Characteristically, the entire supraglottis appears pale pink and massively edematous, sometimes obscuring visualization of the vocal folds. The turban-like appearance of the epiglottis is virtually pathognomonic.[65] Less commonly, some laryngeal sarcoidosis patients present with a few discrete, sometimes hemorrhagic, nodules (up to 1 cm in diameter) on the epiglottis or other supraglottic structures.

Laryngeal sarcoidosis is a diagnosis of exclusion based primarily on finding noncaseating granulomas and diffuse edema with miliary nodules involving mainly the supraglottic structures and on excluding tuberculosis and other fungal diseases. Patients may present with hoarseness and varying degrees of airway obstruction; however, ulcerative lesions and pain are rare.

The use of systemic and intralesional corticosteroids generally results in improvement or apparent resolution of the lesions. In severe cases, however, endoscopic dilatation and laser resection of involved supraglottic tissues, or tracheostomy, may be necessary.

Wegener's Granulomatosis Wegener's granulomatosis is a systemic disease of unknown etiology, characterized by necrotizing granuloma with vasculitis involving the upper respiratory tract, lungs, and kidneys. On presentation, its laryngeal involvement may resemble acute laryngitis, but the eventual development of granulomatous ulcers throughout the larynx may lead the clinician to suspect the diagnosis. Subglottic stenosis occurs in approximately 20% of cases and can lead to significant airway obstruction, requiring tracheostomy and surgical correction of the stenosis.[69] Diagnosis is based on typical histologic findings of necrotizing granulomas and vasculitis. The anticytoplasmic autoantibody test (C-ANCA) is highly specific for Wegener's granulomatosis. Recommended treatment includes cyclophosphamide with corticosteroids.

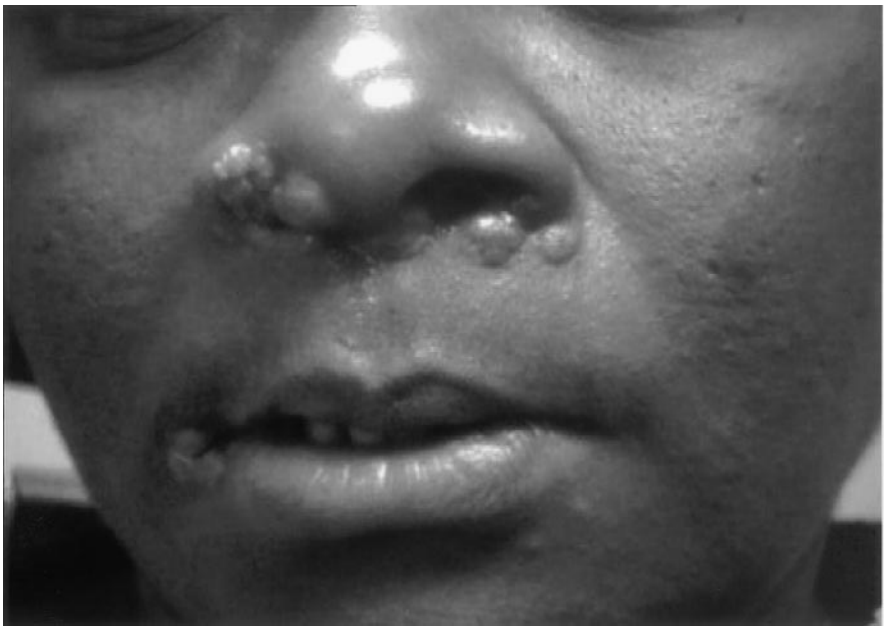


FIGURE 52–7. Sarcoidosis of the nasal rim.

ALLERGIC, IMMUNE, AND IDIOPATHIC DISORDERS

Hypersensitivity Reactions

The term *hypersensitivity* implies an overzealous response of the immune system to an antigenic stimulus. Hypersensitivity reactions include conditions such as allergic rhinitis, contact dermatitis, and urticaria, but these conditions do not produce life-threatening, obstructive edema within the airway.

Anaphylaxis, an acute and profoundly life-threatening immune-mediated allergic response, is made up of a triad of clinical manifestations: (1) flushing, pruritus, and/or urticaria; (2) airway obstruction (angioedema, laryngospasm, and/or bronchospasm); and (3) circulatory collapse (shock).

Angioedema Angioedema, which may occur with or without anaphylaxis, is an acute, allergic, histamine-mediated, inflammatory reaction characterized by acute vascular dilation and capillary permeability. This reaction can occur in many parts of the body; when the airway is acutely affected, the condition is potentially life-threatening. Oral and laryngopharyngeal structures are frequently affected.

In susceptible patients, angioedema can be precipitated by medications (eg, penicillin, aspirin, other nonsteroidal anti-inflammatory drugs, and angiotensin-converting enzyme inhibitors), by food additives and preservatives, and by blood transfusions, infections, or insect bites. In addition, the condition may be associated with a coexisting connective tissue disorder.[70]

Hereditary angioedema is an autosomally dominant inherited deficiency of C1 esterase inhibitor that leads to recurrent attacks of mucocutaneous edema. An acquired form of C1 esterase deficiency has also been reported in patients with angioedema who have occult lymphoma.

Diagnosis is made primarily from the history, although the offending agent may not be readily apparent. Patients present with edema that may involve the face, oral cavity, oropharynx, or larynx. The onset is rapid and may be associated with pruritus. Hoarseness is common when the larynx is involved.

Treatment of both types of angioedema must be prompt and aggressive. Epinephrine, corticosteroids, antihistamines, and aminophylline are the mainstays of therapy. If progressive airway obstruction develops, intubation or tracheostomy may be required. Chronic "pretreatment" of hereditary angioedema with danazol appears to elevate levels of functional C1 esterase inhibitor and to help prevent recurrent episodes.

Stevens-Johnson Syndrome Stevens-Johnson syndrome is a mucocutaneous hypersensitivity reaction (at the severe end of the erythema multiforme spectrum) usually triggered by medications (eg, sulfonamides, phenobarbital, and carbamazepine). It is an acute febrile illness characterized by conjunctivitis, rash, and severe oropharyngeal mucositis. The mucosal and skin lesions rapidly progress to bullae formation and desquamation of the skin in sheets. The most severe form of this syndrome is called toxic epidermal necrolysis, and this variant may involve the larynx.[71] Diagnosis is made by biopsy, and treatment is primarily supportive, including tracheostomy when the airway is compromised. If death occurs, it usually is owing to sepsis from bacterial superinfection.

Immune and Idiopathic Disorders

Infections of the Immunocompromised Host Immunocompromised patients, whether immunocompromised from diabetes, long-term corticosteroid therapy, chemotherapy, or AIDS, are at risk of developing opportunistic infections of the aerodigestive tract. The most commonly reported opportunistic infections that affect the larynx are shown in Table 52–10. Of those listed, laryngeal candidiasis, tuberculosis, and herpes infections are the most commonly encountered in patients with AIDS.

When any immunocompromised patient develops even slight hoarseness, dysphagia, or odynophagia, the otolaryngologist should suspect infection of the laryngopharynx. Even microorganisms that are usually considered indolent may aggressively invade tissue and cause acute inflammation, massive tissue destruction, and rapid obstruction of the airway. When new throat symptoms arise in the immunocompromised patient, the clinician should perform an examination promptly, and if there is evidence of infection, the clinician should aggressively attempt to obtain material for culture and for histologic examination.

TABLE 52–10. Common Laryngeal Infections of the Immunocompromised Host

Viral
Herpes simplex
Herpes zoster
Cytomegalovirus
Papova (papilloma)
Toxoplasmosis
Bacterial
Tuberculosis
Atypical mycobacteria
Actinomycosis
Fungal
Candidiasis
Aspergillosis
Histoplasmosis
Coccidioidomycosis

Rheumatoid Arthritis Rheumatoid arthritis is a systemic autoimmune disorder of unknown cause that can affect any organ in the body. Its most common manifestation is symmetric polyarthritis, but it also can cause inflammation in nonjoint structures, vasculitis, and pulmonary changes.[72] Rheumatoid arthritis may affect the larynx both directly and indirectly.

Rheumatoid involvement of the cricoarytenoid joints may cause hoarseness or airway obstruction. At postmortem examination, up to 87% of patients with rheumatoid arthritis have cricoarytenoid joint changes, but, based on laryngoscopy, only 17 to 33% of patients have clinical signs of laryngeal involvement, namely posterior laryngeal inflammation and decreased arytenoid cartilage mobility.[73]

Rheumatoid nodules may occur anywhere in the larynx or within the substance of the vocal fold itself, leading to hoarseness. The gross appearance of rheumatoid laryngeal nodules is variable. They may appear as white submucosal nodules, as ulcerated friable polypoid lesions, or as ill-defined masses deep within the substance of the vocal fold. Occasionally, unsuspected rheumatoid nodules are discovered during direct laryngoscopy by palpation of the nodule.

Histologically, these lesions show a central area of fibrinoid necrosis surrounded by histiocytes, plasma cells, and lymphocytes. They can be highly vascularized and hyalinized and may have a fibrous capsule. Frequently, rheumatoid nodules of the larynx are misdiagnosed as pyogenic granulomas.[74]

Rheumatoid arthritis, like other collagen vascular diseases, often involves the esophagus, causing esophageal dysmotility and reflux disease. Thus, patients with rheumatoid arthritis may have reflux laryngitis, but it is not known whether such reflux contributes to the arytenoid fixation.

The choice of treatment for rheumatoid airway obstruction secondary to arytenoid cartilage fixation depends on the patient's overall medical condition. Since surgical rehabilitation of arytenoid cartilage function is not possible, endoscopic arytenoidectomy is usually the treatment of choice. This procedure leaves the patient with an adequate airway and a somewhat breathy, dysphonic voice. Sometimes the rheumatoid arthritis so severely affects the neck that endoscopic exposure of the larynx is not possible. In such cases, an open surgical procedure, or simply a tracheostomy, may be performed.

Because rheumatoid nodules of the larynx frequently lie within the substance of the vocal fold and may be inflamed, the vocal fold may be scarred after their removal. As a consequence, most patients with this type of rheumatoid involvement of the larynx have persistent hoarseness following nodule removal.

Systemic Lupus Erythematosus Lupus is a systemic, autoimmune disease of unknown etiology. It affects women more commonly than men and usually presents in the second and third decades of life. Patients with this condition may have autoantibodies to a variety of different tissues, and head and neck manifestations are common. Although the most common manifestations of lupus are arthritis, malar rash, and photosensitivity, up to 40% of patients have mucosal lesions of the aerodigestive tract as well.

The lesions may be varied, for example, petechiae, ulcerations, or raised nonulcerated lesions with erythematous borders. The palate and nose are commonly involved. Painless nasal septal perforations may occur. The larynx may also be involved by these mucosal lesions or, on occasion, by cricoarytenoid arthritis.[72]

Laryngeal involvement usually occurs at times of acute exacerbation of the systemic disease. Airway compromise is uncommon but does occur. Biopsy reveals a mononuclear cell infiltrate. Positive fluorescent antinuclear antibody tests are important for diagnosis and are a key part of ARA criteria. Corticosteroids and symptomatic therapy are the treatment.

Cicatricial Pemphigoid Pemphigus and pemphigoid are idiopathic, autoimmune epithelial disorders. There are several clinical variations; however, in common, they share subepithelial bullae inflammation. The primary distinctions between the two entities are the clinical patterns and histologic features: in pemphigus, there is suprabasilar separation (cantholysis), and in pemphigoid, the bullae are subepidermal. Clinical variations, that is, the manifestations and sites of involvement, are believed to be the result of autoimmunity to distinct basement membrane antigens. Pemphigus and pemphigoid also may be associated with systemic lupus erythematosus. Of this group of uncommon bullous diseases, only cicatricial pemphigoid appears to involve the larynx with any frequency (9 to 20% of cases).

Cicatricial pemphigoid is a painful, unremitting, chronic-inflammatory, vesiculobullous disease. Cicatricial formation may occur at any site of involvement, including the nose, nasopharynx, pharynx, larynx, and esophagus. Cicatricial pemphigoid affects women twice as often as men, and most patients are over 50 years of age. The oral cavity and eyes are most commonly involved; the aerodigestive tract is sometimes involved. The largest series reported in the otolaryngology literature reported 13 (9%) of 142 patients with laryngeal involvement, 3 of whom required airway intervention.[75] All 13 had involvement of other areas of the aerodigestive tract as well, but isolated laryngeal involvement has been reported. The primary symptom of laryngeal pemphigoid is severe odynophagia, and the most common findings are ulcers of the epiglottis and aryepiglottic folds.

Diagnosis is dependent on biopsy, which shows inflammatory subepithelial bullae surrounded by a mixed cellular inflammatory infiltrate. Immunofluorescent studies usually reveal linear deposition of immunoglobulins (IgG and IgM) along the basement membrane.[75] Long-term treatment is with the anti-inflammatory medication dapsone, with systemic corticosteroids and/or immunosuppressive therapy being reserved for dapsone treatment failures. Intralesional corticosteroids are ineffective. Stenosis of the larynx and other sites usually requires surgical intervention.

Relapsing Polychondritis Relapsing polychondritis is a rare, idiopathic, generally progressive, autoimmune disease that causes inflammation of cartilage. It can mimic rheumatoid arthritis and sometimes occurs in patients with other autoimmune diseases, such as Sjögren's syndrome, systemic lupus erythematosus, and psoriatic and rheumatoid arthritis.

Relapsing polychondritis occurs in all age groups, having a bell-shaped distribution and a peak incidence in the fourth decade. Although only 10% of patients present with respiratory tract involvement (larynx and trachea), more than 50% eventually develop such involvement, and 20% require tracheostomy. Of the 20 to 30% of patients who eventually die of the disease, most die of respiratory complications.[76]

This disease is characterized by episodes of inflammation with subsequent destruction of the cartilage of the ears, nose, and larynx. Arthritis involving the large joints is also common. Laryngeal involvement is manifest by hoarseness, dyspnea, stridor, cough, and, sometimes, pain and hemoptysis. On examination, therefore, a clinical diagnosis is confirmed by cartilage biopsy, although most patients with this disorder also have autoantibodies to type II collagen.

Initially, most patients present with bilateral involvement of the ear cartilage. Typically, the ears suddenly become red, swollen, and tender. With or without treatment, the condition may subside within 5 to 10 days. The next most common sites of involvement are the nose and the costal cartilages; there can be involvement of the eye.

Histologically, the normal cartilage is replaced by an eosinophilic material, and acute and chronic infiltrates of lymphocytes and plasma cells are present. The usual basophilic appearance of the cartilage matrix is lost, lacunae are interrupted, and fibrous tissue replaces cartilage. As the disease progresses, fibrosis and chondronecrosis become marked.

Treatment includes corticosteroids and anti-inflammatory medications such as dapsone. Corticosteroid and immunosuppressive medications are used for patients with severe, recalcitrant, or rapidly

progressive disease, especially when the larynx or other airway structures are involved. Tracheostomy may be necessary in the later stages of the disease.

Sjögren's Syndrome Sjögren's syndrome is an idiopathic autoimmune disorder characterized by the clinical triad of xerostomia (dry mouth), conjunctivitis sicca (dry eyes), and rheumatoid arthritis. Patients with Sjögren's syndrome have a relatively high incidence of developing lymphoma. There is also a "limited" form of the disease, occurring without the arthritis, called "sicca syndrome."

Although the cause (or causes) of these syndromes is unknown, both the limited and the full-blown forms have in common autoantibodies to glandular tissue in the eyes, nose, oral cavity, and laryngopharynx. In addition to the lacrimal glands and the major salivary glands, minor salivary and seromucinous glands are usually affected throughout the aerodigestive tract.

The diagnosis is made clinically using a Schirmer's test to document the dryness of the eyes and by salivary gland biopsy. In the major salivary glands, the histologic picture demonstrates (1) an intense lymphoid infiltrate, especially in periductal areas; (2) glandular atrophy; and (3) myoepithelial hyperplasia. Although the salivary glands are virtually always affected, biopsy of a major salivary gland is rarely necessary. Instead, biopsy of minor salivary gland tissue (lip biopsy) is usually sufficient to make the diagnosis. The histopathologic features seen in minor salivary glands are similar to those seen in the major salivary glands, although the myoepithelial hyperplasia is absent.

The seromucinous glands of the larynx may be involved, leading to inflammation of the larynx similar to that seen in the salivary glands. Clinically, this involvement produces edema, erythema, dryness, crusting, and, hence, chronic hoarseness. Laryngeal Sjögren's syndrome, however, does not occur in isolation; that is, patients with laryngeal symptoms and signs of Sjögren's syndrome also have other manifestations of the disease.[77]

In some cases, the mucosa of the posterior commissure appears so hypertrophic that the clinician must consider the possibility of tumor. The larynx in Sjögren's syndrome usually exhibits intense erythema and hypertrophy of the posterior commissure with dry, tenacious mucus between the vocal folds.

Biopsies of the larynx reveal histologic findings similar to those seen in the salivary glands. In addition, patients with Sjögren's syndrome often have impaired esophageal function and gastroesophageal reflux. Treatment is symptomatic, and antireflux and anti-inflammatory medications are sometimes prescribed.

Amyloidosis Amyloidosis is a dysproteinemia in which a characteristic, amorphous, eosinophilic substance is deposited in the tissues of various organs. Amyloidosis is classified as being either "systemic" or "localized." The systemic types are familial, primary (with or without myeloma), and secondary (usually associated with chronic inflammatory disease). Primary amyloidosis has a 5-year survival of only 20%, with the patients dying of renal, central nervous system, or cardiac involvement.[78] The "localized" type of amyloidosis is almost never fatal and is the type that most commonly involves the larynx. On rare occasions, however, the larynx may become involved by either primary or secondary systemic amyloidosis. In other words, most cases of laryngeal amyloidosis occur in isolation, although simultaneous involvement of the trachea and, to a lesser extent, the bronchi occurs in about one-third of those cases.

On laryngoscopy, amyloidosis appears as diffuse mucosal thickening or subepithelial nodules, localized mainly to the anterior part of the subglottis. Patients are usually asymptomatic until the deposits involve the vocal folds or critically narrow the airway. When amyloidosis is suspected, biopsy specimens should be stained with Congo red, which, when viewed with polarized light, shows a pathognomonic apple-green birefringence.

Laryngeal amyloidosis usually has a benign course. Symptomatic cases are best treated by endoscopic carbon dioxide laser excision of the lesions; laryngeal dilatation and tracheostomy are rarely necessary.[79]

MISCELLANEOUS INFLAMMATORY CONDITIONS

Parasitic Infections

Trichinosis Trichinosis in humans is caused by ingesting meat contaminated with the helminthic roundworm *Trichinella spiralis*. Trichinosis is relatively common worldwide. About 30 cases occur

each year in the United States. In the past, contaminated pork was the most common source of infection, but today, most cases are caused by eating feral meat, such as bear or wild boar.

Humans are particularly susceptible to trichinosis infection. Soon after ingestion, the larvae penetrate the intestinal wall, where copulation and multiplication occur. The next generation of larvae then enters the bloodstream, is distributed throughout the body, and finally enters and grows in skeletal muscle. The muscles of the diaphragm, eyes, tongue, chest, shoulders, and calves are often affected. Laryngeal involvement is uncommon.

In tissue, the larvae elicit an eosinophilic and lymphocytic inflammatory response. The severity of the clinical manifestations depends on the location and density of the larvae. The first symptoms occur within 2 days of ingestion. During the initial stage of infection, diarrhea, nausea, and malaise are common. During the muscle invasion stage (lasting 1 to 6 weeks), fever, weakness, skin rash, myalgia, muscle tenderness, and facial and periorbital edema are usually present. Some cases are complicated by urticaria, splinter hemorrhages, and angioedema. The primary symptom of laryngeal involvement is hoarseness. Trichinosis should be suspected by the history and eosinophilia. Diagnosis can be made by serologic testing and muscle biopsy. Treatment is with a 7-day course of thiabendazole. The disease can be prevented by cooking meat products to an internal temperature of 170°F.

Leishmaniasis Leishmaniasis, although uncommon in the United States, is indigenous throughout the rest of the world. It is estimated that there are 1.5 to 2 million cases each year worldwide. The vast majority of the world's cases occur in India, Bangladesh, Nepal, Sudan, and Brazil. The organism infects rodents and dogs, and transmission to humans is usually from an animal, although the bite of an intermediate host, the sandfly, may cause the disease as well. Although there are several clinical forms of the disease, the mucocutaneous form, caused by *Leishmania braziliensis* and *Leishmania mexicana*, is the one that most commonly involves the airway. Espundia, the form caused by *L. mexicana*, is endemic in south-central Texas.

Usually, one or more skin lesions on the lower extremity begin as sores that slowly enlarge and ulcerate over a period of months. These untreated lesions seldom heal. Months or years later, metastatic lesions appear on the lips, nose, and pharynx. Leishmaniasis involves the larynx in approximately one-third of cases.[80]

Fever, anemia, weight loss, and hoarseness are common symptoms. As time passes, extensive soft tissue destruction may lead to grotesque facial disfiguration, as well as progression of the laryngeal disease. Examination of the larynx may reveal a localized, polypoid, inflammatory lesion or diffuse, granular, spongy mucosa throughout the larynx. The lesions may be ulcerated, and airway obstruction may occur. These lesions are often mistaken for laryngeal cancer, tuberculosis, histoplasmosis, or blastomycosis.

Biopsy reveals a chronic granulomatous pattern, with a predominance of lymphocytic and histiocytic cells. The diagnosis can be made by identification of the parasite in biopsy specimens, but, in some cases, the parasites may be difficult to find. A specific agglutination test for leishmaniasis and the leishmaniasis skin test are diagnostic. The mucocutaneous form of leishmaniasis should be treated with antimonials for at least 30 days.

Schistosomiasis Endemic in the tropics and subtropics, schistosomiasis ("bilharziasis" in Egypt) is widespread in 72 countries of the world and infects an estimated 5% of the world's population—200 million people. It is the most prevalent helminthic infection in the world.

There are three *Schistosoma* species parasitic in man: *S. mansoni, S. japonicum*, and *S. haematobium*. *Schistosoma mansoni* and *S. japonicum* inhabit the mesenteric veins; the eggs are found along the wall of the intestine and in the liver and are passed in the feces. *Schistosoma haematobium* invades the veins of the pelvic plexus, and the eggs are passed in the urine. Ectopic lesions occur in 18% of cases and may be found in any part of the body. How the parasite reaches other parts of the body, outside the gastrointestinal and urinary tracts, is unknown.

Humans contract schistosomiasis from water infested by cercariae, the microscopic infective stage. Cercariae penetrate the intact skin, and *schistosomules* form in the skin. After several days, the schistosomules migrate to the lungs and portal veins, where the male and female species mate. Weeks later, depending on the species and the time of transit,

eggs are deposited in the infected organs.

To complete the life cycle, organisms excreted by humans must contaminate water because the snail is the intermediate host. The cercariae grow in the soft tissue of the snail, and after 1 to 2 months, they are released back into the water, where they can again enter the human body, thus completing the organism's life cycle.

Schistosomiasis in the aerodigestive tract is characterized by intense inflammatory granulomas around deposits of schistosome eggs. Granulomas may be as small as a pinhead or as large as an orange. Histologically, coagulation necrosis occurs around the egg deposits, and eosinophils, plasma cells, and lymphocytes predominate in the cellular response.[81]

Patients with laryngeal schistosome granulomas present with hoarseness. The laryngeal lesion has the appearance of a pink-gray, cauliflower-like granuloma, with surrounding inflammation. Ipsilateral paralysis may be present.[81,82] The degree of surrounding fibrosis is determined by the egg density and the duration of the infection.

Diagnosis may be suspected by the histologic examination; however, confirmation of the diagnosis is made by identification of parasitic ova in the urine or feces. Treatment is with antihelminthic medication (praziquantel or oxamniquine).

Syngamosis *Syngamus laryngeus* is a unique gapeworm indigenous to Brazil, Puerto Rico, Martinique, Trinidad, British Guiana, the West Indies, and the Philippines. It invades the upper respiratory tract of cattle, water buffalo, and (very rarely) humans. Transmission to humans is believed to be through consumption of contaminated vegetables. Once ingested, the adult worms migrate to the larynx and upper trachea and firmly attach themselves to those mucosal surfaces.

The primary symptoms of syngamosis are cough, a foreign body sensation in the throat, and, occasionally, hemoptysis. The diagnosis may be suspected if the patient coughs up worms in copula or if they are seen on laryngoscopy. Removal of the worms by direct laryngoscopy is the only known treatment.[83]

Inhalation Laryngitis

When nebulized, radiolabeled, acidic fog is inhaled and scanned, the density of aerosol deposit in the larynx is greater than in any other site in the aerodigestive tract.[84] The size and anatomic configuration of the larynx (having the narrowest and most convoluted lumen of the upper airway) may explain this phenomenon. Perhaps for this reason, the larynx is especially susceptible to the effects of tobacco smoke, dust, and other airborne environmental contaminants. Table 52–11 lists some of the commonly reported substances associated with acute and chronic inhalation injuries of the larynx.

Acute (Thermal) Injury Thermal and toxic injuries of the airway are often seen in patients with facial burns. In this regard, thermal injury usually plays a greater role than chemical injury in the larynx. Intense heat produces excessive vasodilation, capillary permeability, massive edema, and, thus, airway obstruction. The edema usually peaks at 8 to 24 hours following injury and resolves within 4 to 5 days.

The severity of permanent damage depends on the severity of the burn. Complete laryngeal stenosis can occur. Treatment consists of providing humidified oxygen and maintaining the airway. In most cases, corticosteroids and antibiotics should be avoided. Weeks after the acute injury has resolved, laryngeal dilation and surgical procedures to re-establish the airway, if necessary, may be performed.

Pollution and Inhalant Allergy The common clinical manifestations of allergy to inhalants are well known to otolaryngologists: sneezing, watery rhinorrhea, nasal congestion, and obstruction. The most common offending allergens are dusts, pollens, molds, and chemicals. Allergic reactions of the larynx, similar to those seen in the nose, are uncommon but do occur. The diagnosis is made by the history, clinical findings, and intradermal skin testing. Treatment should be removal of the offending allergen(s), if possible, and desensitization. Corticosteroid inhalers are ineffective in laryngeal allergy and should be avoided.

Carcinogens Tobacco and several other carcinogens (see Table 52–11) cause chronic inflammation of the larynx. These substances can cause mucosal thickening, submucosal edema, hyperkeratosis, dysplasia, and, eventually, carcinoma. In the larynx, leukoplakia, pachydermia, and Reinke's edema (polypoid degeneration) should be viewed as precursors to the development of carcinoma.

TABLE 52–11. Causes of Inhalation Laryngitis (Substances Associated with Acute and Chronic Laryngeal Inflammation)

Acute inhalation injury
Steam
Hot dry gases
Smoke
Common allergens
Dusts
Animal danders
Formaldehyde
(Plus the list of pollutants below)
Pollutants
Dusts
Ozone
Ammonia
Chlorine
Nitrous oxide
Hydrogen disulfide
Sulfur dioxide
Carbon monoxide
Sulfuric acid
Hydrochloric acid
Kerosene (heaters)
Insecticides
Pesticides
Known carcinogens
Tobacco
Asbestos
Radon
Nickel
Sulfuric acid
Isopropyl oils
Mustard gas

Leukoplakia, which means "white plaque," is histologically hyperkeratosis. This is a typical metaplastic response of epithelium to chronic irritation, regardless of the cause. When the laryngeal mucosa becomes diffusely thickened (and also usually diffusely hyperkeratotic) to the point of narrowing the lumen of the endolarynx, the condition is termed *pachydermia laryngis*. Pachydermia ("elephant skin") is frequently found in the larynges of patients who have smoked cigarettes for many years.

Reinke's edema also results from chronic laryngeal irritation over a period of many years. It is almost always bilateral, and although it occurs most frequently in elderly female smokers, it is also seen in nonsmoking patients with reflux disease and hypothyroidism.

Leukoplakia, pachydermia, and Reinke's edema may improve with cessation of smoking and antireflux therapy (when the patient has reflux disease), but often patients with these lesions require surgical treatment.

Radiation Injury

Radiation Laryngitis Radiation therapy for laryngeal carcinoma, as well as for tumors in other head and neck sites, may deliver significant radiation doses to normal laryngeal tissue. The initial effects produce an intense inflammatory response, characterized by increased capillary permeability, edema, neutrophilic infiltration, vascular thrombosis, and obliteration of lymphatic channels.

Patients undergoing radiation treatment complain of a globus sensation, dysphagia, odynophagia, dysphonia, and odynophonia. On examination, the larynx appears red and swollen, and a fibrinous exudate may be present. The symptoms tend to worsen as the treatment progresses; they are worst at the completion of treatment and gradually abate thereafter.[85]

Late tissue sequelae consist of degenerative changes and fibrosis in adipose, connective, and glandular tissues and a pronounced obliterative endarteritis of small blood vessels. These changes may take place over a period of years, and the symptoms of many patients worsen with time.

Atrophy of glandular elements of the larynx leads to *laryngitis sicca*, a dry larynx. This, in turn, increases the susceptibility of the larynx to infection and to damage from LPR. Patients often complain about the effects of dryness, including hoarseness, crusting, globus, and cough. Treatment is symptomatic, consisting of hydration, administration of expectorants, and environmental humidification.

Radionecrosis Like laryngitis sicca, radionecrosis is a late complication of laryngeal irradiation, being the result of ischemia of the cartilaginous framework. Patients with laryngeal radionecrosis usually present with dysphagia, weight loss, *fetor oris* (foul breath), severe dysphonia, impaired vocal fold mobility, and relative airway obstruction; the last three depend on the amount and location of swelling. Massive edema of the larynx with airway obstruction is not uncommon.

When a patient presents after radiation therapy with increasing symptoms and laryngeal edema, the clinician must determine the cause or causes. The differential diagnosis of laryngeal radionecrosis also includes tumor recurrence and reflux. Known as "the three R's" (radionecrosis, recurrence, reflux), which of these has caused the patient's increasing symptoms will determine the treatment. Each patient must be evaluated for each of the three possibilities.

Computed tomographic scanning of the larynx may reveal tumor recurrence; however, unless there is relatively massive destruction, scans are notoriously unreliable for evaluating the laryngeal cartilages. Tumor recurrence is usually treated by surgical salvage, and reflux is usually treated with PPI.

Treatment for radionecrosis should be individualized. In some cases, necrotic cartilage and soft tissue can be removed through an endoscopic approach. Often, however, an open surgical procedure is needed. The larynx may then be stented for 2 to 4 weeks. If the necrotic sequestration has been adequately removed, the larynx may heal in a satisfactory manner, even if much of the cartilaginous framework is resected. When available, hyperbaric oxygen treatment may be of significant benefit, completely obviating surgery in selected cases.[86] Laryngectomy should be a last resort and should be considered only when all conservative measures have failed.

Vocal Abuse and Misuse Syndromes

Vocal Fold Hemorrhage Screaming and other forms of vocal abuse can lead to acute submucosal hemorrhage of the vocal folds. By history, such hemorrhage occurs abruptly and produces severe dysphonia. On examination, the appearance of hematoma is unmistakable. Voice rest is the usual treatment, although some laryngologists recommend surgical drainage in selected cases.

Muscle Tension Dysphonias *Muscle tension dysphonia* is a generic term for any "functional" voice disorder caused by chronic vocal abuse or misuse. Patients with vocal nodules, contact ulcers, and granulomas or the Bogart-Bacall syndrome all fit into this category, as do patients with psychogenic voice disorders.[87,88] We classify laryngeal muscle tension into primary and secondary categories. Primary muscle tension dysphonia is owing to nonorganic vocal fold pathology. The behavior may be learned or entirely psychogenic. Secondary, or compensatory, muscle tension dysphonia is a result of an underlying glottal insufficiency. Among a community cohort (N = 100), persons with vocal fold bowing were 17 times more likely to display abnormal muscle tension patterns (MTPs) (predominantly MTP II and III in Figure 52–8) than individuals without vocal fold bowing ($p < .001$).[88,89] The glottal incompetence may be owing to presbylaryngis, vocal fold paralysis, or vocal fold paresis. These findings suggest that supraglottic hyperfunction may be a physiologic attempt by the glottis to compensate for underlying glottal closure problems. Muscle tension dysphonias are common in professional voice users and are frequently initiated or complicated by reflux

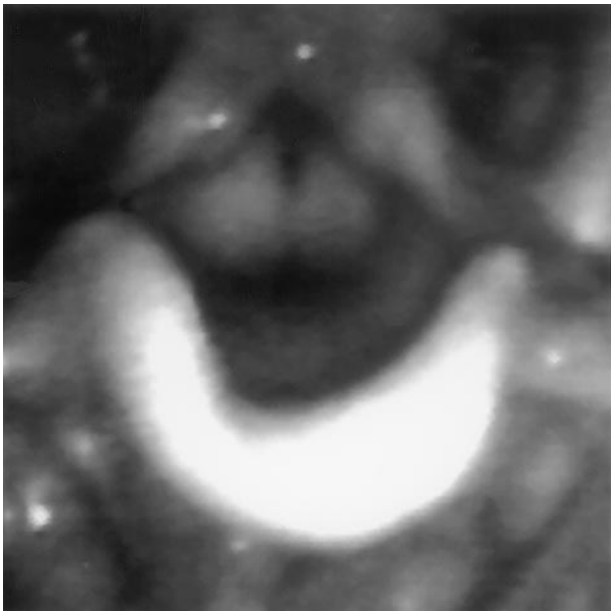


FIGURE 52–8. Muscle tension pattern II—lateral to medial approximation of the false vocal folds. The true vocal folds are completely obscured. Muscle tension pattern III—anterior to posterior contraction of the larynx. The anterior third of the true vocal folds is obscured.

laryngitis. Treatment with voice therapy, surgery, and, when appropriate, antireflux therapy often successfully resolves the problem.

Contact Ulcers and Granulomas Anatomically, only a thin layer of mucosa and perichondrium overlies the vocal processes of the arytenoid cartilages, so ulceration over a vocal process can occur from a variety of insults, including vocal abuse, coughing, viral infection, GER, and endotracheal intubation.[90–92] Because of the shape of the larynx, endotracheal tubes reside in the posterior commissure and thus have contact with the vocal processes. Examination of 99 postintubation laryngeal specimens obtained at autopsy showed that those intubated for less than 12 hours demonstrated pale, oval areas on the mucosa of the vocal processes, which under microscopic examination revealed complete or focal loss of surface epithelium. Those intubated for 12 to 48 hours almost uniformly demonstrated ulcerations over the vocal processes.[93]

Granulomas over the vocal process are five times more common than ulcers. Chronic ulcers over the vocal processes are uncommon because most either heal or go on to form granulomas.[94] The most common sequence of granuloma formation is as follows. First, the mucosa overlying the vocal process is acutely damaged. Then, owing to continued trauma (eg, throat clearing, cough, or LPR), ulceration may occur. Next, as the ulcer attempts to heal, secondary infection and/or reflux perpetuates the inflammation, and a granuloma forms (Figure 52–9).

As recently as 25 years ago, contact ulcers and granulomas were believed to be exclusively caused by vocal abuse; consequently, prolonged periods of voice rest were prescribed. Such treatment was, however, usually unsuccessful. Today, regardless of the inciting cause, patients with these lesions should initially be treated with "voice modification" (not voice rest), to reduce continual vocal process trauma, and antireflux therapy (preferably with a PPI). This regimen will result in healing in the majority of cases within 6 months.

Although surgical removal of a vocal process granuloma is seldom indicated, there are three instances when it should be considered: (1) when the clinician is concerned about the possibility of carcinoma, (2) when the lesion has matured and taken on the appearance of a fibroepithelial polyp, and (3) when the airway is obstructed. The carbon dioxide surgical laser is useful in removing granulomas.[94] Finally, in carefully selected cases, if traumatic vocal behaviors persist and cannot be corrected, a "one-time" botulinum toxin injection may be used "to put the larynx to rest" long enough for recalcitrant granulomas to heal.

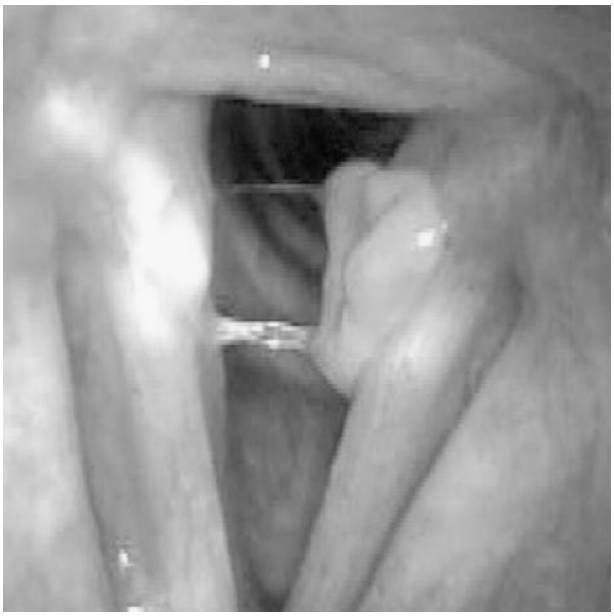


FIGURE 52–9. Granuloma of the left vocal process.

REFERENCES

1. Castell DO, Wu WC, Ott DJ, editors. Gastroesophageal reflux disease, pathogenesis, diagnosis, therapy. Mt. Kisko (NY): Futura; 1985.
2. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD). Laryngoscope 1991;101 Suppl 53:1–78.
3. Koufman JA, Amin MR, Panetti M. Prevalence of reflux in 113 consecutive patients with laryngeal and voice disorders. Otolaryngol Head Neck Surg 2000; 123:385–8.
4. Bain WM, Harrington JW, Thomas LE, et al. Head and neck manifestations of gastroesophageal reflux. Laryngoscope 1983;93:175–9.
5. Little FB, Koufman JA, Kohut RI, et al. Effect of gastric acid on the pathogenesis of subglottic stenosis. Ann Otol Rhinol Laryngol 1985;94:516–9.
6. Ossakow SJ, Elta G, Colturi T, et al. Esophageal reflux and dysmotility as the basis for persistent cervical symptoms. Ann Otol Rhinol Laryngol 1987;96:387–92.

7. Ward PH, Hanson DG. Reflux as an etiological factor of carcinoma of the laryngopharynx. Laryngoscope 1988;98:1195–9.
8. Morrison MD. Is chronic gastroesophageal reflux a causative factor in glottic carcinoma? Otolaryngol Head Neck Surg 1988;99:370–3.
9. Burton DM, Pransky SM, Katz RM, et al. Pediatric airway manifestations of gastroesophageal reflux. Ann Otol Rhinol Laryngol 1992;101:742–9.
10. Jindal JR, Milbrath MM, Shaker R, et al. Gastroesophageal reflux as a likely cause of "idiopathic" subglottic stenosis. Ann Otol Rhinol Laryngol 1994; 103:186–91.
11. Wiener GJ, Koufman JA, Wu WC, et al. Chronic hoarseness secondary to gastroesophageal reflux disease: documentation with 24-H ambulatory pH monitoring. Am J Gastroenterol 1989;84:1503–8.
12. Belafsky P, Postma G, Koufman J. Laryngopharyngeal reflux symptoms improve before changes in physical findings. Laryngoscope 2001;11:979–81.
13. Belafsky P, Postma G, Koufman J. The validity and reliability of the reflux finding score (RFS). Laryngoscope 2001;111:1313–7.
14. Richter JE, editor. Ambulatory esophageal pH monitoring: practical approach and clinical applications. Tokyo: Igaku-Shoin; 1991.
15. Frank M, Komisar A. Ambulatory pH monitoring in the management of reflux. Ann Otol Rhinol Laryngol 1993;102:243–6.
16. Johnson P, Belafsky P, Koufman J, Postma G. Ambulatory double-probe pH monitoring: the importance of manometry [abstract]. The Triological Society meeting, 2001.
17. Richter JE, Bradley LA, DeMeester TR, et al. Normal 24-hour pH values: influence of study center, pH electrode, age, and gender. Dig Dis Sci 1992;37:849–56.
18. Johnson P, Postma G, Koufman J, Belafsky P. pH monitoring in patients with laryngopharyngeal reflux: why the pharyngeal probe is essential [abstract]. Combined Otolaryngology Spring Meeting, 2001.
19. Ott DJ, Cowan RJ, Gelfand DW, et al. The role of diagnostic imaging in evaluating gastroesophageal reflux disease. Postgrad Radiol 1986;6:3–14.
20. Kamel PL, Hanson D, Kahrilas PJ. PPI for the treatment of posterior laryngitis. Am J Med 1994;96:321–6.
21. Wolfe MM, Soll AH. The physiology of gastric acid secretion. N Engl J Med 1988;319:1707–15.
22. Bough ID, Sataloff RT, Castell DO, et al. Gastroesophageal reflux laryngitis resistant to PPI therapy. J Voice 1995;9:205–11.
23. Amin MR, Postma G, Johnson P, et al. Proton pump inhibitor resistance in the treatment of laryngopharyngeal reflux. Otolaryngol Head Neck Surg 2001; 125:374–8.
24. Loughlin CJ, Koufman JA. Paroxysmal laryngospasm secondary to gastroesophageal reflux. Laryngoscope 1996;106:1502–5.
25. Contencin P, Narcy P. Gastropharyngeal reflux in infants and children. A pharyngeal pH monitoring study. Arch Otolaryngol Head Neck Surg 1992; 118:1028–30.
26. Cressman WR, Meyer CM. Diagnosis and management of croup and epiglottitis. Pediatr Clin North Am 1994;41:265–76.
27. Heath PT, Booy R, Azzopardi HJ, et al. Non-type b *Haemophilus influenzae* disease: clinical and epidemiologic characteristics in the *Haemophilus influenzae* type b vaccine era. Pediatr Infect Dis J 2001;20:300–5
28. Crockett DM, Healy GB, McGill TJ, Friedman EM. Airway management of acute supraglottitis at the Children's Hospital, Boston: 1980–1985. Ann Otol Rhinol Laryngol 1988;97:114–9.
29. Cantrell RW, Bell RA, Morioka WT. Acute epiglottitis: intubation versus tracheotomy. Laryngoscope 1978;88:994-1005.
30. Havaldar PV. Diphtheria in the eighties: experience in a south Indian district hospital. J Indian Med Assoc 1992;90:155–6.
31. Deeb ZE, Einhorn KH. Infectious adult croup. Laryngoscope 1990;100:455–7.
32. Nadel S, Offit PA, Hodinka RL, et al. Upper airway obstruction in association with perinatally acquired herpes simplex virus. J Pediatr 1992;120:127–9.
33. D'Angelo AJ, Zwillenberg S, Olezszyk, et al. Adult supraglottis due to herpes simplex virus. J Otolaryngol 1990;19:179–81.
34. Solomon P, Weisbrod M, Irish JC, Gullane PJ. Adult epiglottitis: the Toronto Hospital experience. J Otolaryngol 1998;27:332–6.
35. Murrage KJ, Janzen VD, Ruby RR. Epiglottitis: adult and pediatric comparisons. J Otolaryngol 1988; 17:194–8.
36. Deeb ZE, Yenson AC, DeFries HO. Acute epiglottitis in the adult. Laryngoscope 1985;95:289–91.
37. Souliere CR, Kirchner JA. Laryngeal perichondritis and abscess. Arch Otolaryngol 1985;111:481–4.
38. Ramadan HH, Tarazi AE, Baroudy FM. Laryngeal tuberculosis: presentation of 16 cases and a review of the literature. J Otolaryngol 1993;22:39–41.

39. Harney M, Hone S, Timon C, Donnelly M. Laryngeal tuberculosis: an important diagnosis. J Laryngol Otol 2000;114:878–80.
40. Soda A, Rubio H, Salazar M, et al. Tuberculosis of the larynx: clinical aspects in 19 patients. Laryngoscope 1989;99:1147–50.
41. Bailey CM, Windle-Taylor PC. Tuberculosis laryngitis: a series of 37 patients. Laryngoscope 1981;91:93–100.
42. Soni NK. Leprosy of the larynx. J Laryngol Otol. 1992; 106:518–20.
43. Younus M. Leprosy in ENT. J Laryngol Otol 1986;100: 1437–42.
44. Flanagan PM, McIlwain JC. Tuberculosis of the larynx in a lepromatous patient. J Laryngol Otol 1993; 107:845–7.
45. Holinger PH, Gelman HK, Wolfe CK. Rhinoscleroma of the lower respiratory tract. Laryngoscope 1977;87:1–9.
46. Stiernberg CM, Clark WD. Rhinoscleroma—a diagnostic challenge. Laryngoscope 1983;93:866–70.
47. Jay J, Green RP, Lucente FE. Isolated laryngeal rhinoscleroma. Otolaryngol Head Neck Surg 1985;93:669–73.
48. Alfaro-Monge JM, Fernandez-Espinosa J, Soda-Morhy A. Scleroma of the lower respiratory tract: case report and review of the literature. J Laryngol Otol 1994;108:161–3.
49. Abou El-Hamd KA. Scleroma of the larynx and associated squamous cell carcinoma: a causal relationship? Ann Otol Rhinol Laryngol 2000;109:306–10.
50. Nelson EG, Tybor AG. Actinomycosis of the larynx. Ear Nose Throat J 1992;71:356–8.
51. Everett ED, Templer JW. Oropharyngeal tularemia. Arch Otolaryngol 1980;106:237–8.
52. Wilkinson L. Glanders: medicine and veterinary medicine in common pursuit of a contagious disease. Med Hist 1981;25:363–84.
53. Musher DM. Syphilis, neurosyphilis, penicillin, and AIDS. J Infect Dis 1991;163:1201–6.
54. McNulty JS, Fassett RL. Syphilis: an otolaryngologic perspective. Laryngoscope 1981;91:889–905.
55. Burton DM, Seid AB, Kearns DB, Pransky SM. Candida laryngotracheitis: a complication of combined steroid and antibiotic usage in croup. Int J Pediatr Otorhinolaryngol 1992;23:171–5.
56. Vrabec DP. Fungal infections of the larynx. Otolaryngol Clin North Am 1993;26:1091–14.
57. Tashjian LS, Peacock JE. Laryngeal candidiasis. Report of seven cases and review of the literature. Arch Otolaryngol 1984;110:806–9.
58. Reder PA, Neel HB. Blastomycosis in otolaryngology: review of a large series. Laryngoscope 1993;103:53–8.
59. Lester CF, Conrad FG, Atwell RJ. Primary laryngeal blastomycosis. Review of the literature and presentation of a case. Am J Med 1958;24:305–9.
60. Boyle JO, Coulthard SW, Mandel RM. Laryngeal involvement in disseminated coccidioidomycosis. Arch Otolaryngol Head Neck Surg 1991;117:433–8.
61. Ward PH, Berci G, Morledge D, Schwartz H. Coccidioidomycosis of the larynx in infants and adults. Ann Otol Rhinol Laryngol 1977;86:655–60.
62. Bolivar R, Gomez LG, Luna M, et al. *Aspergillus* epiglottis. Cancer 1983;51:367–70.
63. Remington PL, Vergeront JM, Stoebig JF, et al. Sporotrichosis in Wisconsin. Wis Med J 1983;82:25–7.
64. Agger WA, Seager GM. Granulomas of the vocal folds caused by *Sporothrix schenckii*. Laryngoscope 1985;95:595–6.
65. Neel HB, McDonald TJ. Laryngeal sarcoidosis: report of 13 patients. Ann Otol Rhinol Laryngol 1982;91:359–62.
66. Gallivan GJ, Landis JN. Sarcoidosis of the larynx: preserving and restoring airway and professional voice. J Voice 1993;7:81–94.
67. Jorizzo JL, Koufman JA, Thompson JN, et al. Sarcoidosis of the upper respiratory tract in patients with nasal rim lesions: a pilot study. J Am Acad Dermatol 1990;22:439–43.
68. Bower JS, Belen JE, Weg JG, Dantzker DR. Manifestations and treatment of laryngeal sarcoidosis. Am Rev Respir Dis 1980;122:325–32.
69. Lebovics RS, Hoffman GS, Leavitt RY, et al. The management of subglottic stenosis in patients with Wegener's granulomatosis. Laryngoscope 1992;102: 1341–5.
70. Werber JL, Pincus RL. Oropharyngeal angioedema associated with the use of angiotensin-converting enzyme inhibitors. Otolaryngol Head Neck Surg 1989;101:96–8.
71. Wahle D, Beste DJ, Conley SF. Laryngeal involvement in toxic epidermal necrolysis. Otolaryngol Head Neck Surg 1992;107:796–9.
72. Campbell SM, Montanaro A, Bardana EJ. Head and neck manifestations of autoimmune disease. Am J Otolaryngol 1983;4:187–216.
73. Jurik AG, Pedersen U, Nrgård A. Rheumatoid arthritis of the cricoarytenoid joints: a case of laryngeal obstruction due to acute and chronic joint changes. Laryngoscope 1985;95:846–8.

74. Friedman BA, Rice DH. Rheumatoid nodules of the larynx. Arch Otolaryngol 1975;101:361–3.
75. Hanson RD, Olsen KD, Rogers RS. Upper aerodigestive tract manifestations of cicatricial pemphigoid. Ann Otol Rhinol Laryngol 1988;97:493–9.
76. McAdam LP, O'Hanlan MA, Bluestone C, Pearson CM. Relapsing polychrondritis: prospective study of 23 patients and a review of the literature. Medicine 1976;55:193–215.
77. Barrs DM, McDonald TJ, Duffy J. Sjögren's syndrome involving the larynx. Report of a case. J Laryngol Otol 1979;93:933–6.
78. Gertz MA, Kyle RA. Primary systemic amyloidosis—a diagnostic primer. Mayo Clin Proc 1989;64: 1505–19.
79. Lewis JE, Olsen KD, Kurtin PJ, Kyle RA. Laryngeal amyloidosis: a clinicopathologic and immunohistochemical review. Otolaryngol Head Neck Surg 1992;106:372–7.
80. Marsden PD. Mucosal leishmaniasis ("espundia" Escomel, 1911). Trans R Soc Trop Med Hyg 1986; 80:859–76.
81. Toppozada HH. Laryngeal bilharzia. J Laryngol Otol 1985;99:1039–41.
82. Manni HJ, Lema PN, van Raalte JA, Westerbeek GF. Schistosomiasis in otorhinolaryngology: review of the literature and case report. J Laryngol Otol 1983; 97:1177–81.
83. Weinstein L, Molavi A. *Syngamus laryngeus* infection (syngamosis) with chronic cough. Ann Intern Med 1971;74:577–80.
84. Bowes SM, Laube BL, Links JM, Frank R. Regional disposition of inhaled fog droplets: preliminary observations. Environ Health Perspect 1989;79: 151–7.
85. Chandler JR. Radiation fibrosis and necrosis of the larynx. Ann Otol Rhinol Laryngol 1979;88:509–14.
86. Ferguson BJ, Hudson WR, Farmer JC. Hyperbaric oxygen therapy for laryngeal radionecrosis. Ann Otol Rhinol Laryngol 1987;96:1–6.
87. Koufman JA, Blalock PD. Vocal fatigue and dysphonia in the professional voice user: Bogart-Bacall syndrome. Laryngoscope 1988;98:493–8.
88. Koufman JA, Blalock PD. Functional voice disorders. Otolaryngol Clin North Am 1991;24:1059–73.
89. Belafsky PC, Postma GN, Reulbach TR, et al. Muscle tension dysphonia as a sign of underlying glottal insufficiency [abstract]. Annual Meeting of the American Academy of Otolaryngology-Head and Neck Surgery, Denver, CO, 2001.
90. Feder RJ, Mitchell MJ. Hyperfunctional, hyperacidic, and intubation granulomas. Arch Otolaryngol 1984; 110:582–4.
91. Cherry J, Margulies SI. Contact ulcer of the larynx. Laryngoscope 1968;78:1937–40.
92. Delahunty JE, Cherry J. Experimentally produced vocal fold granulomas. Laryngoscope 1968;78:1941–7.
93. Donnelly WH. Histopathology of endotracheal intubation. An autopsy study of 99 cases. Arch Pathol 1969;88:511–20.
94. Koufman JA. Contact ulcer and granuloma of the larynx. In: Gates GA, editor. Current therapy in otolaryngology-head and neck surgery. 5th ed. St. Louis: Mosby; 1994. p. 456–9.

SUGGESTED READING

Baugh R, Gilmore BB. Infectious croup: a critical review. Otolaryngol Head Neck Surg 1986;95:40–6.

Caldarelli DD, Friedberg SA, Harris AA. Medical and surgical aspects of the granulomatous diseases of the larynx. Otolaryngol Clin North Am 1979;12: 767–81.

Cann CI, Fried MP, Rothman KJ. Epidemiology of squamous cell cancer of the head and neck. Otolaryngol Clin North Am 1985;18:367–88.

Klinkenberg-Knol EC. Recent advances in the diagnosis and management of gastro-oesophageal reflux disease: the role of PPI in clinical practice. Amsterdam: VU University Press; 1990.

Lucente FE. Impact of the acquired immunodeficiency syndrome epidemic on the practice of laryngology. Ann Otol Rhinol Laryngol 1993;102 Suppl 161:1–24.

Pillsbury HC, Sasaki CT. Granulomatous diseases of the larynx. Otolaryngol Clin North Am 1982;15:539–51.

Richtsmeier WJ, Johns ME. Bacterial causes of granulomatous disease. Otolaryngol Clin North Am 1982; 15:473–92.

Sataloff RT. The impact of pollution on the voice. Otolaryngol Head Neck Surg 1992;106:701–5.

CHAPTER 53

Neurogenic and Functional Disorders of the Larynx

Christy L. Ludlow, PhD, Eric A. Mann, MD, PhD

Normal speech requires the complex integration of peripheral and central motor control mechanisms. The voice component of speech requires (1) neural control of the intrinsic and extrinsic laryngeal muscles to shape the glottis and (2) a steady stream of airflow from the respiratory system to support regular, symmetric, and synchronous vibration of the vocal folds. Thus, the central nervous system must coordinate the respiratory, phonatory, and speech mechanisms to produce normal phonation during speech.

Effects of Speech on Laryngeal View during Endoscopy

The larynx does not function in isolation from other structures in the vocal tract. Muscles shaping the pharynx and oral cavity, particularly those controlling the tongue, lip, and jaw, alter laryngeal tension as they change vocal tract posture. In normal speech, the vocal tract moves rapidly between different target shapes for specific speech sounds.

Valving at the glottis, velum, tongue, lips, and jaw amplifies certain harmonics of the fundamental frequency of vocal fold vibration, based on the posture and shape of the vocal tract. Depending on the shape of the tract, harmonics in the region of certain resonant frequencies are amplified, resulting in sound energy in formants. The frequency regions of the formants are responsible for the vowel sounds. A simple example is illustrated in Figure 53–1. Sagittal sections show specific vocal tract configurations for the vowels /i/ (labeled "iy"), /u/ (labeled "uw"), /æ/ (labeled "ae"), and /a/ (labeled "aa"). In high-frequency vowel sounds such as /i/ or "ee," the posture of the tongue is high and forward, exerting tension from the tip to the base of the tongue. Coupling between the hyoid bone and suspended thyroid cartilage distributes tension to extrinsic and intrinsic laryngeal musculature that may tilt the thyroid cartilage. During /i/, the pharyngeal cavity is large, making conditions for airflow different than for production of /a/ or "ah," which is characterized by a constricted pharyngeal cavity because of the low and posterior placement of the tongue.

These differences in pharyngeal constriction also affect videoendoscopy; the vowel /i/ allows a better view of the larynx because the tongue is forward and high, opening the pharynx and bringing the epiglottis forward. In contrast, during /a/, the larynx is much less visible because the tongue is back; consequently, the epiglottis covers the larynx, obstructing the view.

Multidimensional Aspects of Voice Function and Disorders

Pharyngeal and/or laryngeal constriction can produce turbulence in airflow in the case of neurogenic disorders. Conversely, reduced vocal fold adduction can produce breathiness because of air escaping and not being modulated to produce voice. The acoustic product is thus affected by changes in constriction and air pressures in the pharynx and oral cavity. Both peripheral and central neurogenic disorders, therefore, may alter the voice by changing the configuration of the larynx and vocal tract. In functional voice disorders, changes in vocal tract tension can also alter voice to a similar degree as in neurogenic voice disorders. Abnormalities in vocal tract tension owing to voluntary or involuntary muscle contractions occur in neurogenic and/or functional disorders. The origins may by traumatic, neoplastic, vascular, infectious, degenerative, or idiopathic. They are often present in conjunction with disease processes that alter not only the laryngeal musculature but also lip, tongue, and jaw control and other motor systems such as respiration.

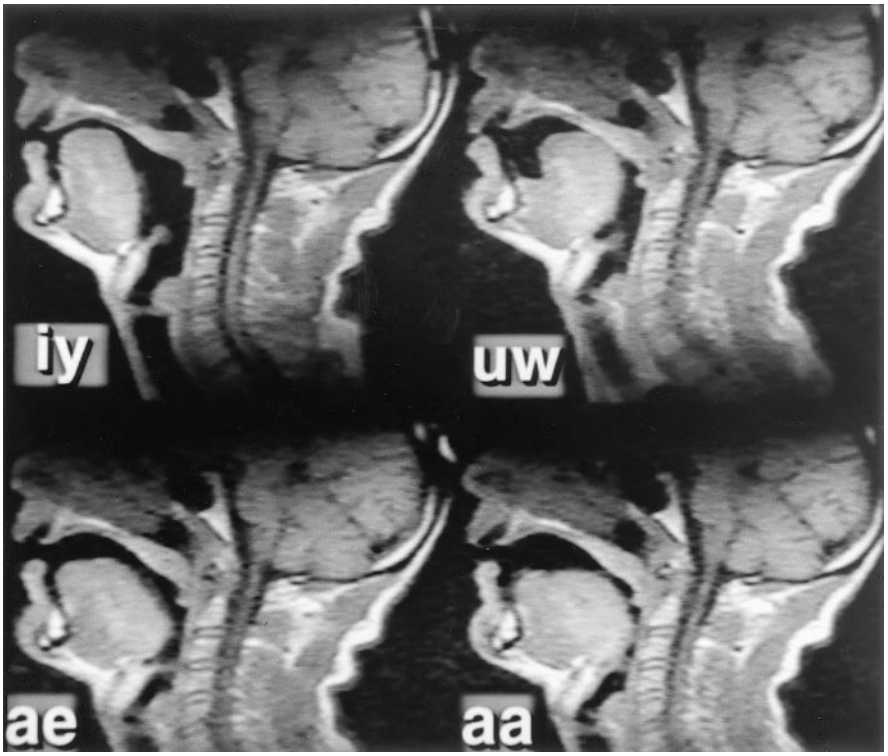


FIGURE 53–1. Sagittal sections from magnetic resonance images of the vocal tract illustrate vocal tract shape during production of four vowels. Adapted from Baer T, Gore JC, Gracco LC, Nye PW. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. J Acoust Soc Am 1991;90:799–828.

In addition to the degree of motor impairment, the diagnostic process must also take into account the body's capacity for reorganization and compensation that may alter the perceived severity of speech disturbance. Perceptual voice parameters such as pitch, loudness, voice quality, and their variability may help to characterize neurogenic and functional voice disorders. Quantification of acoustic and perceptual characteristics of voice and speech may be useful based on a variety of instrumental techniques.[1,2] Computer-assisted quantification of the acoustic product may be a valuable adjunct to the perceptual parameterization of voice; however, the two are not synonymous.[3] Measures based on pressure and flow transduction during phonation provide information about the functional relationship between the respiratory and phonatory systems. Further, the use of diagnostic imaging techniques, including rigid and flexible fiber-optic videoendoscopy, allows observation of laryngeal and pharyngeal movement. Flexible fiber-optic examination is required when the voice disorder occurs more during connected speech, such as in spasmodic dysphonia (SD). Stroboscopic examination is useful in assessing the vibratory characteristics of the vocal folds.[4] High-speed photography, ultrasonography, videofluoroscopy, and magnetic resonance imaging are other imaging tools that have value in the diagnostic protocol.[5,6] Use of electromyography (EMG) accompanied by magnetic stimulation of the laryngeal nerve may be valuable in differentiating peripheral and central neurogenic dysfunction.[7] The goal of instrumental and perceptual techniques has been to characterize vocal fold movement control in terms of timing, speed, and accuracy of movement of vocal tract structures and to quantify these parameters for the documentation of the status and progression of disease for therapeutic purposes.

NEUROGENIC DISORDERS

Dysphonia can appear as the first sign of neurogenic disease, whereas the other speech symptoms (dysarthria) may become evident only as the disease progresses. Therefore, caution should be used in the treatment of neurogenic voice symptoms until it can be ensured that the laryngeal disorder is focal and not part of a progressive neurodegenerative disease. For example, what may initially appear to be a unilateral idiopathic vocal fold paralysis might later progress to become a bilateral vocal fold paralysis as part of a peripheral neuropathy[8] or multiple systems atrophy.[9] A unilateral thyroplasty could be detrimental if bilateral vocal fold paraly-

sis developed subsequently resulting in airway compromise. For this reason, a complete neurologic examination is important before planning intervention in these disorders.

Functionally, neurogenic disorders resulting in dysphonia may be divided into the following categories:

1. Consistent neurogenic voice disorders are characterized by constant vocal quality, loudness, or pitch deviations during speech and sustained vowels. Examples include flaccid dysarthria owing to lower motoneuron disease loss, a peripheral neuropathy, or laryngeal nerve injury. These can produce paralysis or paresis of the adductor and/or abductor muscles, causing asymmetries in movement. On the other hand, tension in the vocal folds may be altered when either the cricothyroid muscle or the thyroarytenoid muscles are affected as both contribute to vocal fold tension.[10]
2. Spastic dysarthria, including dysphonia, is associated with upper motoneuron disease involving the corticobulbar tracts. For example, hypertension of the vocal tract, including the vocal folds, may be seen in multiple sclerosis (MS).
3. Arrhythmically fluctuating neurogenic voice disorders are most frequent. These are characterized by unpredictable, irregular variations in quality, loudness, and pitch during speech. Sustained vowel production may accentuate irregularities in some cases. Ataxic, choreic, and dystonic dysphonias, including the SDs, display this type of irregularity.
4. Rhythmically fluctuating neurogenic voice disorders include palatopharyngolaryngeal myoclonus and essential voice tremor. These dysphonias are marked by regular or rhythmic fluctuations in voice, pitch, and loudness. Paroxysmal neurogenic voice disorders exhibit bursts of dysphonic voice, as in Tourette's syndrome.
5. Neurogenic voice disorders associated with loss of volitional control of voice production, including apraxia of phonation, respiration, or speech and akinetic mutism, usually follow a cerebrovascular accident or cortical injury.

Upper Motoneuron Disorders

Upper motoneuron diseases with characteristic voice components include Parkinson's disease and related syndromes. Parkinsonism is a pathophysiologic state reflecting failure of the nigrostriatal dopaminergic neuronal system owing to degeneration or destruction of the substantia nigra. This can occur by depleting neuronal dopamine, inhibiting its synthesis, or blocking the striatal dopamine receptors.[11] The basal ganglia modulate cortical control via the thalamus. Neural activity at the cortex, in turn, controls the lower motoneurons. Hence, damage to the basal ganglia can release inhibition of nerve impulses affecting the lower motoneurons, resulting in rigidity and reduced rate of movement (bradykinesia).

Hypophonia in Parkinson's Disease Although parkinsonism can be described as a syndrome of diverse origins, distinct syndromes are recognized from clinical features and secondary or associated manifestations, although differentiation is often difficult. Parkinson's disease can occasionally be familial, and genetic predisposition can be both mendelian or a complex interaction of a genetic predisposition with environmental factors.[12] Onset is typically in the sixties or later, but the disease may appear as early as the middle thirties. Rare cases have been reported as early as 15 years of age.[11]

Reduced loudness and breathy vocal quality, referred to as hypophonia, are the hallmark of voice disorders in early Parkinson's disease. An essential feature is a voice quality that fades into breathiness in contextual speech. The patient may have difficulty with production of glottal stops or voice onset after voiceless consonants such as /s/. For this reason, persons with early onset of Parkinson's disorder may have voice symptoms similar to abductor SD.[13] There may be reduced range and speed of vocal fold movement, particularly on the more affected side. Distinctions between voiceless and voiced sounds become reduced owing to impaired ability to produce glottal stops to mark word boundaries. Patients tend to maintain sounds as all voiced or all voiceless.[14] In later stages of the disease, the patient may be unable to produce phonation even with instruction. In advanced disease, severe "on-off" drug-related phenomena may occur, during which patients may experience breathy voice followed by periods of propulsive speech with a strained quality. Vocal fold imaging studies in Parkinson's disease often illustrate a glottic gap and bowing of the vocal folds[13] associated with breathiness and reduced loudness. Asymmetries have also been demonstrated

in vocal fold closure, thinning of the mucosa, and vocal fold bowing.

Patients can increase voice intensity on demand,[15] and voice therapy aimed at increasing vocal intensity can improve speech intelligibility and voice when combined with dopaminergic enhancement therapy.[16] Focus on upper articulator (lip, tongue, and jaw) movement produces little in the way of increased intelligibility for this population. Recently, percutaneous collagen augmentation of the vocal folds has been used to improve the breathy hypophonia of Parkinson's disease in persons with vocal fold bowing and glottic insufficiency.[17] Collagen therapy does not, however, improve overall speech intelligibility in patients with preexisting dysarthria producing speech intelligibility and articulatory precision problems.

Parkinson Plus Syndromes: Progressive Supranuclear Palsy Progressive supranuclear palsy (PSP) is a rare degenerative disorder characterized by supranuclear ophthalmoplegia, complaints of falling backward, nuchal dystonia in extension, moderate axial dystonia, pseudobulbar palsy, difficulty in swallowing, dysarthria, bradykinesia, masked facies, nonspecific changes in personality, lability, sleep disturbance, and performance decrements on various neuropsychological tasks. Progressive supranuclear palsy may be differentiated from Parkinson's disease in that supranuclear ophthalmoplegia is characteristic of PSP, and the tremor present in parkinsonism is typically absent in PSP. Patients with PSP may have voice symptoms similar to those in parkinsonism, with the possible exception of vocal tremor. Dysphagia is a major problem for these patients. Hypophonia is present, with unilaterally reduced vocal fold range and speed of movement. The associated dysarthria may include palilalia (uncontrolled syllable repetition) and oral motor rigidity.[18]

Multiple Systems Atrophy Multiple systems atrophy is another disease that is included as a Parkinson plus syndrome. This is a rare degenerative movement disorder with lesions in the cerebellum, brainstem, and basal ganglia.[19] Three clinical variants are recognized: (1) spinocerebellar degeneration (olivopontocerebellar atrophy), (2) progressive autonomic failure (Shy-Drager syndrome), and (3) atypical parkinsonism (striatonigral degeneration).[20]

Olivopontocerebellar atrophy is characterized by progressive cerebellar ataxia, and coordination of the laryngeal muscles is affected, as in ataxic dysarthria. One sees loss of muscle coordination (dyssynergia), loss of ability to gauge range of motion (dysmetria), and tremor during voluntary movement (intention tremor). The degree of laryngeal impairment depends on the severity of the ataxia. Dysphonia may take one of several forms: sudden bursts of loudness, irregular increases in pitch and loudness, or coarse voice tremor.[21]

In Shy-Drager syndrome, voice symptoms typical of Parkinson's disease are present, along with progressive autonomic dysfunction.[22] Voice may be more severely affected owing to reduced range and speed of abduction of the vocal folds. Speech is slowed, with a reduction in vocal intensity, and there is an absence of glottal stops and poor voicing contrasts. Patients may also be predisposed to laryngospasm. The most prominent feature is airway compromise secondary to bilateral abductor paresis, and patients often require tracheostomy in the late stages of the disease.[23]

Voice treatment for these syndromes largely depends on the stage of disease and may include compensatory strategies as well as supportive nonverbal communication alternatives in end stages.

Pseudobulbar Palsy Pseudobulbar palsy results from neuronal loss above the level of the nucleus ambiguus involving the corticobulbar tracts bilaterally. It can occur with vascular and degenerative lesions involving the cortical motor areas bilaterally, vascular lesions and tumors of the internal capsule or brainstem, degenerative lesions involving the entire corticobulbar tract system, or infections. Dysphonia associated with these types of lesions is characterized by a strained-strangled, harsh voice, likely the result of hyperadduction of true and false vocal folds. This is thought to be the result of loss of inhibition of excitation of vagal nuclei. Individuals with pseudobulbar palsy have a breathy voice with vocal fold asymmetry.[21]

Multiple Sclerosis Multiple sclerosis is a progressive demyelinating disorder, having sensory and motor impairments, cognitive problems, spasticity, and tremor. Although MS is one of the most common causes of ataxia, patients with MS usually have evidence of central nervous system disease outside the cerebellum and its pathways.[23] For example,

optic nerve and corticospinal tract involvement are usually present. Dysarthria, tremor, and decomposition of movement are frequently associated with gait and limb ataxia in MS.

Essential features include staccato speech and harsh voice quality with intermittent hyperadduction of the vocal folds. Spasticity and tremor may affect all or some regions of the vocal tract, thereby reducing speech intelligibility and making speech effortful. Breathy voice and vocal fold asymmetry are common findings. Depending on the systems involved, voicing may or may not be affected in some stages of this disease. Given the predominant ataxia that may influence speech motor systems, voicing is characteristic of other ataxic dysarthrias (see the section of this chapter on cerebellar degeneration). In one case of vocal fold paralysis associated with MS, botulinum toxin injection into the abductor musculature appeared to result in improved vocal fold mobility and voice,[24] presumably by allowing "rebalancing" of the forces with the intralaryngeal musculature.

Myoclonus Myoclonus is a sudden, brief, shocklike involuntary movement caused by active muscular contractions.[25] The distribution may be focal, multifocal, or generalized, with the presentation spontaneous, active, and stimulus sensitive. The source of the neural discharge may be cortical, brainstem, or spinal. Four main pathophysiologic categories of myoclonus have been described: cortical, subcortical, cortical-subcortical, and spinal myclonus differ in the location of neurons involved in producing myoclonus.[25] Sometimes these jerks are produced in response to external stimuli, but most occur both at rest and during purposeful movements.

Palatal myoclonus is focal, affecting the movements of the palate either unilaterally or bilaterally at 1.5 to 3 Hz. It is frequently accompanied by synchronous movements of adjacent muscles such as the extraocular muscles, tongue, larynx, face, neck, or diaphragm.[26] When the soft palate is involved, abrupt, rhythmic, anteroposterior, and vertical movements occur. Pharyngeal muscle contractions can open and close the eustachian tube, thereby producing a bruit or clicking sound transmitted by the tympanic membrane and sometimes heard by others. The constant clicking sounds are often the patient's primary complaint and can be benefited by injection of the tensor veli palatini muscles with small dosages of botulinum toxin.[27] Movements are often present during sleep and may occur without the patient's knowledge. Myoclonus may be idiopathic or may be seen in association with stroke, MS, tumors, trauma, or metabolic encephalopathy.[28] Pathology affecting the inferior olivary nuclei may be the origin of palatal myoclonus.[29] When the pharynx and larynx are involved, rhythmic adductor movements of the vocal folds and gross upward and downward movements of the larynx cause momentary and rhythmic phonatory interruptions. Voicing deficits are not often perceived during connected speech but will be apparent on vowel prolongation.

Lower Motoneuron Disorders

Amyotrophic Lateral Sclerosis Amyotrophic lateral sclerosis, a motoneuron disease, may produce a mixed dysarthria because of both upper motoneuron involvement and some lower motoneuron involvement, particularly in the late stages of the disease (flaccid-spastic paralysis). Amyotrophic lateral sclerosis is a degenerative disease of the corticobulbar tracts and lower motoneuron nuclei. The speech and voice symptoms may vary depending on the predominance of spastic or flaccid components. Flaccid symptoms present as hypoadduction of one or both vocal folds and pooling of saliva in the piriform sinuses on laryngeal examination. Voicing is characterized by breathy, hypernasal quality with reduced intensity range. A "wet" phonation is the result of poor management of secretions owing to flaccidity. Dysphagia is common. If spastic components prevail, voicing will be strained and harsh because of hyperadduction of the true and false vocal folds. A mixed form of dysphonia may result in voicing characterized by both flaccid and spastic components.

Amyotrophic lateral sclerosis occurs in the fifth to seventh decade of life and may present with primarily pyramidal tract signs or lower motoneuron signs of progressive muscular atrophy. Fasciculations are more common in the pyramidal tract form. When the disease affects the brainstem rather than the spinal cord, it may progress more rapidly. Speech symptoms may be the first signs in the bulbar type.[30] Facial muscle weakness, palatal weakness, and lip, tongue, and jaw weakness with tongue fasciculations are predominant and cause poor speech intelligibility. Voicing is weak and breathy because of flaccidity of the vocal folds.

Myasthenia Gravis Myasthenia gravis is a disorder of acetylcholine transfer at the myoneural junction, characterized by weakness and fatigability of striated muscle. Muscle contraction, dependent on stimulation of the motor end plate by acetylcholine, is weakened or reduced by the reduction of acetylcholine receptors. This disorder causes a flaccid dysphonia, characterized by breathy, weak phonation. Sometimes stridor can develop with bilateral abductor muscle weakness.[31] The voice intensity range is reduced, and sustained effort causes progressive weakness. This disorder may affect phonation (larynx), resonation (velum), and articulation (lip, tongue, and jaw), and these systems may be affected separately or serially as the disease progresses. The larynx is less frequently affected, whereas the extraocular muscles are usually the first affected. Dysphagia can be severe.[32] This and other early disorders of neurogenic origin may present with reduced movement range and speed of the vocal folds on laryngoscopic examination and are often mistaken for psychogenic voice disorders. Flaccid dysphonia may be an early symptom of neurogenic disease.[30]

Wallenberg's Syndrome Occlusion of the posterior inferior cerebellar artery may produce infarction of the lateral medulla, resulting in Wallenberg's syndrome, also known as lateral medullary syndrome. The medial and descending vestibular nuclei are usually included in the zone of infarction consisting of a wedge of the dorsolateral medulla just posterior to the olive. This syndrome is marked by dysarthria and dysphagia, ipsilateral impairment of pain and temperature sensation on the face, and contralateral loss of pain and temperature in the trunk and extremities. Major symptoms include vertigo, nausea, vomiting, intractable hiccupping, ipsilateral facial pain, and diplopia. Unilateral vocal fold paralysis and flaccid dysphonia occur when the nucleus ambiguus or corticobulbar tracts leading to it are affected.[33] If the paralysis persists past 1 year, both voice and swallowing can sometimes be benefited by a thyroplasty to augment the paralyzed side.

Postpolio Syndrome Of the 250,000 survivors of the poliomyelitis epidemics, approximately 25% experience progressive muscle weakness known as postpolio syndrome. Postpolio patients who complain of swallowing difficulties are at risk for laryngeal dysfunction. This syndrome is characterized by the new onset of progressive muscle weakness, fatigue, and pain. This may occur 30 to 40 years after the initial infection with polio. Electrodiagnosis of neuronal dropout or axonal loss in these patients is consistent with neurogenic change. Patients with previous bulbar symptoms show evidence of neurogenic change. The course of this disease is variable. In some cases, patients may develop progressive vocal fold involvement leading to bilateral vocal fold paralysis and acute respiratory distress. Patients with active swallowing complaints require thorough laryngeal and voice assessment to evaluate the coexistent laryngeal pathologic process in addition to appropriate therapy for dysphagia.[34]

Disorders of the Peripheral Nervous System

The recurrent laryngeal nerves (RLNs) and external branches of the superior laryngeal nerves (SLNs) supply motor innervation to the laryngeal muscles, providing support for fine control, as described in Table 53–1. The conduction velocity is high, and the innervation ratio per motor unit is low, estimated at 100 to 200 cells per motor unit. This represents a smaller innervation ratio than the hand, indicating a very fine degree of contraction and ability to control precisely phonatory parameters such as fundamental frequency.[35]

Lesions of the tenth cranial nerve at any point along its pathway from the nucleus ambiguus in the brainstem to the musculature cause paresis or paralysis of the laryngeal muscles, resulting in dysphonia or even aphonia. The extent of weakness and the degree of dysphonia depend on the location of the lesion along this pathway (Table 53–2).

Evaluation of Voice Function

Clinical evaluation of dysphonia includes diagnostic procedures to determine the cause of the voice disorder, the degree of phonatory deficit, and the prognosis for recovery of function. Based on these assessments, therapeutic recommendations may involve behavioral, medical, or surgical treatment. The most effective treatment plan often includes a multidisciplinary approach, using complementary techniques. For example, patients who present with vocal fold paralysis may require medialization surgery. To optimize the surgical effects on phonation,

TABLE 53–1. Characteristic Functions of the Laryngeal Muscles in Vocal Fold Adjustments

	CT	*VOC*	*LCA*	*IA*	*PCA*
Position	Paramed	*Adduct*	*Adduct*	*Adduct*	*Abduct*
Level	Lower	Lower	*Lower*	0	*Elevate*
Length	*Elongate*	*Shorten*	Elongate	(Shorten)	*Elongate*
Thickness	*Thin*	*Thicken*	Thin	(Thicken)	Thin
Edge	*Sharpen*	*Round*	Sharpen	0	Round
Muscle (body)	*Stiffen*	*Stiffen*	Stiffen	(Slacken)	Stiffen
Mucosa (cover and transition)	*Stiffen*	*Slacken*	Stiffen	(Slacken)	Stiffen

0 = no effect; ( ) = slightly; italics = markedly; CT = cricothyroid muscle; VOC = vocalis muscle; LCA = lateral cricoarytenoid muscle; IA = interarytenoid muscle; PCA = posterior cricoarytenoid muscle.
Reproduced with permission from Hirano M.[2]

however, voice therapy is often recommended pre- and postoperatively to help the patient to undo the previous compensatory techniques before surgery and to learn to use the "new vocal instrument" after surgery. Some diagnostic parameters measure the degree and nature of vocal impairment to determine appropriate intervention. Monitoring the patient's progress throughout treatment is essential to successful outcome. Because phonosurgery does not always result in a normal voice, the speech-language pathologist or laryngologist must decide whether difficulties in ease of phonation are structurally or behaviorally based. This is best accomplished with a combination of instrumental and perceptual tests.

Videostroboscopy The importance of adequate visualization of the larynx during phonation and respiration cannot be overestimated. These observations are easily made using laryngeal videostroboscopy, which allows viewing of the vibratory characteristics of the vocal folds, as well as opening and closing gestures. A description of vibratory characteristics of the vocal folds should include basic information about the (1) symmetry of bilateral movements, (2) regularity, (3) glottal closure, (4) amplitude, (5) mucosal wave, (6) extent of nonvibratory portions, and (7) the opening-closing pattern.[36]

Acoustic and Perceptual Measures Perceptual assessment and objective acoustic measures are also useful in documenting the progress and success of treatment. The degree of impairment is necessarily reflected in the nature and degree of vocal control available to the patient. Measures of frequency and intensity characteristics of phonation should include mean, range fluctuation, extent, and periodicity. Fundamental frequency is the acoustic correlate of pitch, measured in hertz. Intensity, measured in decibels, is the acoustic correlate of loudness. The range in each of these parameters indicates the limits of laryngeal movement dynamics available to the subject. Instrumental assessment of these factors as well as noise and harmonic structure in the voice signal is now available in a variety of computer-based systems designed specifically for voice analysis. Measures of the irregularity in frequency and intensity of the voice signal, also known as perturbation measures, may be useful in monitoring vocal performance[1] in normal voices but often cannot be accurately used in dysphonia because of tracking difficulties for an abnormal signal. Jitter refers to cycle-to-cycle variation in time or period of vibration (frequency). Shimmer refers to variation in amplitude from cycle to cycle (intensity). Harmonics-to-noise ratio is often used to assess the degree of energy contained in harmonics of the fundamental frequency compared to turbulent noise between harmonics present during phonation. These values are thought to reflect abnormalities or asymmetries in mass, neural control, tension, and biomechanical characteristics of the vocal folds.

Aerodynamic Measures Some measure of airflow or volume velocity is useful in determining how rap-

TABLE 53–2. Effects of Neurogenic Lesions

	Effect on Vocal Folds		*Effect on Phonation*		*Effect on Soft Palate*			
Level of Lesion	*Unilateral Lesion*	*Bilateral Lesion*	*Unilateral Lesion*	*Bilateral Lesion*	*Unilateral Lesion*	*Bilateral Lesion*	*Effect on Nasal Resonation*	*Associated Signs*
Above origin of pharyngeal, superior laryngeal, and recurrent laryngeal nerves	One vocal fold fixed in abducted position	Both vocal folds fixed in abducted position	Breathy, moderate, reduced loudness and pitch	Extremely breathy to whispered (aphonia)	One side low, immobile	Both sides low, immobile	Hypernasality, nasal emission	Glottal croup and cough absent, weak, or mushy; difficulty in swallowing; nasal regurgitation of food; aspiration of secretions; pharyngeal paralysis
Above origin of superior laryngeal and recurrent laryngeal nerves but below origin of pharyngeal nerve	Same as above	Same as above	Same as above	Same as above	None	None	None	Same as above, except no pharyngeal paralysis or difficulty in swallowing
Superior laryngeal nerve	Both vocal folds able to adduct, affected vocal fold shorter, asymmetric shift of epiglottis and anterior larynx toward intact side on phonation	Absence of tilt of thyroid on cricoid cartilage, inability to view full length of vocal folds because of epiglottic overhang, vocal folds bowed	Breathy, hoarse	Breathy, hoarse, reduced loudness, restricted pitch range	None	None	None	None

Continued

TABLE 53–2. *Continued*

	Effect on Vocal Folds		*Effect on Phonation*		*Effect on Soft Palate*			
Level of Lesion	*Unilateral Lesion*	*Bilateral Lesion*	*Unilateral Lesion*	*Bilateral Lesion*	*Unilateral Lesion*	*Bilateral Lesion*	*Effect on Nasal Resonation*	*Associated Signs*
Recurrent laryngeal nerve	One vocal fold fixed in paramedian position	Both vocal folds fixed in paramedian position	Breathy, hoarse, reduced loudness, diplophonia (not in all cases)	Breathy, hoarse, reduced loudness	None	None	None	Unilateral: marginal airway, weak cough Bilateral: severe difficulty on inhaling for life purposes, inhalatory stridor, tracheostomy often necessary
Myoneural junction (myasthenia gravis)	Not applicable	Restriction of adductor-abductor movements	Not applicable	Breathy, hoarse, reduced loudness; symptoms worsen with sustained speaking	Not applicable	Both sides low, immobile	Hypernasality, nasal emission; symptoms worsen with sustained speaking	Difficulty in swallowing, nasal regurgitation of food, inhalatory stridor, articulation defects

From Darley FL et al.[21]

idly the air passes through the glottis. Mean airflow may be obtained from averaged measures over several glottal cycles. Mean airflow rates are useful in documenting change following phonosurgery, especially in RLN paralysis, because preoperative airflow is high. These measures may be made using a pneumotachograph.[1,2] A spirometer is used to determine flow volume loops, which are altered by glottal narrowing.

Subglottal pressure is important for vocal fold vibration and for modulation of vocal intensity. Subglottal pressure may be measured indirectly by a pressure transducer placed in the oral cavity.[37] During production of the sound "p," requiring lip closure, the vocal tract is a closed tube with equal pressure throughout. Pressure measured in the oral cavity reflects pressure beneath the glottis. In a related measure, glottal resistance is calculated by dividing subglottal pressure by mean airflow. Glottal resistance increases with vocal intensity and is thought to be low in case of RLN paralysis. Phonation threshold pressure is the lowest subglottal pressure that a person can use to achieve phonation and, as such, can be used as a measure of vocal fold stiffness.[38] These and other quantifiable measures are currently routinely used in the evaluation of voice deficits, including vocal fold paralysis. The acquisition of perceptual and objective data is key in the treatment program of patients with neurogenic voice disorders.[2,4]

Neurophysiologic Measures of Voice and Laryngeal Function Electromyography is a useful adjunct in the assessment of neuromuscular disorders and is often used in prognostic judgments about patients with those disorders. Electrical silence, fibrillation potentials, polyphasic potentials, high-amplitude potentials, and percentage of normal potentials are the basis for interpretation of such examinations. Electromyography may be useful in differentiating a variety of voice disorders: peripheral and central vocal fold paralysis, functional disorders, and arytenoid joint fixation (Table 53–3).[2,39] Others have suggested the use of neuromyographic assessment that involves the direct stimulation of a laryngeal nerve in conjunction with the use of EMG to monitor resulting muscle activity.[40] During this examination, the contraction of laryngeal muscles is monitored with EMG, while the skin surface is stimulated in the area of the SLN. This produces a short latency response that is brainstem mediated and effected by the RLN to the thyroarytenoid muscle. This method is thought to be one means of obtaining an early estimate of the type and degree of RLN injury. Compared with standard EMG, this technique may offer some advantage in monitoring the reinnervation of paralyzed laryngeal muscles following different reinnervation techniques.[39,40]

Phonosurgery for Vocal Fold Paralysis

The glottis must remain closed for airway protection during swallowing, be patent for respiration, and oscillate for vocalization.[41–43] Neither of the first two vital functions can be compromised in surgery aimed at improving phonation. Because speech requires effective motor coordination of muscles of respiration and the pharynx, mouth, and tongue as well as the larynx, if these other functions are impaired, surgery for vocal fold paralysis for speech may not be warranted, such as in patients with severe dysarthria.

The vocal fold has three functional layers: (1) the cover, which consists of the epithelium and the superficial layer of the lamina propria, or Reinke's space; (2) the vocal ligament, which consists of the confluence of the intermediate and deep layers of the lamina propria; and (3) the vocalis muscle. The cover is the most compliant layer and vibrates most during phonation. Scarring within this layer must be avoided because it can severely disturb the mucosal wave and thus the quality of the voice. The mechanical relationships between these layers determine the vibratory properties of the vocal fold and the quality of the phonation. Any proposed phonosurgery must take into account not only the effect on the position of the vocal fold as a whole but also the effect on the mechanical relationships between the layers. This becomes particularly important in the selection of locations for vocal fold injections aimed at augmentation. The movement of the vocal folds during phonation is at least two-dimensional and includes both lateral-medial and superior-inferior motion during vibration. The shape of the vocal fold changes during vibration as the wave in the mucosa travels from the inferior to the superior level as well as from the medial to lateral surface. New methods of high-speed imaging are now showing that vibration can involve much of the laryngeal tissue during phonation in three dimensions. It is not yet known

TABLE 53–3. Electromyographic Patterns in Relation to the Duration of Vocal Fold Paralysis

Muscle	*Duration EMG*	*–2W*	*2W–1M*	*1M–2M*	*2M–3M*	*3M–6M*	*6M–12M*	*12M+*	*Total*
VOC	S	0	3	1	1	2	3	6	16
	F	1	3	11	5	6	0	5	31
	P	0	0	0	1	4	2	0	7
	H	0	0	1	0	0	0	0	1
	N or n	3	3	15	2	5	4	3	35
	Total	4	9	28	9	17	9	14	90
LCA	S	1	2	0	0	3	2	3	11
	F	1	6	7	7	8	0	4	33
	P	0	0	0	3	6	5	1	15
	H	0	0	1	0	1	1	0	3
	N or n	2	8	25	5	9	6	13	68
	Total	4	16	33	15	27	14	21	130
PCA	S	0	2	0	0	3	2	6	13
	F	0	5	11	7	4	0	4	31
	P	0	0	0	1	0	3	0	4
	H	0	0	1	0	0	0	0	1
	N or n	2	3	14	4	5	7	9	44
	Total	2	10	26	12	12	12	19	93

VOC = vocalis; W = weeks of paralysis; M = months of paralysis; S = electrical silence; F = fibrillation potential; P = polyphasic potential; H = high-amplitude potential; n = normal potential (reduced in number); N = normal potential; LCA = lateral arytenoid; PCA = posterior arytenoid. When two or more different patterns were observed within a given muscle, the categorization in this table was made according to the following rule: F + S → F, P + x → P (x: any pattern), H + x → H, n + S and/or F → n, N + S and/or F → N.
From Hirano M.[2]

to what degree the vibration of other parts of the sound propagation.

Perioperative Management Numerous disease processes can negatively influence the results of phonosurgery. Because phonosurgery is elective, these conditions must be treated before surgery is performed. Preoperatively, all patients should discontinue tobacco use. Many patients benefit from an appropriate program of aerobic or vocal exercise. It is also important to optimize management of nutrition; diabetes; pulmonary disease, especially in patients with chronic obstructive pulmonary disease (COPD); chronic allergy; and gastroesophageal reflux disease (GERD). The benefits of the most technically superb phonosurgery are likely to elude patients with either (1) inability to sustain phonation because of active COPD, vocal edema owing to chronic allergy or smoking, or continued inflammation in the larynx owing to persistent reflux or (2) a postoperative wound infection owing to uncontrolled diabetes and poor nutrition. Postoperatively, many surgeons recommend a short period of complete voice rest (less than 1 week), although the benefit from complete vocal rest has not been proven. Alternatively, modified voice rest can be recommended for which extended speaking without adequate rest and loud phonation are avoided. Whispering can be damaging and is not indicated after any phonosurgery.

Unilateral Paralysis Before phonosurgery, a careful search must be made for the primary cause of the dysfunction. Vocal fold immobility may be attributable to surgical damage to the RLN, such as during thyroid or cancer surgery, or may result from invasion of the nerve by tumor, from neuromuscular disorders or neurodegenerative disease, or from trauma to or disease of the cricoarytenoid joint. Procedures that are appropriate to perform while waiting for recovery and compensation include medialization thyroplasty, collagen, and Gelfoam or fat injection, which have less chance of causing permanent damage. Voice and swallowing therapy should be initiated during this period.[44] For patients with known terminal or unresectable disease, such as advanced cancer invading the RLN, time should not be wasted, and phonosurgery, if desired, should proceed at once.

Surgery for Glottic Incompetence Glottic closure is essential not only for phonation but also for airway protection. Procedures to correct glottic insufficiency are of three major types: augmentation, medialization, and reinnervation. With all techniques, a delicate balance must be reached: providing sufficient glottic closure for phonation and airway protection while leaving an adequate opening to allow respiration.

Augmentation. Vocal fold injection places material within or lateral to the vocal fold to move it medially or to smooth its medial edge. It is best performed under direct laryngoscopy with the patient awake to allow for intraoperative assessment of results. Topical lidocaine anesthesia and intravenous sedation are sufficient for most patients.[45] Injection is accomplished with a Bruening syringe, which delivers about 0.2 mL of material with each click of its ratcheted piston. One or two injection sites are usually sufficient.

Polytetrafluoroethylene paste (Teflon) has been widely used in the past for the treatment of vocal fold paralysis since its introduction by Arnold in 1962.[46] It was injected lateral to the thyroarytenoid muscle to prevent extrusion and to minimize alteration of the mechanical properties of the vocal fold layers. Many patients have achieved initially good voice results following Teflon injection, but local inflammatory responses to this foreign material (ie, Teflon granulomas) are now well documented[47] and believed by some authors to occur in all cases over time.[48] Overinjection or misplaced injection of Teflon appears to increase the risk for development of granuloma.[49] Enlargement of the granuloma may cause the free edge of the vocal fold to become irregular and stiff, resulting in worsening dysphonia. Airway obstruction can also occur owing to progressive enlargement. Surgical management of laryngeal Teflon granuloma is still evolving but includes endoscopic laser vaporization[50] or resection of the granuloma via an open approach.[51] Unfortunately, these procedures may result in persistent dysphonia owing to vocal fold scarring. Other disadvantages of Teflon are related to its irreversible nature, potential interference with the mucosal wave if injected superficially, paste migration, and, most importantly, exclusion of future motor recovery. Many laryngologists now perform Teflon injections only infrequently because of these disadvantages and owing to the availability of newer techniques of vocal fold medialization without these disadvantages.[51]

Gelfoam paste can be injected when return of vocal fold function is anticipated and only temporary medialization is desired.[52] An effect lasting 6

weeks to 3 months is typical. Although scarring and deformation of the vocal fold can occur, this material is usually absorbed, and no permanent damage occurs if normal function returns to the vocal fold.

Autologous fat injection for treatment of vocal fold paralysis was first reported by Mikaelian et al in 1991.[53] Cited advantages include the lack of foreign body inflammatory response, the availability and ease of harvesting graft material, and the soft, pliable texture of fat that does not impair the normal vibratory characteristics of the vocal fold. It must be carefully harvested to prevent fat cell destruction and is injected into two to three locations within the thyroarytenoid muscle.[53] Problems with resorption of the injected fat over time have led most surgeons to overinject the vocal fold by 30 to 50%. Rosen reported no noticeable resorption of fat beyond that which occurs within the first postoperative month.[45] Several techniques of fat harvesting have been proposed to enhance long-term graft survival.[54,55]

Collagen injection is a preferred technique for correction of gaps caused by vocal fold atrophy and other small defects. Collagen should be injected superficially, into the deep layer of the lamina propria to prevent rapid resorption. If this rule is followed, the correction present after 3 months tends to persist.[56] Over time, collagen has a softening effect on the tissues, so reinjection at a later date is usually easy. Earlier studies used injectable bovine collagen preparations,[56,57] which carry the risk of hypersensitivity reactions or transmission of infectious agents. Autologous collagen has been proposed as a safer alternative to the bovine preparation.[58–60] Autologous fascia may also prove to be useful for vocal fold augmentation in the future.[61]

Vocal fold injection is a simple technique but is not without pitfalls. The ideal material for injection has not yet been found. Nonetheless, vocal fold injection can provide marked benefit for patients with glottic insufficiency causing aspiration, weak or absent cough, short maximum phonation time, breathy voice, insufficient loudness, and limited pitch range.

Medialization Thyroplasty. Laryngeal framework surgery provides methods to medialize, separate, shorten, or lengthen the vocal folds. These procedures, referred to as thyroplasty types I through IV, respectively, were popularized by Isshiki and bear his name. If the preoperative laryngeal examination reveals contact of the mobile vocal process with the paralyzed vocal process during phonation, then type I thyroplasty alone usually results in significant improvement.[62] If a wide posterior glottal gap is noted during phonation, a concomitant arytenoid adduction procedure (described below) may be required to achieve an optimal voice result. The lateral compression test is sometimes useful for preoperative assessment of these patients. In this maneuver, the examiner uses the thumb and forefinger to compress the thyroid alae medially while the patient sustains a vowel sound. Subjective and objective measures (ie, maximum phonation time, videostroboscopy, acoustic analysis) of change, if any, are noted. If improvement is noted with the lateral compression test, then type I thyroplasty is usually successful. However, it may still be successful even if no improvement is noted on manual compression, especially in older patients with ossified, inflexible thyroid cartilages.

Like vocal fold injection, thyroplasty is best performed with the patient awake to allow intraoperative monitoring of results. The patient is placed supine on the operating room table. Intraoperative flexible fiber-optic laryngoscopy is used to verify the side of paralysis. The entire procedure is accomplished under local anesthesia consisting of 1% lidocaine and 1:100,000 epinephrine. A 3 to 4 cm transverse incision is made along a skin crease overlying the involved thyroid ala from a point on the anterior cervical midline extending laterally (Figure 53–2). Superior and inferior subplatysmal flaps are raised, and the retraction of the strap muscles exposes the outer perichondrium of the thyroid cartilage ala.

Using a caliper, a horizontal midline of the thyroid ala is determined. The size of the alar window measures approximately 6×13 mm. The window is carefully positioned below the horizontal midline 5 to 7 mm from the anterior commissure (Figure 53–3). The window should be as close to the inferior border of the thyroid ala as possible while leaving a sufficiently strong inferior cartilage strut to support the implant (3 mm).[62] The exact placement and size of the window should take into account the anticipated shape and size of the implant and the overall size of the larynx.[63]

An inferiorly based perichondrial flap is raised according to the calculated window dimensions. The cartilage window is cut with a 2 to 3 mm otologic bur. Care is taken to avoid violating the inner alar perichondrium. The cartilage window is then

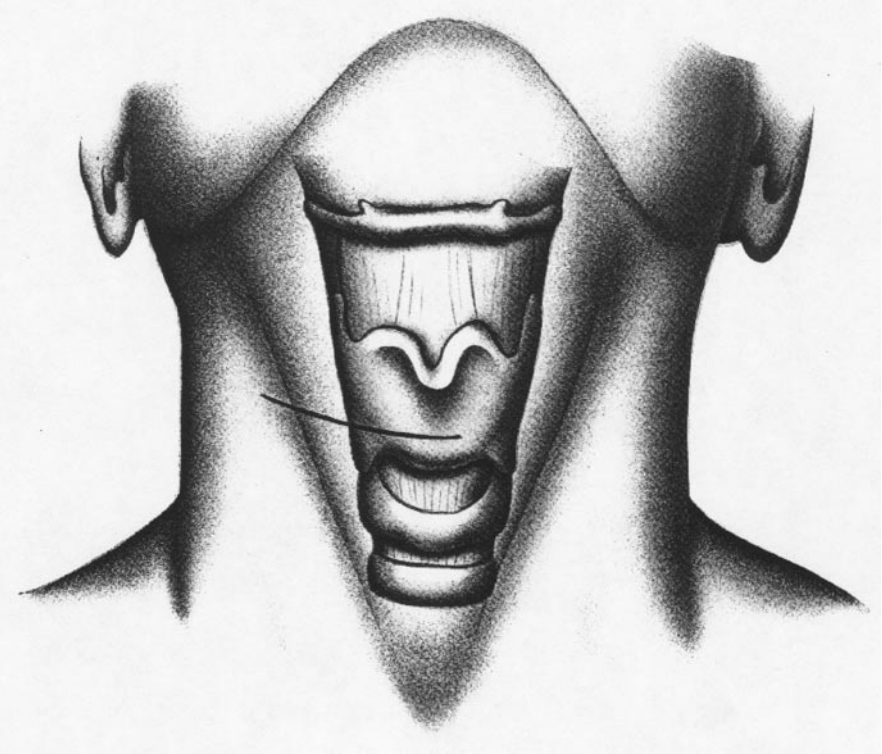


FIGURE 53–2. The incision is outlined in a cervical crease overlying the thyroid cartilage extending more to the side of the paralysis. If indicated, simultaneous procedures such as arytenoid adduction or reinnervation can be performed through one incision. Reproduced with permission from Wanamaker JR et al.[62]

removed, and the inner perichondrium is carefully elevated from the medial surface of the thyroid cartilage surrounding the window using a Woodson elevator. Too high an elevation from the horizontal midline may lead to perforation at the laryngeal ventricle. Some authors recommend creation of an anteriorly based flap of the inner perichondrium to enhance the accuracy and effectiveness of medialization (Figure 53–4).[62] The Silastic prosthesis is fashioned and placed into the alar window. When optimal medialization has been established, the prosthesis is scored flush with the outer perichondrium and is cut across with a scalpel to result in an embedded edge (Figure 53–5). The inferior perichondrial flap is sutured back to its anatomic position with interrupted 5-0 nylon sutures, and the wound is closed in a two-layer fashion over a quarter-inch Penrose drain.

McCulloch and Hoffman introduced the use of expanded polytetrafluoroethylene (Gore-Tex) as an alternative to Silastic for medialization.[64] Cited advantages of this material include the ability to make incremental and easily reversible adjustments in medialization, the lack of precision required in creating the cartilage window or shaping the implant, and the lack of special instrumentation required.[65]

Advantages of type I thyroplasty over Teflon injection include easy intraoperative adjustability and the ability to remove the implant if vocal fold movement returns. Thus, in patients who may spontaneously recover neuromuscular function, corrective surgery can be performed before the usual 6- to 8-month waiting period. Type I thyroplasty allows the medialization of the entire fold rather than an isolated or uneven segment sometimes observed following injection techniques. Following successful type I thyroplasty, an immediate improvement is seen both in breathiness and maximum phonation time. Over time, reductions in voice strain are also common.

Poor or suboptimal voice results most commonly from placing the window too far superiorly or anteriorly. This may result in excessive convex displacement of the anterior part of the vocal fold and diminished voice quality. Of even more concern, excessive anterior placement predisposes to penetration of the ventricular mucosa, granuloma formation, and implant extrusion.[58,63,66]

Arytenoid Adduction. In 1978, Isshiki et al described the arytenoid adduction procedure for patients in whom vocal fold paralysis resulted in a

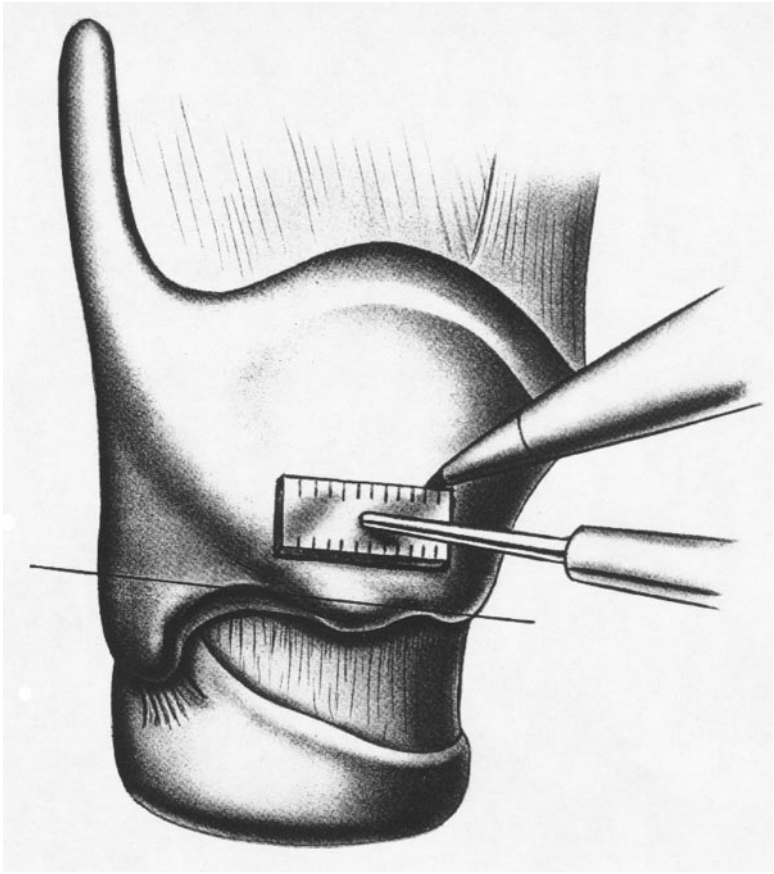


FIGURE 53–3. The window is outlined using the 6 × 13 mm window template parallel to the lower border of the thyroid cartilage. When outlining the lower border, one must elevate both anterior and posterior to the muscular process of the attachment of the cricothyroid muscle. This process is quite variable in size. The true parallel line is obtained by lining up the anterior part of the lower border of the thyroid cartilage with the ridge in the cartilage posterior to this process. The window is outlined as low as possible while leaving an inferior strut (3 mm) that is strong enough to support the implant. Reproduced with permission from Wanamaker JR et al.[62]

large posterior glottal chink or a difference in the level of the vocal folds during phonation.[67] The goal of the procedure is to approximate the physiologic position of the arytenoid during phonation. Under local anesthesia, the cricoarytenoid joint is exposed through dislocation of the cricothyroid joint and subperichondrial dissection around the posterior border and inner surface of the thyroid ala. Careful superior elevation of the piriform sinus mucosa exposes the posterior cricoarytenoid (PCA) muscle and the palpable arytenoid muscular process. The cricoarytenoid joint is opened, and two 3-0 nylon sutures are tied across the muscular process. The sutures are then passed anteromedially through the paraglottic space, through the anterior part of the thyroid ala, and are tied on the outer surface of the ala. The vector of force of these sutures causes a medial-inferior rotation of the arytenoid cartilage, which mimics the action of the lateral cricoarytenoid and lateral thyroarytenoid muscles.

Netterville et al facilitate exposure of the muscular process through removal of a cartilaginous window on the posterior border of the thyroid ala.[68] The fixation suture is then placed through the lateral surface of the muscular process without disrupting the cricothyroid or cricoarytenoid joints. Arytenoid adduction appears to enhance the aerodynamic parameters and perceptual voice quality of many patients undergoing type I thyroplasty, and some laryngologists now routinely expose the muscular portion of the arytenoid cartilage as part of their laryngeal framework surgery to assess the benefit of arytenoid adduction in conjunction with Silastic or Gore-Tex medialization.[69] Arytenoid adduction reduces the posterior glottic aperture (ie, the ventilatory portion of the glottis), and this procedure should be used with caution, if at all, in patients with COPD, with paresis of the contralateral vocal fold, or who require a wide glottal aperture for intensive aerobic activities.

Arytenoid Arytenopexy. Noting the excessively inferior displacement of the vocal fold that may result from the classic arytenoid adduction procedure, Zeitels et al designed the adduction arytenopexy procedure to reflect more closely the simultaneous, three-dimensional effects of the lateral cricoarytenoid, PCA, interarytenoid, and thyroarytenoid

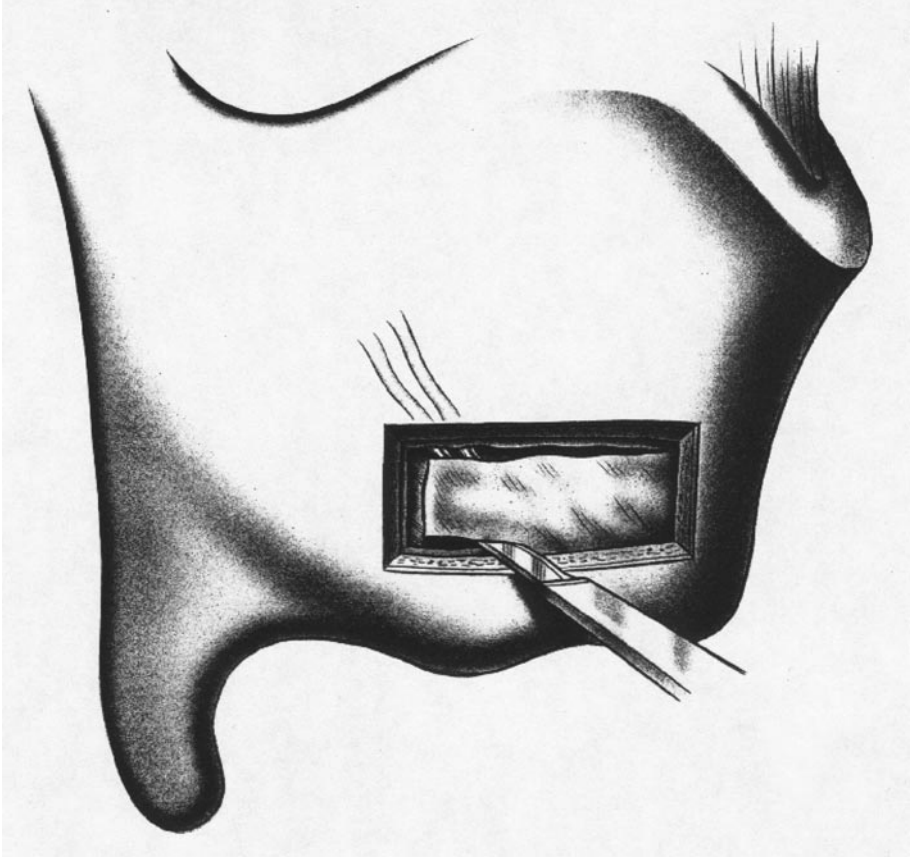


FIGURE 53–4. The perichondrium is carefully incised along the superior, inferior, and posterior margins of the window. Care is taken to minimize injury to the underlying thyroarytenoid fascia, which protects the muscle and the distal branches of the superior laryngeal vessels. Reproduced with permission from Wanamaker JR et al.[62]

muscles on arytenoid position and stabilization during phonation.[70,71] Following exposure and opening of the cricoarytenoid joint, a 4-0 polypropylene suture is placed through the posterior plate of the cricoid cartilage medial to the joint facet. The suture is then passed through the body of the arytenoid cartilage and back out underneath the facet through the posterior plate of the cricoid cartilage where it is tied following endoscopic verification of proper arytenoid position. Silastic medialization (type I thyroplasty) is then performed to optimize the position of the membranous vocal fold. The addition of a cricothyroid subluxation suture has also been proposed to increase the length and tension of the paralyzed fold as needed to enhance the final voice result.[72]

Reinnervation. Reinnervation techniques succeed to the degree that they can reproduce the normal tone and action of the intrinsic laryngeal muscles. Tucker described use of a nerve muscle pedicle from the omohyoid muscle to restore adductory function to the thyroarytenoid muscle in nine patients.[73] All patients had a unilateral vocal fold paralysis with involvement of both SLNs and RLNs, resulting in a large posterior glottal defect and bowing of the affected fold. After creating a window in the lower half of the ipsilateral thyroid ala, the nerve muscle pedicle was sutured to the lateral portion of the thyroarytenoid muscle. Between 2 and 12 weeks postoperatively, some observable return of adductory function and improvement in voice was noted in all patients. Four of the nine patients achieved appropriate adduction and tensing of the affected fold during phonation. Since the ipsilateral recurrent nerve is not disturbed during the dissection, this procedure does not preclude any spontaneous reinnervation that may occur by the RLN.

Crumley et al reported excellent voice quality in patients undergoing ansa cervicalis nerve transfer (NT) to RLN for unilateral RLN paralysis.[74] The ansa's branch to the sternothyroid is sutured to the distal end of the transected RLN using 10-0 monofilament nylon sutures. The best results were seen in patients without prior medialization procedures and in whom the paralysis had occurred within 24 months of NT. Although adductory or abductory movements of the operated side were not observed, NT is thought to impart static tone and mass to the

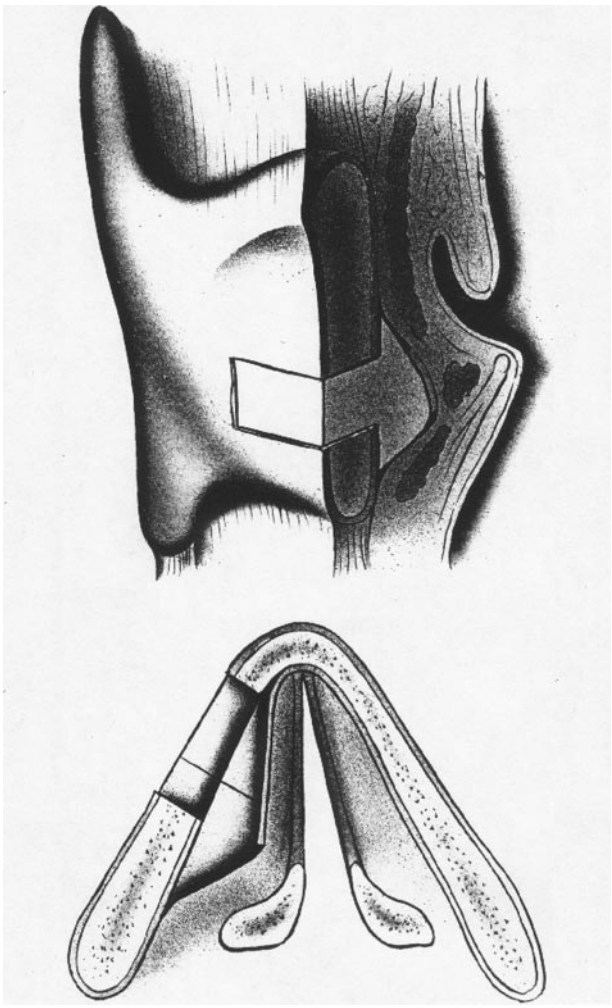


FIGURE 53–5. The final position of the implant (*above*) shows the plane of maximal medialization at the lower margin of the window. This helps to prevent medialization of the false vocal fold and the ventricular mucosa. The posterior extension of the implant (*below*) medializes the vocal process of the arytenoid as well as the mobile true vocal fold. This results in some adduction of the arytenoid cartilage. Reproduced with permission from Wanamaker JR et al.[62]

intrinsic laryngeal muscles, thereby stabilizing the arytenoid cartilage into a more symmetric position compared with the opposite side. The jerky, synkinetic movements sometimes seen following RLN injury have not been observed following NT; the sternothyroid branch of the ansa appears to provide weak, tonic innervation, which maintains vocal fold tone without producing spasms. Crumley et al recommended NT as the procedure of choice in younger patients and voice professionals because the voice quality results, in their experience, are superior to those achieved with Teflon or thyroplasty. The disadvantages of the technique include a waiting period of up to 6 to 12 months postoperatively until improvement is seen, and transection of the RLN disrupts any spontaneous reinnervation that has occurred. Despite these promising results, NT is infrequently performed in the general otolaryngology community for the treatment of unilateral vocal fold paralysis.[75]

Bilateral Vocal Fold Paralysis

Patients with bilateral vocal fold paralysis usually require treatment because of loss of abductor function leading to airway obstruction. Not only is the glottic chink narrowed by the resting position of the vocal folds, but also, with inspiration, the Bernoulli effect causes the folds to move medially as the velocity of the airflow increases. Surgical treatment of this problem usually involves one of three strategies: bypassing the site of obstruction, increasing the caliber of the airway, or attempting to restore abductor function. In weighing treatment options, points to consider include effectiveness, preservation of vocal quality, and possible complications.

Tracheostomy Tracheostomy bypasses the site of obstruction, providing immediate, effective treatment. Although long-term tracheostomy may be less appealing than other treatments, it often plays a role in the treatment of acute bilateral vocal fold paralysis by stabilizing the patient. When quick recovery of function is expected, it may be possible to avoid this operation altogether with the use of continuous positive airway pressure to splint the airway.[76] In other cases, tracheostomy may be useful as an interim treatment, allowing other techniques to be performed under less pressing circumstances.

Increasing Airway Caliber Techniques for enlarging the airway usually involve lateralization of one of the vocal folds and/or the removal of tissue, particularly from the posterior glottis. Work by Hirano and colleagues has shown that the posterior glottis accounts for 50 to 65% of the area of the total glottis.[77] Thus, enlarging the posterior glottis has a greater effect on the airway while preserving vocal quality. These techniques have been applied using both transcervical and endoscopic approaches.

Transcervical Approaches. Chevalier Jackson in 1922 used the midline thyrotomy to resect the vocal fold and ventricle, resulting in a satisfactory airway but a poor-quality voice.[78] Subsequent techniques that used a laryngofissure included cordopexy,[79] submucous resection of the vocal fold,[80] and arytenoidectomy with cordopexy.[81] Later this approach was used to perform arytenoidectomy alone,[82] without lateralization, which was successful in expanding the airway while preserving voice.[83] In 1946, Woodman modified a lateral open approach to perform arytenoidectomy, avoiding laryngofissure.[84] While the arytenoid was excised, the vocal process was preserved and used to lateralize the cord by suturing it to the lesser cornu of the thyroid cartilage. Although this procedure was widely used for many years and resulted in a predictably good airway, the voice results were often poor.

Endoscopic Approaches. Thornell introduced the endoscopic arytenoidectomy in 1948.[85] Good results have been reported with this procedure, and it has been widely used. El Chazly et al reported better airway results when posterior cordectomy was added to arytenoidectomy, although vocal quality was reportedly impaired.[86] Problems with hemostasis during traditional endoscopic arytenoidectomy were overcome with the introduction of the carbon dioxide laser arytenoidectomy by Ossoff et al in 1984.[87] They reported a success rate of 86% for this procedure, which includes vaporization of a portion of the vocalis muscle as well as the arytenoid.[88] In contrast, a procedure that involves vaporization of only the medial portion of the arytenoid has been described in patients with more moderate airway compromise.[89] This procedure is reportedly superior in preserving voice and can be done as a staged bilateral procedure if necessary. Benninger et al stressed the importance of maintaining a medial mucosal flap for coverage of remaining exposed cartilage to reduce postoperative granuloma formation.[90] Temporary stitch lateralization of the operated vocal fold has also been used to ensure an adequate airway until postoperative edema resolves in patients without tracheostomy.[90]

Some endoscopic techniques do not involve arytenoidectomy. Kirchner described a procedure in which he resected a portion of the thyroarytenoid muscle and then lateralized the vocal fold with temporary sutures.[91] Success has also been reported for endoscopic lateralization with permanent sutures without any tissue resection.[92] Other techniques have made use of the carbon dioxide laser. Linder and Lindholm described using the laser to reduce the bulk of the vocal fold laterally, followed by fibrin glue to maintain it in the lateral position.[93] The laser has also been used to remove a portion of the posterior vocal fold (posterior cordectomy), most recently by Dennis and Kashima,[94] who created a crescent-shaped 3.5 to 4 mm defect anterior to the vocal process. Even less tissue was removed with the subsequent development of transverse cordotomy, in which the incision anterior to the vocal process leads to the development of a wedge-shaped defect as the thyroarytenoid muscle retracts.[95] Transverse cordotomy may also be combined with anterolateral arytenoidectomy.[90]

Restoration of Function The techniques described previously often involve a trade off between airway and vocal quality. Thus, the ideal treatment of bilateral vocal fold paralysis remains restoration of function via reinnervation. Direct repair of the injured RLN has not led to satisfactory results; vocal fold motion tended to be uncoordinated and synkinetic.[96,97] Better results were obtained in animals with selective reinnervation of the laryngeal abductor muscle, the PCA.[98] Use of a cable graft directly to the PCA was not successful in humans, however.[99] A more successful procedure has been the use of a nerve muscle pedicle based on the ansa hypoglossi to reinnervate the PCA.[100] In 1989, Tucker reported a long-term success rate of 74% in more than 200 patients with bilateral vocal fold paralysis, with success based on the rate of decannulation.[101] Other centers have not been able to achieve this success rate, however.

Another promising strategy for restoring function in bilateral RLN paralysis involves the use of electrical pacing devices.[102] Electrical stimulation of the PCA using an implantable nerve stimulator has been shown to abduct the vocal folds sufficiently to permit decannulation in five of seven patients implanted worldwide thus far.[103] Presently, the device is turned on by the patient as needed and supplies a programmed schedule of stimulations to which the patient must entrain his/her breathing pattern. Ultimately, a pressure-sensing device may be added to the system to coordinate PCA stimulation with inspirations. The technology and surgical approaches for laryngeal pacing are currently

undergoing improvements and refinements, which will hopefully lead to functional rehabilitation of the paralyzed larynx.

In summary, tracheostomy and endoscopic procedures, particularly those that make use of the carbon dioxide laser, are currently the mainstays of treatment for bilateral vocal fold paralysis. Because of the tradeoff often involved between airway and vocal quality, however, much research continues to focus on the development of techniques to restore function to the larynx.

Laryngeal Paralysis and Aspiration

Vocal fold paralysis alone, whether unilateral or bilateral, rarely results in aspiration. When vocal fold paralysis occurs in conjunction with other motor or sensory dysfunction, however, the combination may significantly impair the protective function of the larynx, leading to aspiration. Neurogenic causes of aspiration may include cerebrovascular accidents, degenerative diseases, neuromuscular disorders, peripheral nerve disorders, intracranial neoplasms, and anoxic or traumatic brain injury.[104]

Evaluation of the patient with aspiration begins with a thorough history and physical examination and requires multidisciplinary evaluation.[105] The history should elicit information about symptoms of any associated neurogenic disorder as well as information about respiratory symptoms that may indicate the degree of severity of aspiration. Physical examination should include visualization of the larynx. Other testing should include chest radiography and modified barium swallow[42] using contrast material of different consistencies performed with the guidance of a speech-language pathologist.

Nonsurgical management of the patient with aspiration usually consists of discontinuing oral intake and providing alternative methods of alimentation. Enteral alimentation is preferable and may be provided with a nasogastric feeding tube. For long-term feeding, however, gastrostomy or jejunostomy may be performed. Attention should also be given to pulmonary toilet. Tracheostomy may be helpful in caring for patients with copious secretions; however, the presence of a tracheostomy may also contribute to problems with aspiration.[42,43]

Specific surgical treatment of the patient with significant aspiration and vocal fold paralysis includes a variety of procedures to medialize the vocal fold discussed elsewhere in this chapter.[43] In patients in whom recovery of function is expected, procedures that are temporary or reversible are indicated, such as Gelfoam paste injection[106] or thyroplasty.[107]

Other surgical techniques are not specifically directed at treatment of vocal fold dysfunction; rather, they involve separation of the upper digestive tract from the upper respiratory tract. Unfortunately, patients undergoing such procedures frequently lose the ability to phonate and may also require permanent tracheostomy. Such procedures can be performed with low morbidity, however, and some of these procedures are at least theoretically reversible should function return.

Narrow-field laryngectomy remains the oldest and one of the most effective surgical treatments of aspiration. Tracheoesophageal puncture can be used to restore phonation. Reluctance to sacrifice an otherwise normal larynx has led to the development of other procedures to close the larynx. Montgomery described a glottic closure technique in which the true and false vocal cords were approximated.[108] Closure was improved by Sasaki and associates with the interposition of a sternohyoid muscle flap.[109] This procedure was extremely effective but did not permit phonation, which was sometimes possible with supraglottic laryngeal closure methods. Since first being described in 1972,[110] the epiglottic flap closure technique has undergone certain modifications, including intentionally leaving an opening posteriorly to permit phonation.[111] Successful reversal of this procedure has been reported by an endoscopic approach.[112] Biller and colleagues have also described a supraglottic closure technique, vertical laryngoplasty, intended for use in patients undergoing total glossectomy.[113] This method also permits speech.

Efforts to devise a procedure that is completely reversible have led to the development of different endolaryngeal stents. Weisberger and Huebsch used a solid Silastic stent in conjunction with a tracheostomy,[114] whereas Eliachar and Nguyen devised vented silicone stents that permit phonation.[115] The advantages of stenting include the use of the endoscopic approach and ease of reversibility. Successful control of aspiration has not been uniform, however, and long-term use of stents carries the risk of endolaryngeal injury, limiting their utility.

Tracheoesophageal diversion was developed by Lindeman with the goal of controlling aspiration

definitively while preserving the larynx and RLNs and thus the potential for reversal.[116] In this procedure, the trachea is divided at the level of the fourth or fifth ring. The proximal trachea is anastomosed to the esophagus, whereas the distal trachea is anastomosed to the skin. Thus, aspirated secretions are diverted back into the gastrointestinal tract. In the modified laryngotracheal separation procedure, the proximal segment is instead closed as a blind pouch.[117] Both techniques have provided good control of aspiration, and both have been successfully reversed.[118] Because no current method is completely satisfactory, investigations continue in an attempt to find a safe, effective means of controlling aspiration without disrupting respiratory or phonatory function. One possibility for the future may be implantable electronic systems now being developed in animals that can produce tight glottic closure during swallow.[119]

FUNCTIONAL VOICE DISORDERS

The term "functional voice disorder" is used to designate a voice disorder resulting from difficulties with the use of the larynx for voice production. There are no structural abnormalities and no peripheral nerve injuries to account for abnormalities in laryngeal function. The disorder occurs only when the person attempts to use their larynx for a particular function such as vocalization or breathing. These disorders are thought to be owing to abnormalities in central nervous system control or behavior.

Functional voice disorders may be (a) idiopathic, (b) attributable to misuse of the larynx for voice production (eg, the development of compensatory behaviors for voice production), or (c) psychogenic, owing to psychological difficulties interfering with normal voice production.

As knowledge and understanding of these disorders progress, fewer will be considered functional after an underlying abnormality has been identified. For example, the SDs are currently classified as idiopathic disorders but are also classified as laryngeal dystonias, a motor control disorder.[120] Until the underlying defect is identified, however, these are considered idiopathic disorders.

Most of these disorders have distinctive symptom patterns, allowing them to be recognized as separate disorders. Further, they usually respond differently to various treatment approaches. Identification depends on the person's voice symptoms.[121] Because diagnosis is based on voice characteristics, consensus has not yet been reached, and diagnoses may differ even among specialized voice centers. Because treatments for controlling symptoms in some of the functional voice disorders are invasive and may alter the larynx, accurate diagnosis is important before such intervention is warranted.

Idiopathic Disorders

In these voice disorders, volitional control of vocal fold movement during the specific laryngeal functions is affected, such as speaking in the SDs and respiration in paradoxical vocal fold movement.

Spasmodic Dysphonias The SDs include *adductor* SD (uncontrolled closing of the vocal folds),[122] *abductor* SD (prolonged vocal fold opening for voiceless sounds extending into vowels),[123] or *vocal fold tremor* (modulations in phonatory pitch and loudness most evident during prolonged vowels). These disorders are characterized by involuntary changes in the ability to maintain voicing during speech either because of intermittent glottal catches (voice breaks) in the adductor type or breathy breaks owing to prolonged vocal fold abductions for voiceless consonants in the abductor type. Often the problem first appears with a very specific task,[124] gesture or posturing of the larynx, such as vowel production during speech, and then may progress to interfere with other tasks or postures such as singing and speaking using falsetto. In mild to moderate forms of SD, patients may have uncontrolled spasms affecting only the speaking voice, with normal singing and falsetto phonation.

Onset often follows an upper respiratory infection, laryngeal injury or inflammation, a period of excessive voice use, or occupational or emotional stress.[125] Increased effort is one of the major patient complaints, along with loss of control and an increased difficulty with prolonged voice use or stress. Onset is between 30 and 50 years of age, and at least 60% of those affected are women. The characteristics of these disorders include the following:

1. They are specific to a particular task, gesture, or posture of the larynx.
2. They are action induced, that is, they appear only with voluntary movement and are not usually apparent at rest.

3. They become worse with prolonged speaking, practice, or performance.
4. Onset is usually gradual, often following some upper respiratory infection, inflammation, or stressful event.
5. Reflexive and emotional aspects of voice function are unaffected, such as coughing, crying, shouting, and laughter.
6. In professional voice users, they may appear during busy professional schedules or following injury.

In some patients, particularly those with abductor SD, some degree of vocal fold asymmetry may be apparent within the first 6 to 12 months, suggesting injury.[126,127] With time, this asymmetry often disappears, but the disorder remains.

Movement abnormalities occur when the patient attempts specific gestures such as phonation of vowels in adductor SD, voiceless consonants (such as s, h, f, t, k) followed by vowels in abductor SD, or using a modal speaking voice in vocal tremor.

No information is available on the etiology of these disorders. Sometimes they occur in persons with a family history of dystonia.[120] However, in clinics specializing in voice disorders, few patients with SD have a positive family history for dystonia.

Diagnosis and Assessment. Diagnosis and management of SDs are best accomplished by a voice team including an otolaryngologist, a speech pathologist, a neurologist, and, in some cases, a psychologist. Because these movement disorders affect the larynx, diagnosis depends on observing the vocal folds during speaking and nonspeech gestures. In addition, the larynx must be visualized to rule out other disorders that could account for the symptoms. The laryngologist rules out vocal fold nodules, polyps, carcinoma, cysts, contact ulcers, inflammation (laryngitis), vocal fold paresis, or paralysis using fiber-optic laryngoscopy. A neurologic examination is necessary to rule out amyotrophic lateral sclerosis, Parkinson's disease, or supranuclear palsy, which can all produce vocal fold movement abnormalities.[128] Some patients may also have concomitant focal dystonias such as writer's cramp or blepharospasm, which can be successfully managed by a neurologist. Many patients may have some degree of laryngeal tremor in addition to spasmodic hyperadduction or hyperabduction. These patients are usually included as a subtype of the SDs[129,130] and may have a more severe disorder.[131]

An extensive history, a trial of voice therapy, and a psychosocial interview may be needed to rule out psychogenic dysphonia.[132,133] Patients may have psychosocial reactions as a result of their voice disorder, confounding the ability to differentiate idiopathic SD from psychogenic disorders through history and interview. For example, many patients will no longer use the telephone and avoid social gatherings as a result of having a speech disorder. Voice disorders have not been found to produce mental illness, however.[134]

Tasks To Be Examined During Fiber-optic Laryngoscopy. Both speech and nonspeech tasks must be sampled to identify (1) movement control abnormalities during vocal fold abduction (opening) and adduction (closing) in speech and (2) movement during nonspeech items such as respiration, sniffing, throat clearing, whistling, and singing. For example, in the paradoxical vocal fold movement disorder, vocal fold abnormalities appear only during inspiratory breathing but not talking (Table 53–4).

Fiber-optic laryngoscopy is useful when examining dysphonia associated with many of the neurogenic disorders to determine how vocal fold movement is affected for adduction and abduction and the speed of motion. Usually, stroboscopy is less helpful because patients with tremor or spasms do not have a regular phonatory cycle that can be tracked from the contact microphone or electroglottographic signal. Further, in patients with other functional voice disorders such as muscular tension dysphonia, the constant hoarseness similarly interferes with tracking. The emerging use of kymography and eventually high-speed video will be particularly useful for examining such patients.

Speech Testing. Voice symptoms should be compared during three tasks to discriminate among adductor SD, abductor SD, and vocal tremor:

1. Prolonged vowel phonation usually manifests tremor if it is present; prolonged vowel production is affected only in the more severe forms of adductor and abductor SD.
2. Production of sentences in which all of the sounds are voiced and frequent glottal stops are at word boundaries (eg, "We mow our lawn all

TABLE 53–4. Movement Characteristics Observed during Fiber-optic Laryngoscopy in Various Types of Functional Voice Disorders

Task	*Movement Examined*	*Adductor SD*	*Abductor SD*	*Tremor Adductor*	*Tremor Abductor*	*Paradoxical Vocal Fold Movement Disorder*	*Muscular Tension Dysphonia*
Deep inhalation	Abduction range	Normal abduction	May be increased	Normal	May have tremor	Vocal fold adduction	Normal
Prolonged vowel "ee"	Adduction for voice	Normal or intermittent hyperadductions	Normal	Repetitive hyperadductions (5 Hz)	Repetitive abductions (5 Hz)	Normal	Constant hyper-adduction
Throat clear 3 times	Adduction	Normal	Normal	Normal	Normal	Normal	Normal
Whisper	Vocal fold partial abduction	May have hyperadduction	May have increased abduction	Normal	May have tremor	Normal	May have constant adduction
Whistling	Vocal fold abduction and adduction	Normal	Normal	Normal	Normal	Normal	Normal
Repeated and quick sniffs	Speed of abduction	Normal	May be increased in range	Normal	May have tremor on abduction	Adduction on sniff	Normal
Alternating between sniff and vowel "ee"	Speed and range of abduction and adduction	Intermittent hyperadduction on vowel	Normal	Tremor on vowel	Tremor on vowel	May have hyper-adduction on sniff	Constant hyper-adduction on vowel

Continued

TABLE 53–4. *Continued*

Task	*Movement Examined*	*Adductor SD*	*Abductor SD*	*Tremor Adductor*	*Tremor Abductor*	*Paradoxical Vocal Fold Movement Disorder*	*Muscular Tension Dysphonia*
Rapid repetition of vowel "ee" 6 times	Phonation offset owing to glottal stop	Prolonged glottal stops	Normal	Tremor on vowel	Tremor on vowel	Normal	Constant hyper-adduction with antero-posterior squeeze
Rapid repetition of "see" 6 times	Speed of abduction and adduction	Normal or intermittent adductions on vowel	Prolonged abductions during "s"	Adductor tremor on vowel	Abductor tremor on vowel	Normal	Constant hyper-adduction
Ascending and descending glides	Controlled lengthening and shortening of folds	Normal or intermittent adductions on low end	Normal	Adductor tremor on low end	Abductor tremor	Normal	Constant hyper-adduction
Sentences: "We eat eels everyday"	Glottal stops in sentences	Prolonged glottal stops	Normal	Adductor tremor	Abductor tremor	Normal	Constant hyper-adduction
"The waves were rolling along"	Constant voicing in sentences	Intermittent spasms in vowels	Normal	Adductor tremor	Abductor tremor	Normal	Constant hyper-adduction
"He will keep the keys"	Voiceless consonants in sentences	Normal	Prolonged voiceless consonants "he," "k"	Adductor tremor	Abductor tremor	Normal	Constant hyper-adduction

SD = spasmodic dysphonia.

year" and "We eat eels everyday") is usually most difficult and demonstrates frequent breaks or voice arrests in adductor SD.

3. Production of sentences with voiceless consonants (s, t, p, k, h) ("She speaks pleasingly," "Keep Tom at the party," and "When he comes home, we'll feed him") is usually most difficult in abductor SD. Sentences with predominantly voiced sounds are much easier to produce and smoother for these patients.

Electromyography. The laryngeal muscle activation abnormalities that can produce problems differ greatly across patients and can account for the wide variety of symptoms manifested.[135] The laryngeal muscles should be examined to determine which muscles contain spasms concurrent with a patient's voice symptoms during speech. In adductor SD, the thyroarytenoid is the most often affected.[136] In abductor SD, spasms can be seen in the cricothyroid in some patients[123] and in the posterior cricoarytenoid in others, or heightened activity can be seen on one side in the thyroarytenoid muscle.

In vocal tremor, a variety of muscles can be involved.[137] Most often the thyroarytenoid is affected, but other muscles, including the strap muscles, can be affected, including the thyrohyoid, the sternothyroid, and, in some patients with abductor tremor, either the cricothyroid or PCA or both.

By using a concentric EMG electrode connected to an amplifier with a dual-channel storage oscilloscope, one channel for the EMG and another for the speech waveform, at a slow sweep speed, quick identification can be made of the muscles having spasms during voice breaks. Descriptions of EMG techniques are provided in greater detail in Hirano and Ohala.[138]

For the thyroarytenoid and lateral cricoarytenoid muscles, items to be examined for spasmodic bursts concurrent with voice breaks include prolonged "ee," repeated "ee," and all voiced sentences, such as "We mow our lawn all year." When examining abductor SD, while recording from the cricothyroid or PCA, speech should include "see-see-see," "pea-pea-pea," "he-he-he," "Kathy took a potato," and "Keep Tom at the party." Muscles with bursts of activity before and during voice breaks can be considered for injection with botulinum toxin injection.

Treatment. In the last 15 years, the following treatments have been used for managing symptoms in adductor SD.

Voice Therapy. This can assist mildly affected patients but may not have long-lasting effects. When uncertain of the diagnosis of a patient's voice disorder, a trial of voice therapy is recommended. Usually within three sessions, a speech pathologist experienced in voice therapy will report whether voice therapy might be beneficial for a patient. Voice therapy may also be helpful after botulinum toxin injection by prolonging the benefit period.[139]

Inhalation Voice Therapy. Speaking on inhalation is sometimes helpful in mild to moderate cases but may not have long-lasting effects. This technique is difficult for some patients to learn given that they often have respiratory control problems in addition to their voice problems.

Recurrent Laryngeal Nerve Section. Dedo first described unilateral removal of a 1 cm segment of the RLN below the thyroid isthmus for the treatment of adductor SD in 1976.[140] The procedure resulted in an initial dramatic reduction or elimination of voice spasms, but symptoms may recur in up to 64% of patients.[141] Recurrence of symptoms has been attributed to reinnervation of the operated side based on EMG and histologic studies. Those patients failing surgery may still respond to botulinum toxin injections.[142] Owing to the high rate of symptom recurrence, the permanent alteration in laryngeal function, and the requirements for subsequent medialization surgery in some patients with excessive breathiness, RLN section is less preferred to botulinum toxin injections at the present time.

Recurrent Laryngeal Nerve Avulsion. Netterville and his colleagues proposed a more extensive removal of the recurrent nerve in an effort to reduce the risk of reinnervation into the distal stump of the sectioned RLN.[143] After mobilization of the ipsilateral thyroid lobe, all branches of the RLN are identified, traced, and avulsed from their muscular insertions deep to the cricopharyngeus muscle. The total length of RLN removed averages 9 cm, in contrast to an average of 2 cm reported in previous RLN section studies. Long-term follow-up (3 to 7 years) on 18 patients following RLN avulsion revealed that 16 patients (89%) were free of SD symptoms at 3-

year follow-up.[144] Two of these patients, however, later developed recurrent spasms following medialization laryngoplasty for treatment of weak voice. Thus, an overall success rate of 78% (14/18) was reported for the series, which compares favorably to long-term recurrence rates for RLN section.[141] Weed et al recommended that RLN avulsion be reserved for patients who do not benefit from or do not tolerate botulinum toxin injections or have failed prior RLN section.[144]

Selective Laryngeal Adductor Denervation-Reinnervation. Berke et al published preliminary results with selective bilateral denervation of the laryngeal adductor muscles in an attempt to achieve a permanent bilateral adductor weakness that mimics the transient effects of botulinum toxin.[145] The new procedure involves bilateral section of the adductor RLN branches to the thyroarytenoid and lateral cricoarytenoid muscles with intentional reinnervation of the proximal thyroarytenoid branches using branches of the ansa cervicalis nerve. The aim of this directed reinnervation is to prevent unwanted reinnervation by RLN efferents and to preserve adductor muscle tone.

The study presented 1- to 5-year follow-up on 21 patients with adductor SD who underwent the procedure. In general, the symptoms and overall severity of SD improved from moderate to severe ratings preoperatively to mild to absent ratings postoperatively. All patients experienced severe vocal fold bowing and breathiness in the early postoperative period that improved after 3 to 6 months. Aspiration of greater than 2 weeks duration was noted in two patients, one of whom required hospitalization for aspiration pneumonia. Additional treatments (botulinum toxin, voice therapy, collagen injection, thyroarytenoid myotomy) were also performed in several patients to enhance the postoperative voice result. As noted by the authors, long-term follow-up in a larger patient population will be required to define the role of this procedure in the treatment of adductor SD. In particular, the effects of bilateral denervation on airway protection and fine motor control of the larynx (pitch variability, loudness range) require further careful assessment. Because the procedure is technically difficult and produces permanent structural and functional changes in the larynx, it should be reserved for patients with relatively severe disease. However, the authors reported promising results with normal conversational voice intensity and good inflection in most adductor SD patients treated thus far.

Thyroplasty. This surgical approach pushes the thyroid keel posteriorly, thus shortening and adducting the vocal folds.[146] In most cases, this can be detrimental to the voice, resulting in pitch lowering. This surgery cannot be easily reversed owing to the long-term effects of fibrosis and scarring. Because other more successful approaches are available, it is not recommended for patients with adductor SD.

Botulinum Toxin. This approach was developed in 1987[147] and has been evaluated using unilateral or bilateral thyroarytenoid injections in adductor SD.[148] The results seem excellent, with restoration of normal voice in 70% of patients for up to 3 months, followed by gradual symptom return within 4 to 5 months. Because the effect is temporary, it can be used on a trial basis in certain muscles in a particular patient. At 3-month intervals, injections can be administered in different muscle combinations to identify the correct muscles and dosage to be used for a particular patient.

The most commonly used method of injection is the percutaneous EMG-guided approach through the cricothyroid membrane.[149] The injection needle is Teflon coated, except for the bared tip, which serves as a monopolar EMG electrode when the bared hub is connected to a physiologic amplifier. This allows monitoring for a characteristic physiologic signal to determine that the needle tip is located in muscle before injecting. However, because the monopolar electrode has a large recording field and may pick up signals from any of the laryngeal adductor muscles in the region, such monitoring is not necessarily specific for the targeted thyroarytenoid muscles.[7]

To approach the thyroarytenoid percutaneously, the needle is angled superiorly and slightly laterally to approximate the mediolateral aspect of the vocalis portion of the TA muscle.[138] However, the needle tip can easily be placed more posteriorly than intended, medial to the lateral cricoarytenoid muscle and medial-superior to the cricothyroid muscle.[150] This can be seen in coronal sections of the adult human larynx.[151] From this, it is apparent that posteriorly placed injections may denervate portions of the lateral cricoarytenoid and cricothyroid muscles, in addition to the thyroarytenoid,

either by direct injection of these muscles or by diffusion across fascial planes.[152] This could account for the variability in response that is often seen (a) among patients and (b) in the same patient from consecutive EMG-guided injections several months apart, even though the same technique and dosages are used.

Two approaches have been used: large unilateral injections producing unilateral vocal fold paralysis[153] or small bilateral injections not altering the range of vocal fold movement on either side.[154] The advantage of the unilateral injection technique is that muscles on only one side of the larynx are affected by botulinum toxin. The long-term effects of repeated needle insertion and repeated denervation by botulinum toxin on laryngeal muscles are as yet unknown. Muscle biopsies in blepharospasm patients treated repeatedly with botulinum toxin injection have demonstrated significant muscle scarring and atrophy.[155] Similar changes in both thyroarytenoid muscles could have long-term effects on airway protection. The disadvantage of the unilateral injection approach is that a large dosage of toxin must be administered and that immobility of the injected fold often results.[153]

The advantage of the bilateral approach is that much smaller dosages of botulinum toxin can be used, reducing the cost. Further, effective symptom control can be obtained without any observable change in range of motion. Partial denervation seems to be effective in reducing the degree of hyperadduction, resulting in less interference with phonation. The disadvantages are that injections and needle insertions are performed on both sides of the larynx. With repeated injections every 4 to 6 months over many years, damage might result to both sides of the larynx, reducing airway closure and protection during swallowing.[156] Another difficulty with the bilateral approach is that if symptom control is not obtained with the usual dosage of 2.5 U on each side, further bilateral injections have the risk of causing bilateral adductor paralysis. With unilateral injections, repeated injection at higher dosages on the same side can be administered without danger of aspiration.

Other injection techniques that use visualization of the vocal folds rather than EMG to guide placement of the injection needle are the indirect laryngoscopic peroral method[157] and the transcartilaginous method.[158] Recently, the availability of the flexible fiber-optic nasolaryngoscope with a working channel has made possible an endoscopic method of injecting the thyroarytenoid muscle using a flexible catheter needle.[159] The usual dosages used by each technique are presented in Table 53–5.

Within 5 days of thyroarytenoid botulinum toxin injection, the number of breaks in phonation on vowels is reduced, the presence of voice roughness or hoarseness is reduced on spectrographic analyses, and speech rate is increased.[153] Patients' subjective reports include a reduction in speech effort and tension, reduced number of breaks, and greater voice control.[160] Two major side effects are reported: breathiness and swallowing difficulties. These side effects do not occur in all patients but can be significant in males following bilateral injections.[160]

Caution should be exercised in the use of botulinum toxin, particularly in undiagnosed patients with functional voice disorders. In such patients, the injections should be as small as possible (see Table 53–5) and combined with voice training. The injections should be only in those muscles demonstrated to be affected through a thorough EMG study during patient performance of affected and unaffected tasks.

Botulinum toxin injections have been employed in abductor SD in both the cricothyroid muscles[123] and the PCA.[161] Only about two-thirds of persons with abductor SD are benefited and to a lesser degree than those with adductor SD following thyroarytenoid muscle injections.[160]

The effects of vocal tremor on voice production can also be reduced by botulinum toxin injection[160]; however, the outcome is less predictable, and not all persons with vocal tremor are benefited.[162,163] In those patients with tremor affecting only the vocal folds and involving just the thyroarytenoid muscles, this treatment can be as effective as in adductor SD. However, many patients with this disorder have involvement of the strap muscles and have limited benefit even when these additional muscles are injected.

Paradoxical Vocal Fold Movement Paradoxical vocal fold movement is the adduction of the vocal folds during the inspiratory phase of respiration producing either a complete stoppage of air or stridor. These patients have normal vocal fold movement during speech, however, and often can inspire

TABLE 53–5. Dosages Used for Unilateral and Bilateral Injection of the Thyroarytenoid Muscle in Adductor Spasmodic Dysphonia

Unilateral/Bilateral	*EMG Guided*[149] *(U)*	*Injection Technique Peroral*[157] *(U)*	*Transcartilaginous*[158] *(U)*	*Endoscopic*[159] *(U)*
Unilateral	15[153]	2.5		6
Bilateral	2.5 each[154]	2.5	2 each	2 each
		+ 3 opposite		
Total	5	5.5	4	4

Superscript numbers indicate references.
EMG = electromyography.

between phrases while speaking. There are several different categories of patients with these disorders: (1) as an idiopathic focal dystonia or part of Meige's syndrome,[164] (2) associated with or masquerading as asthma,[165] (3) exercise-induced stridor,[166] (4) psychogenic,[167,168] and (5) associated with gastroesophageal reflux.[169]

These movement disorders are very rare but have become increasingly recognized in the last 5 years.[170,171] Most patients were previously diagnosed as psychogenic,[165] although it is becoming increasingly apparent that some are not.[164,172] The first presentation of this as a focal laryngeal dystonia specific to inspiration suggested that in some patients this is not psychogenic but rather caused by involuntary spasmodic bursts in the thyroarytenoid muscle.[164] Patients with a focal dystonia report that the symptoms are not present during sleep but that the symptoms are often present during the daytime, when they may become exacerbated by exercise and stress. Treatment with botulinum toxin is difficult because swallowing problems can be experienced as a result.[173] Some success has been reported in selected cases not responsive to speech therapy, psychotherapy, and pharmacotherapy.[164,174]

Patients with Meige's syndrome, an orofacial dyskinesia, can develop pharyngeal spasms producing intermittent airway obstruction, either because of pharyngeal constriction or obstruction at the epiglottis. Often the cricopharyngeus muscle is overly active in such patients. Again, the symptoms are absent during sleep, almost always present during the day, but exacerbated by stress.

Patients with chronic asthma can have laryngeal vocal fold adduction during asthmatic attacks.[166,175] However, it is not known whether this is owing to psychological overlay, might be a learned response, or is a dystonic reaction following many years of use of inhalants for symptom management. Exercise-induced asthma has also been reported to have a laryngeal adduction component.[166]

Most cases in the literature of paradoxical vocal fold dysfunction causing stridor or episodes of airway obstruction have been reported as psychogenic. In such cases, the symptoms abate following psychotherapy and speech therapy.[176]

Koufman has suggested that gastroesophageal reflux is a common cause of paroxysmal laryngospasms.[169] He reported 12 patients who presented with intermittent, sudden-onset, noisy, obstructed breathing that some termed choking episodes. These attacks of stridor often occurred following a meal, after the start of exercise, or after bending over. Sometimes they occurred at night. Some had them up to twice a day, others much less frequently, only a few a year. These attacks usually lasted only a few minutes. All complained of symptoms of reflux such as a lump in the throat, chronic throat clearing, cough, intermittent hoarseness, and difficulty swallowing. He reported that all responded well to lifestyle modification and omeprazole 20 mg twice daily.

However, caution must be exercised in diagnosis and treatment of this group of disorders,[171] which can be life threatening.[170]

Disorders of Vocal Misuse

Disorders of vocal misuse are often thought of as psychogenic; however, these disorders are best conceptualized as disorders of muscle misuse.[177]

Muscular Tension Dysphonia Muscular tension dysphonia is a complex first described by Morrison and colleagues[177,178] as dysphonia resulting from increased muscular tension in the larynx and neck associated with (1) palpably increased phonatory muscle tension in the paralaryngeal and suprahyoid muscles, (2) elevation of the larynx in the neck on increasing vocal pitch, (3) an open posterior glottic chink between the arytenoid cartilages on phonation, and (4) variable degrees of mucosal changes such as vocal nodules or chronic laryngitis. Morrison et al have now named this "laryngeal isometric" to denote the high level of generalized muscle tension in all of the laryngeal muscles.[177] These patients often have a high level of anxiety and may have learned an abnormal pattern of laryngeal muscle use. Aronson included these patients as psychogenic voice disorders and recommended laryngeal maneuvering, progressive relaxation, and voice training, going through several stages from simple to more complex voice production.[122] Roy and colleagues have shown that manual laryngeal manipulation is an effective intervention for this disorder.[179]

Other variants of vocal misuse owing to muscle tension include lateral-to-medial hypercontraction, which produces a tense harsh voice with high laryngeal resistance. This posture of the larynx may be triggered by an infection or chronic reflux and is often associated with erythema or diffuse thickening of the mucosa. Morrison and colleagues have termed this "irritable larynx," suggesting that peripheral inflammation may be the basis for the disorder.[180] Poor vocal habits can be reversed by therapy working directly on changing vocal posture and technique. Fiber-optic videotaping can be an instructive feedback tool.

Supraglottic hyperconstriction is a third variant of muscular tension. Hyperadduction of the ventricular folds along with the vocal folds produces a squeaky voice with a great deal of effort. Very rarely, phonation is produced at the ventricular folds, but most often both the vocal folds and ventricular folds are involved in voice production.

Anteroposterior supraglottic contraction produces a "pinhole"-type larynx in the most severe form with no voice or an anteroposterior shortening in milder forms, which results in a rough, hoarse, low-pitched voice.

The etiology of these misuse disorders is unknown; rarely are these patients' voice production abnormalities helped by psychological counseling alone. Usually, they respond to voice re-education and life management counseling to change their patterns of voice use.

The distinction between these patients and adductor SD patients is the degree to which the abnormal laryngeal posture is used consistently for voice production (see Table 53–4). In muscular tension dysphonia and other variants of vocal misuse, the abnormal vocal fold positioning seen on fiber-optic videoendoscopy is consistent regardless of the type of speech. In the SDs, the abnormalities are intermittent, depending on the voicing gesture required, and contain irregular rapid adductor or abductor movements causing intermittent voice breaks on particular types of speech sounds, vowels in adductor SD and prolonged voiceless consonants in abductor SD. In benign essential tremor, these intermittent voice breaks are regular, usually around 5 Hz. However, in some of the more extreme forms of SD and/or muscular tension dysphonia, the differentiation can be difficult because the spasmodic involuntary movement abnormalities may be difficult to identify when patients have extreme muscular tension. Whenever a possibility of either diagnosis exists, a trial of voice therapy by a speech pathologist experienced in treatment of these disorders is recommended before considering botulinum toxin. If botulinum toxin is employed, a combined regimen of voice therapy following injection can be helpful in training patients to eliminate those abnormal postures that may be contributing to their voice disorder in addition to the spasmodic involuntary movements.

Voice Fatigue Syndrome Voice fatigue syndrome is diagnosed only after other disorders have been excluded, including myasthenia gravis, Parkinson's disease, amyotrophic lateral sclerosis, and muscle atrophy associated with aging, Epstein-Barr virus, and sulcus vocalis. When a misuse disorder is present, patients use excessive muscular tension in one muscle group that interferes with voice production for extended periods. Often patients must use their voice for authority or change their use for special demands such as classroom teaching. Patients then complain about their loss of voice and fatigue by the end of the day. Fiber-optic videotaping can be beneficial in retraining normal voice production.

Voice fatigue can also be associated with aging when an apparent loss of muscle bulk can become

apparent. If neurogenic disease can be excluded, then a regimen of increased physical exercise may be beneficial in increasing vocal function.

Abnormal Loudness Abnormal loudness can be identified as a misuse disorder when patients employ increased and excessive intensity, often when they have experienced a rapid diminution in their hearing or that of a family member. These patients may have assumed abnormal voice production patterns with excessive tension similar to the variants of muscular tension dysphonia.

Abnormal Pitch The continuation of a high pitch sometimes occurs in males following pubescent voice change. It can also be referred to as puberphonia or mutational falsetto. The misuse of a tense elongated vocal fold position during voicing can be seen. Often when this condition continues into adulthood, it is difficult to reverse because the misuse has become part of the voicing gesture. Laryngeal maneuvering can sometimes be successful in reducing the abnormal posturing.[122,177] The combination of an initial treatment with botulinum toxin into the cricothyroid muscles in association with voice retraining can be beneficial in the most resistant cases.[181]

False Vocal Fold Phonation (Dysphonia Plicae Ventricularis) This was described above as a variant of muscular tension dysphonia in which there is excessive lateral-to-medial approximation. Phonation involving only the ventricular folds is extremely rare and usually occurs following injury to the true vocal folds from intubation, radiation, vocal fold atrophy owing to vocal fold paralysis, or surgery for carcinoma. In most functional cases, hyperadduction of both the glottis and ventricular folds is involved. When this condition occurs with functional vocal folds, it is best treated using fiber-optic videoendoscopy to demonstrate gestures such as throat clear, humming, and sighing when glottal phonation is most likely independent of ventricular adduction.

Psychogenic Voice Disorders

These voice disorders are related to psychological processes in the patient resulting in inappropriate use of a normal vocal mechanism for speech communication. A psychogenic origin should be considered only when all other possible structural, neurogenic, or functional voice disorders have been excluded. Sometimes patients can have a psychological history similar to that usually found in patients with a psychogenic voice disorder but have an unrelated voice disorder, for example, a neurogenic disorder. Many of the other functional disorders, such as the SDs, were considered psychogenic voice disorders until recently. Only the behavioral characteristics are available for differentiating among these disorders.[129,182–186] In general, psychogenic voice disorders do not produce chronic habitual abuse or overuse of the peripheral vocal mechanism and rarely produce secondary peripheral tissue abnormalities. Sometimes the abnormal voice production can be intermittent with normal periods of voice production.

At least five types of psychogenic voice disorders can be identified and are described below: conversion reaction dysphonia, malingering, psychogenic dysphonia, elective mutism, and psychological overlay. Patients with vocal misuse are sometimes classified as having psychogenic voice disorders, and successful management can include vocal retraining alone. In psychogenic voice disorders, management is most successful when vocal retraining, psychological or psychiatric counseling, and life situation change are simultaneously employed; otherwise, the voice symptoms may return intermittently.

Most psychogenic dysphonias are hypofunctional vocal behaviors, "whispering" or "aphonic" dysphonia, but they may also appear as falsetto or hyperfunctional behaviors or as a variable mix of the two.

Conversion Reaction Dysphonia The essential feature of conversion reaction dysphonia is the experience of severe psychological or emotional trauma that is closely associated in time with the onset of the dysphonia.[187] The traumatic experience may be concrete and easily identified, such as the loss of a loved one or the experience of a direct threat to the patient or a loved one, or it may be more subtle and difficult to identify, such as the experience of psychological or physical abuse or the patient's fear of losing a loved one.

Voice quality in conversion reaction dysphonia is typically, but not exclusively, hypofunctional, or aphonic.[188] Sometimes the voice can be tremorous as

well.[189] Laryngeal function for vegetative purposes (eg, swallowing, coughing) is within normal limits. Fiberoptic evaluation of the larynx is remarkable for lack of vocal fold adduction during attempts to produce voicing for speech without evidence of vocal fold paresis, paralysis, arytenoid fixation, or other movement difficulties. Vocal fold adductions for coughing and Valsalva's maneuver are within normal limits.

In some cases, conversion reaction dysphonia occurs as hyperfunctional dysphonia and may accompany complaints regarding inconsistent, mild respiratory distress. Conversion disorders most frequently affect young women in their twenties and thirties. Vocal symptoms appear in their final form rather than gradually increasing in severity.

Because diagnosis requires the confirmation of a sudden onset of vocal symptoms closely associated with the experience of an event that may have caused severe psychological trauma, the patient's history, provided either by the patient or others in the patient's environment, is essential.[190] In some cases, the patient is aware of the event and talks about it freely. Sometimes, however, extensive interview and counseling may be required before the patient may be willing or able to talk about the event directly. Interviews with close relatives or friends may be important in such cases, along with referrals for professional psychological evaluation and counseling.

In hypofunctional cases, weak, breathy voice quality with complete or near-complete absence of vocal fold contact or vibratory activity is common. In hyperfunction, however, vocal fold tension is increased such that vocal fold vibration for speech is difficult or impossible to achieve. Such tension is frequently associated with palpable tension in the extrinsic laryngeal muscles.[177] In some cases, symptoms may be more noticeable during formal testing situations than during informal communication.

Malingering Dysphonia Although the presentation is very similar to conversion reaction dysphonia in symptoms, malingering dysphonia involves a conscious intent on the part of the patient to simulate a voice disorder for some gain, often financial or emotional. The patient may be pursuing a medical malpractice suit, an injury or accident case, a legal redress against an employer, or special disability assistance. Other agendas may be less concrete although no less motivating for the patient, such as attempting to avoid communication with an individual or attempting to avoid an experience that the patient considers objectionable. A carefully obtained, complete patient history and interview are the most useful tools for identifying a malingering dysphonia from conversion reaction dysphonia.

Voice quality in malingering dysphonia is usually hypofunctional or aphonic for speech communication purposes, although hyperfunctional or variable cases may be found. Voice quality associated with vegetative functions such as coughing or throat clearing is usually normal. No evidence of laryngeal physiologic abnormality is present. Cognitive distracters, such as counting backwards in sevens or reading difficult material, are often important in eliciting voice production inconsistencies.

Psychogenic Dysphonia In most psychogenic dysphonias, patients exhibit inconsistent voice symptoms, such as intermittent aphonia, abnormal vocal pitch, or periods of severe or markedly abnormal voice may be interspersed with periods of normal voice. These symptoms may be in reaction to psychosocial stresses either in the intermediate or distant past. Symptoms range from a noticeable voice abnormality in pitch to stuttering-like voice breaks to complete aphonia disrupting all attempts to communicate. Often patients express a lack of responsibility for their voice problem and poor insight into their voice difficulties. These patients usually have a history of frequent somatic complaints and are unable to take responsibility for their life situations. They may respond to symptom-oriented voice therapy early in their course or may become resistant to voice therapy and require psychosocial management. They often become dependent on others for communication; in fact, the voice abnormalities may be used unconsciously by the patient to avoid certain social or occupational situations that they find stressful.[188]

Diagnosis may depend on identifying significant psychosocial trauma or physical/sexual abuse, a dissociative personality, or a history of frequent somatizations. An inconsistency of the voice symptoms is often noted by others in the patient's environment. Often the voice symptoms become less exaggerated during distraction such as reading difficult material orally and counting backwards or during examinations such as fiber-optic nasolaryngoscopy. The disorder is more often seen in females who are unable to control their environment.

In cases of aphonia, the patient does not adduct vocal folds for speech but has normal Valsalva's maneuver and swallow. Although patients may have secondary gains from their disorder, they are usually not conscious of the secondary gain. Symptoms may range from periods of intermittently abnormal pitch or hoarseness with strained voice to a consistent aphonia with an absence of vocal fold adduction only for speech.

Elective Mutism Although rare, this usually appears in young children immediately following a traumatic event or in extreme anger. It is characterized by a complete withholding of oral communication in the presence of normal laryngeal structure and function and language skills. It frequently occurs with psychosocial problems relating to the family situation (abuse, conflicts within the family or with the community). This appears most often at preschool or kindergarten age, and the child becomes mute on entering school. It may occur at other times during the school-age years in response to traumatic events and sometimes in adults. Without intervention, it can persist for several years. The withholding of vocal communication is related to issues of control, emotional disengagement, manipulation, secondary gain, and low self-esteem and indicates maladaptive psychosocial development. The family may have a history of frequent somatizations, hostility toward school or community, and disturbed interpersonal relationships.

The child may be described as shy and cooperative and attentive in class. Normal results on tests of receptive language, reading, writing, and other school-related abilities are found, as are normal results when the structure and function of the larynx are visualized. The parents may minimize the problem, suggesting that there is no difficulty at home and that the child does not like to talk at school. The child may have a continuous and consistent refusal to talk in almost all situations (eg, school, school bus, playground) but demonstrates the ability to comprehend spoken language and read and speak appropriately when not observed. A team approach including voice therapy, psychosocial counseling, and intervention is beneficial.

Psychogenic Overlay Frequently, patients have complex voice disorders; they may develop emotional complaints as the result of having a voice disorder, which may then appear psychogenic in origin,[191] or they may either consciously or unconsciously exaggerate their symptoms to draw attention, receive sympathy, or receive further secondary gains resulting from their voice disorder. Management is most effective by combining appropriate treatment of the primary voice disorder (phonosurgery, voice therapy, pharmacotherapy) with counseling to emphasize the advantages of recovery.

Conclusions

Because of the complexity and diagnostic uncertainty of most of the functional voice disorders, caution must be exercised. Usually, the disorder is not acute or life threatening, except in the case of acute airway obstruction in paradoxical vocal fold adduction. Therefore, a trial of voice therapy is often warranted before embarking on phonosurgery, botulinum toxin injections, pharmacologic management techniques, or psychological counseling.

REFERENCES

1. Baken RJ. Clinical measurement of speech and voice. Boston: College-Hill Press; 1987.
2. Hirano M. Clinical examination of voice. New York: Springer-Verlag Wien; 1981.
3. Rabinov CR, Kreiman J, Gerratt BR, Bielamowicz S. Comparing reliability of perceptual ratings of roughness and acoustic measure of jitter. J Speech Hear Res 1995;38:26–32.
4. Bless DM, Hirano M, Feder R. Videostroboscopic examination of the larynx. Ear Nose Throat J 1987; 66:289–96.
5. Stone M, Shawker T, Talbot TRA. Cross-sectional tongue shape during vowel production. J Acoust Soc Am 1988;83:1586–96.
6. Baer T, Gore JC, Gracco LC, Nye PW. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. J Acoust Soc Am 1991; 90:799–828.
7. Ludlow CL. Neurophysiological assessment of patients with vocal motor control disorders. In: Cooper JA, editor. Assessment of speech and voice production: research and clinical applications. Bethesda (MD): National Institute on Deafness and Other Communication Disorders Monograph; 1991. p. 161–71.

8. Dyck PJ, Litchy WJ, Minnerath S, et al. Hereditary motor and sensory neuropathy with diaphragm and vocal cord paresis. Ann Neurol 1994;35: 608–15.
9. Hanson DG, Ludlow CL, Bassich CJ. Vocal fold paresis in Shy-Drager syndrome. Ann Otol Rhinol Laryngol 1983;92:85–90.
10. Titze IR, Luschei ES, Hirano M. Role of the thyroarytenoid muscle in regulation of fundamental frequency. J Voice 1989;3:213–24.
11. Duvoisin RC. The cause of Parkinson's disease. Movement disorders. London: Butterworth; 1982. p. 2.
12. Gasser T. Autosomal-dominantly inherited forms of Parkinson's disease. J Neural Transm Suppl 2000; 31–40.
13. Hanson DG, Gerratt BR, Ward PH. Cinegraphic observations of laryngeal function in Parkinson's disease. Laryngoscope 1984;94:348–53.
14. Logemann JA, Fisher HB, Boshes B, Blonsky ER. Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. J Speech Hear Disord 1978;43:47–57.
15. Ramig LO. Speech therapy for patients with Parkinson's disease. In: Therapy of Parkinson's disease. New York: Marcel Dekker; 1994.
16. Ramig LO, Countryman S, O'Brien C, et al. Intensive speech treatment for patients with Parkinson's disease: short- and long-term comparison of two techniques. Neurology 1996;47:1496–504.
17. Berke GS, Gerratt B, Kreiman J, Jackson K. Treatment of Parkinson hypophonia with percutaneous collagen augmentation. 1999;109:1295–9.
18. Steele JC, Richardson JC, Olszeweski J. Progressive supranuclear palsy. A heterogenous degeneration involving the brainstem, basal ganglia and cerebellum with vertical gaze, pseudobular palsy, nuchal dystonia and dementia. Arch Neurol 1964;10:333.
19. Jankovic J, Tolosa E, editors. Parkinson's disease and movement disorders. Baltimore: Williams and Wilkins; 1993.
20. Oppenheimer DR. Diseases of the basal ganglia, cerebellum and motor neurons. In: Greenfields' neuropathology. London: Edward Arnold; 1984. p. 700.
21. Darley FL, Aronson AE, Brown JR. Motor speech disorders. Philadelphia: WB Saunders; 1975.
22. Bassich CJ, Ludlow CL, Polinsky RJ. Speech symptoms associated with early signs of Shy Drager syndrome. J Neurol Neurosurg Psychiatry 1984;47:995–1001.
23. Lechtenberg R. Mutiple sclerosis fact book. Philadelphia: FA Davis; 1983.
24. Rontal E, Rontal M, Wald J, Rontal D. Botulinum toxin injection in the treatment of vocal fold paralysis associated with multiple sclerosis: a case report. J Voice 1999;13:274–9.
25. Obeso JA, Artieda J, Mardsen CD. Different clinical presentations of myoclonus in Parkinson's disease and movement disorders. In: Jankovic J, Tolosa F, editors. Parkinson's disease and movement disorders. Baltimore: Williams and Wilkins; 1984.
26. Lapresle J. Palatal myoclonus. Adv Neurol 1986; 43:265–73.
27. Bryce GE, Morrison MD. Botulinum toxin treatment of essential palatal myoclonus tinnitus. J Otolaryngol 1998;27:213–6.
28. Marsden CD, Hallet M, Fahn S. The nosology and patho-physiology of myoclonus. In: Marsden CD, Fahn S, editors. Movement disorders. London: Butterworth; 1982. p. 196–248.
29. Deuschl GA, Ebner A, Hanners R, Luching CH. Differences of cortical activation in spontaneous and reflex myoclonus. Electroencephalogr Clin Neurophysiol 1991;80:326–8.
30. Rose FC. Motor neuron disease. New York: Grune and Stratton; 1977.
31. Job A, Raman R, Gnanamuthu C. Laryngeal stridor in myasthenia gravis. J Laryngol Otol 1992;106:633–4.
32. Kluin KJ, Bromberg MB, Feldman EL, Simmons Z. Dysphagia in elderly men with myasthenia gravis. J Neurol Sci 1996;138:49–52.
33. Goodhart SP, Davison C. Syndrome of the posterior inferior and anterior inferior cerebellar arteries and their branches. Arch Neurol Psychiatry 1936;35:501.
34. Driscoll BP, Gracco C, Coelho C, et al. Laryngeal function in postpolio patients. Laryngoscope 1995; 105:35–41.
35. Martensson A. The functional organization of the intrinsic laryngeal muscles. Ann N Y Acad Sci 1968; 2:381.
36. Hirano M, Bless DM. Videostroboscopic examination of the larynx. San Diego (CA): Singular Publishing; 1993.
37. Smitheran JR, Hixon TJ. A clinical method for estimating laryngeal airway resistance during vowel production. J Speech Hear Disord 1981;46:138–46.
38. Titze IR. Phonation threshold pressure: a missing link in glottal aerodynamics. J Acoust Soc Am 1992; 91:2926–35.
39. Kotby MN. Electromyography and neurography in neurolaryngology. J Voice 1991;5:316.

40. Thumfart F. From larynx to vocal ability: new electrophysiological data. Acta Otolaryngol (Stockh) 1988;105:425–31.
41. Logemann JA. Treatment for aspiration related to dysphagia: an overview. Dysphagia 1986;1:34–8.
42. Sasaki CT. The effect of tracheostomy on the laryngeal closure reflex. Dysphagia 1977;87:1428–33.
43. Nash M. Swallowing problems in the tracheotomized patient. Otolaryngol Clin North Am 1988;21:701–9.
44. Heuer RJ, Sataloff RT, Emerich K, et al. Unilateral recurrent laryngeal nerve paralysis: the importance of "preoperative" voice therapy. J Voice 1997;11:88–94.
45. Rosen CA. Phonosurgical vocal fold injection: indications and techniques. Oper Tech Otolaryngol Head Neck Surg 1998;9:203–9.
46. Arnold GE. Vocal rehabilitation of the paralytic dysphonia. VIII. Phonatric methods of vocal compensation. Arch Otolaryngol 1962;76:76–83.
47. Benjamin B, Robb P, Clifford A, Eckstein R. Giant Teflon granuloma of the larynx. Head Neck 1991; 13:453–6.
48. Varvares MA, Montgomery WW, Hillman RE. Teflon granuloma of the larynx: etiology, pathophysiology, and management. Ann Otol Rhinol Laryngol 1995; 104:511–5.
49. Dedo HH. Injection and removal of Teflon for unilateral vocal fold paralysis. Ann Otol Rhinol Laryngol 1992;101:81–6.
50. Ossoff RH, Koriwchak MJ, Netterville JL, Duncavage JA. Difficulties in endoscopic removal of Teflon granulomas of the vocal fold. Ann Otol Rhinol Laryngol 1993;102:405–12.
51. Coleman JR, Miller FR, Netterville JL. Teflon granuloma excision via a lateral laryngotomy. Oper Tech Otolaryngol Head Neck Surg 1999;10:29–35.
52. Schramm VL, May M, Lavorato AS. Gelfoam paste injection for vocal cord paralysis: temporary rehabilitation of glottic incompetence. Laryngoscope 1978;88:1268–73.
53. Mikaelian DO, Lowry LD, Sataloff RT. Lipoinjection for unilateral vocal cord paralysis. Laryngoscope 1991;101:465–8.
54. Shaw GY, Szewczyk MA, Searle J, Woodroof J. Autologous fat injection into the vocal folds: technical considerations and long-term follow-up. Laryngoscope 1997;107:177–86.
55. Mikus JL, Koufman JA, Kilpatrick SE. Fate of liposuctioned and purified autologous fat injections in the canine vocal fold. Laryngoscope 1995;105:17–22.
56. Ford CN, Bless DM. Clinical experience with injectable collagen for vocal fold augmentation. Laryngoscope 1986;96:863–9.
57. Ford CN, Bless DM, Loftus JM. Role of injectable collagen in the treatment of glottic insufficiency: a study of 119 patients. Ann Otol Rhino Laryngol 1992;101:237–47.
58. Ford CN. Advances and refinements in phonosurgery. Laryngoscope 1999;109:1891–900.
59. Ford CN, Staskowski PA, Bless DM. Autologous collagen vocal fold injections: a preliminary clinical study. Laryngoscope 1995;105:944–8.
60. Remacle M, Lawson GKJ, Jamart J. Use of autologous collagen for correcting glottic gaps: initial results. J Voice 1999;13:280–8.
61. Tsunoda K, Takanosawa M, Niimi S. Autologous transplantation of fascia into the vocal fold: a new technique for glottal incompetence. Laryngoscope 1999;109:504–8.
62. Wanamaker JR, Netterville JL, Ossoff RH. Phonosurgery: Silastic medialization for unilateral vocal fold paralysis. Oper Tech Otolaryngol Head Neck Surg 1993;4:207–17.
63. Ford CN. Intraoperative decision making in medialization laryngoplasty. Oper Tech Otolaryngol Head Neck Surg 1999;10:17–21.
64. McCulloch TM, Hoffman HT. Medialization laryngoplasty with expanded polytetrafluoroethylene: surgical technique and preliminary results. Ann Otol Rhinol Laryngol 1998;107:427–32.
65. Hoffman HT, McCulloch TM. Medialization laryngoplasty with Gore-Tex. Oper Tech Otolaryngol Head Neck Surg 1999;10:6–8.
66. Koufman JA, Postma GN. Revision laryngoplasty. Oper Tech Otolaryngol Head Neck Surg 1999;10: 61–5.
67. Isshiki N, Tanabe M, Sawada M. Arytenoid adduction for unilateral vocal cord paralysis. Arch Otolaryngol 1978;104:555–8.
68. Netterville JL, Stone RE, Luken ES, et al. Silastic medialization and arytenoid adduction: the Vanderbilt experience. Ann Otol Rhinol Laryngol 1993; 102:413–24.
69. McCulloch TM, Hoffman HT, Andrews BT, Karnell MP. Arytenoid adduction combined with Gore-Tex medialization thyroplasty. Laryngoscope 2000;110: 1306–11.
70. Zeitels SM, Hochman I, Hillman RE. Adduction arytenopexy: a new procedure for paralytic dysphonia

with implications for implant medialization. Ann Otol Rhinol Laryngol 1998;107:2–24.
71. Zeitels SM. Adduction arytenopexy with medialization laryngoplasty: a new approach to paralytic dysphonia. Oper Tech Otolaryngol Head Neck Surg 1999;10:9–16.
72. Zeitels SM, Hillman RE, Desloge RB, Bunting GA. Cricothyroid subluxation: a new innovation for enhancing the voice with laryngoplastic phonosurgery. Ann Otol Rhinol Laryngol 1999;108:1126–31.
73. Tucker HM. Reinnervation of the unilaterally paralyzed larynx. Ann Otol Rhinol Laryngol 1977;86: 789–94.
74. Crumley RL, Izdebski K, McMicken B. Nerve transfer versus Teflon injection for vocal cord paralysis: a comparison. Laryngoscope 1988;98:11:1200–3.
75. Benninger MS, Crumley RL, Ford CN, et al. Evaluation and treatment of the unilateral paralyzed vocal fold. Otolaryngol Head Neck Surg 1994; 111:497–508.
76. Zitsch RP. Continuous positive airway pressure. Use in bilateral vocal cord paralysis. Arch Otolaryngol Head Neck Surg 1992;118:875–6.
77. Hirano M, Kurita S, Kiyokawa K, Sato K. Posterior glottis: morphological study in excised human larynges. Ann Otol Rhinol Laryngol 1986;95: 576–81.
78. Jackson C. Ventriculocordectomy: a new operation for the cure of goitrous glottic stenosis. Arch Surg 1922;4:257.
79. Moore J. Operative procedures in the treatment of stenosis of the larynx caused by bilateral paralysis of the abductor muscles. Proc R Soc Med 1923;16:32.
80. Hoover WB. Bilateral abductor paralysis, operative treatment of submucous resection of the vocal cord. Arch Otolaryngol 1932;15:337.
81. Lore J. A suggested operative procedure for the relief of stenosis in double abductor paralysis: an anatomic study. Ann Otol Rhinol Laryngol 1936;45:679.
82. Helmus C. Microsurgical thyrotomy and arytenoidectomy for bilateral recurrent laryngeal nerve paralysis. Laryngoscope 1972;32:491–503.
83. Singer MI, Hamaker RC, Miller SM. Restoration of the airway following bilateral recurrent laryngeal nerve paralysis. Laryngoscope 1985;95:1204–7.
84. Woodman D. A modification of the extralaryngeal approach to arytenoidectomy for bilateral abductor paralysis. Arch Otolaryngol Head Neck Surg 1946; 43:63.
85. Thornell WC. Intralaryngeal approach for arytenoidectomy in bilateral abductor vocal cord paralysis. Arch Otolaryngol 1948;47:505.
86. El Chazly M, Rifai M, El Ezz AA. Arytenoidectomy and posterior cordectomy for bilateral abductor paralysis. J Laryngol Otol 1991;105:454–5.
87. Ossoff RH, Sisson GH, Duncavage JA, et al. Endoscopic laser arytenoidectomy for the treatment of bilateral vocal cord paralysis. Laryngoscope 1984; 94:1293–7.
88. Ossoff RH, Duncavage JA, Shapshay SM, Krespi YP, Sisson GA. Endoscopic laser arytenoidectomy revisted. Ann Otol Rhinol Laryngol 1990;99:764–71.
89. Crumley RL. Endoscopic laser medial arytenoidectomy for airway management in bilateral laryngeal paralysis. Ann Otol Rhinol Laryngol 1993;102:81–4.
90. Benninger MS, Bhattacharyya N, Fried MP. Surgical management of bilateral vocal fold paralysis. Oper Tech Otolaryngol Head Neck Surg 1998;9:224–9.
91. Kirchner FR. Endoscopic lateralization of the vocal cords. Laryngoscope 1979;89:1779–83.
92. Geterud A, Ejnell H, Stenborg R, Bake B. Long-term results with a simple surgical treatment of bilateral vocal cord paralysis. Laryngoscope 1990;100:1005–8.
93. Linder A, Lindholm CE. Vocal fold lateralization using carbon dioxide laser and fibrin glue. J Laryngol Otol 1992;106:226–30.
94. Dennis DP, Kashima H. Carbon dioxide laser posterior cordectomy for treatment of bilateral vocal cord paralysis. Ann Otol Rhinol Laryngol 1989;98: 930–4.
95. Kashima H. Bilateral vocal fold motion impairment: pathophysiology and management by transverse cordotomy. Ann Otol Rhinol Laryngol 1991;100:717–21.
96. Doyle PJ, Brummet RE, Everts EC. Results of surgical section and repair of the recurrent laryngeal nerve. Laryngoscope 1967;77:1245–51.
97. Gordon JH, McCabe BF. The effect of accurate neurorrhaphy on reinnervation and return of laryngeal function. Laryngoscope 1968;78:236–50.
98. Crumley RL. Experiments in laryngeal reinnervation. Laryngoscope 1982;92 Suppl 30:1–27.
99. Rice DH. Laryngeal reinnervation. Laryngoscope 1982;92:1049–59.
100. Crumley RL. Phrenic nerve graft for bilateral vocal cord paralysis. Laryngoscope 1983;93:425–8.
101. Tucker HM. Long-term results of nerve-muscle pedicle reinnervation for laryngeal paralysis. Ann Otol Rhinol Laryngol 1989;98:674–6.

102. de Heyning PV, Sanders I, Goding GS, et al. Human trials of an implanted laryngeal pacer for bilateral vocal fold paralysis. Otolaryngol Head Neck Surg 1996;115:P49.
103. Hillel AD, Benninger M, Blitzer A, et al. Evaluation and management of bilateral vocal cord immobility. Otolaryngol Head Neck Surg 1999;121:760–5.
104. Eisele DW. Chronic aspiration. In: Cummings CW, Fredrickson JM, Harker LA, et al, editors. Otolaryngology-head and neck surgery. St. Louis: CV Mosby; 1998. p. 1989–2000.
105. Martens L, Cameron T, Simonsen M. Effects of a multidisciplinary management program on neurologically impaired patients with dysphagia. Dysphagia 1990;5:147–51.
106. Rontal E. Vocal cord injection in the treatment of acute and chronic aspiration. Laryngoscope 1976; 86:625–34.
107. Isshiki N, Okamura H, Ishikawa T. Thyroplasty type I (lateral compression) for dysphonia due to vocal cord paralysis or atrophy. Acta Otolaryngol (Stockh) 1975;80:465–73.
108. Montgomery WW. Surgery to prevent aspiration. Otolaryngol Head Neck Surg 1975;101:679–82.
109. Sasaki CT, Milmoe G, Yanagisawa E, et al. Surgical closure of the larynx for aspiration. Arch Otolaryngol Head Neck Surg 1980;106:422–3.
110. Habal MB, Murray JE. Surgical treatment of life-endangering chronic aspiration pneumonia: use of an epiglottic flap to the arytenoids. Plast Reconstr Surg 1972;49:305–11.
111. Vecchione TR, Habal MB, Murray JE. Further experiences with the arytenoid-epiglottic flap for chronic aspiration pneumonia. Plast Reconstr Surg 1975;55:318–23.
112. Strome M, Fried MP. Rehabilitative surgery for aspiration: a clinical analysis. Arch Otolaryngol Head Neck Surg 1983;109:809–11.
113. Biller HF, Lawson W, Baek SML. Total glossectomy: a technique of reconstruction eliminating laryngectomy. Arch Otolaryngol 1983;109:69–73.
114. Weisberger C, Huebsch SA. Endoscopic treatment of aspiration using a laryngeal stent. Otolaryngol Head Neck Surg 1982;90:215–22.
115. Eliachar I, Nguyen D. Laryngotracheal stent for internal support and control of aspiration without loss of phonation. Otolaryngol Head Neck Surg 1990;103:837–40.
116. Lindeman RC. Diverting the paralyzed larynx: a reversible procedure for intractable aspiration. Laryngoscope 1975;85:157–80.
117. Lindeman RC, Yarington CT, Sutton D. Clinical experience with the tracheoesophageal anastomosis for intractable aspiration. Ann Otol Rhinol Laryngol 1976;85:609–12.
118. Eisele DW, Yarrington CT, Lindeman RC, Larabee WR. The tracheoesophageal diversion and laryngotracheal separation procedures for treatment of intractable aspiration. Am J Surg 1989;157:230–6.
119. Broniatowski M. Dynamic control of the larynx and future perspectives in the management of deglutive aspiration. Dysphagia 1993;8:334–6.
120. Blitzer A, Brin MF, Fahn S, Lovelace RE. Clinical and laboratory characteristics of focal laryngeal dystonia: study of 110 cases. Laryngoscope 1988; 98:636–40.
121. Gates GA. Spasmodic dysphonia. J Voice 1992;6:293.
122. Aronson AE. Clinical voice disorders: an interdisciplinary approach. 2nd ed. New York: Thieme-Stratton; 1985.
123. Ludlow CL, Naunton RF, Terada S, Anderson BJ. Successful treatment of selected cases of abductor spasmodic dysphonia using botulinum toxin injection. Otolaryngol Head Neck Surg 1991;104:849–55.
124. Rosenbaum F, Jankovic J. Task-specific focal tremor and dystonia: categorization of occupational movememnt disorders. Neurology 1988;38:522–7.
125. Izdebski K, Dedo HH, Boles L. Spastic dysphonia: a patient profile of 200 cases. Am J Otolaryngol 1984; 5:7–14.
126. Blitzer A, Lovelace RE, Brin MF, et al. Electromyographic findings in focal laryngeal dystonia (spastic dysphonia). Ann Otol Rhinol Laryngol 1985;94: 591–4.
127. Dedo HH, Townsend JJ, Izdebski K. Current evidence for the organic etiology of spastic dysphonia. Otolaryngol Head Neck Surg 1978;86:875–80.
128. Hartman DE. Neurogenic dysphonia. Ann Otol Rhinol Laryngol 1984;93:57–64.
129. Aronson AE, Brown JR, Litin EM, Pearson JS. Spastic dysphonia: II. Comparison with essential (voice) tremor and other neurologic and psychogenic dysphonias. J Speech Hear Disord 1968; 33:219–31.
130. Aronson AE, Hartman DE. Adductor spastic dysphonia as a sign of essential (voice) tremor. J Speech Hear Disord 1981;46:52–8.
131. Schaefer SD. Neuropathology of spasmodic dysphonia. Laryngoscope 1983;93:1183–204.
132. Aronson AE, Peterson HW, Litin EM. Psychiatric symptomatology in functional dysphonia and aphonia. J Speech Hear Disord 1966;31:115–27.

133. Aronson AE. Importance of the psychosocial interview in the diagnosis and treatment of "functional" voice disorders. J Voice 1990;4:287–9.
134. Roy N, Bless DM, Heisey D. Personality and voice disorders: a multitrait-multidisorder analysis. J Voice 2000;14:521–48.
135. Watson BC, Schaefer SD, Freeman FJ, et al. Laryngeal electromyographic activity in adductor and abductor spasmodic dysphonia. J Speech Hear Res 1991;34:473–82.
136. Nash EA, Ludlow CL. Laryngeal muscle activity during speech breaks in adductor spasmodic dysphonia. Laryngoscope 1996;106:484–9.
137. Koda J, Ludlow CL. An evaluation of laryngeal muscle activation in patients with voice tremor. Otolaryngol Head Neck Surg 1992;107:684–96.
138. Hirano M, Ohala J. Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. J Speech Hear Res 1969;12:362–73.
139. Murry T, Woodson GE. Combined-modality treatment of adductor spasmodic dysphonia with botulinum toxin and voice therapy. J Voice 1995;9:460–5.
140. Dedo HH. Recurrent laryngeal nerve section for spastic dysphonia. Ann Otol Rhinol Laryngol 1976; 85:451–9.
141. Aronson AE, DeSanto LW. Adductor spasmodic dysphonia: three years after recurrent nerve section. Laryngoscope 1983;93:1–8.
142. Ludlow CL, Naunton RF, Fujita M, Sedory SE. Spasmodic dysphonia: botulinum toxin injection after recurrent nerve surgery. Otolaryngol Head Neck Surg 1990;102:122–31.
143. Netterville JL, Stone RF, Rainey C, et al. Recurrent laryngeal nerve avulsion for treatment of spastic dysphonia. Ann Otol Rhinol Laryngol 1991;100:10–4.
144. Weed DT, Jewett BS, Rainey C, et al. Long-term followup of recurrent laryngeal nerve avulsion for the treatment of spastic dysphonia. Ann Otol Rhinol Laryngol 1996;105:592–601.
145. Berke GS, Blackwell KE, Gerratt RR, et al. Selective laryngeal adductor denervation-reinnervation: a new surgical treatment for adductor spasmodic dysphonia. Ann Otol Rhinol Laryngol 1999;108:227–31.
146. Tucker HM. Laryngeal framework surgery in the management of spasmodic dysphonia: preliminary report. Ann Otol Rhinol Laryngol 1989; 98:52–4.
147. Miller RH, Woodson GE, Jankovic J. Botulinum toxin injection of the vocal fold for spasmodic dysphonia. Arch Otolaryngol Head Neck Surg 1987; 113:603–5.
148. Blitzer A, Brin MF. Laryngeal dystonia: a series with botulinum toxin therapy. Ann Otol Rhinol Laryngol 1991;100:85–9.
149. NIH consensus development conference statement: clinical use of botulinum toxin. Arch Neurol 1991; 48:1294–8.
150. Ludlow CL, Yeh J, Cohen LG, et al. Limitations of laryngeal electromyography and magnetic stimulation for assessing laryngeal muscle control. Ann Otol Rhinol Laryngol 1994;103:16–27.
151. Hirano M, Sato K. Histological color atlas of the human larynx. San Diego (CA): Singular Publishing Group; 1993.
152. George EF, Zimbler M, Wu B, et al. Quantitative mapping of the effect of botulinum toxin injections in the thyroarytenoid muscle. Ann Otol Rhinol Laryngol 1992;101:888–92.
153. Ludlow CL, Naunton RF, Sedory SE, et al. Effects of botulinum toxin injections on speech in adductor spasmodic dysphonia. Neurology 1988; 38:1220–5.
154. Blitzer A, Brin MF, Fahn S, Lovelace RE. Localized injections of botulinum toxin for the treatment of focal laryngeal dystonia (spastic dysphonia). Laryngoscope 1988;98:193–7.
155. Borodic GE, Ferrante R. Effects of repeated botulinum toxin injections on orbicularis oculi muscle. J Clin Neuroopthalmol 1992;12:121–7.
156. Jaffe DM, Solomon NP, Robinson RA, et al. Comparison of concentric needle versus hooked-wire electrodes in the canine larynx. Otolaryngol Head Neck Surg 1998;118:655–62.
157. Ford CN, Bless DM, Lowery JD. Indirect laryngoscopic approach for injection of botulinum toxin in spasmodic dysphonia. Otolaryngol Head Neck Surg 1990;103:752–8.
158. Green DC, Berke GS, Ward PH, Gerratt BR. Point-touch technique of botulinum toxin injection for the treatment of spasmodic dysphonia. Ann Otol Rhinol Laryngol 1992;101:883–7.
159. Rhew K, Fiedler DA, Ludlow CL. Technique for injection of botulinum toxin through the flexible nasolaryngoscope. Otolaryngol Head Neck Surg 1994;111:787–94.
160. Ludlow CL, Bagley JA, Yin SG, Koda J. A comparison of different injection techniques in the treatment of spasmodic dysphonia with botulinum toxin. J Voice 1992;6:380–6.
161. Blitzer A, Brin M, Stewart C, et al. Abductor laryngeal dystonia: a series treated with botulinum toxin. Laryngoscope 1992;102:163–7.

162. Hertegard S, Granqvist S, Lindestad PA. Botulinum toxin injections for essential voice tremor. Ann Otol Rhinol Laryngol 2000;109:204–9.
163. Warrick P, Dromey C, Irish JC, et al. Botulinum toxin for essential tremor of the voice with multiple anatomical sites of tremor: a crossover design study of unilateral versus bilateral injection. Laryngoscope 2000;110:1366–74.
164. Marion MH, Klap P, Perrin A, Cohen M. Stridor and focal laryngeal dystonia. Lancet 1992;339: 457–8.
165. Christopher KL, Wood R, Eckert RC, et al. Vocal-cord dysfunction presenting as asthma. N Engl J Med 1983;306:1566–70.
166. Hurbis CG, Schild JA. Laryngeal changes during exercise and exercise-induced asthma. Ann Otol Rhinol Laryngol 1991;100:34–7.
167. Cormier Y, Camus P, Desmeules M. Non-organic acute upper airway obstruction. Am Rev Respir Dis 1980;121:147–50.
168. Craig T, Sitz K, Squire E, et al. Vocal cord dysfunction during wartime. Mil Med 1992;157:614–6.
169. Koufman JA. Gastroesophageal reflux and voice disorders. In: Rubin JS, Sataloff R, Korovin G, Gould WJ, editors. Diagnosis and treatment of voice disorders. New York: Ogaku-Shoin Publishers; 1995. p. 161–75.
170. Altman KW, Mirza N, Ruiz C, Sataloff RT. Paradoxical vocal fold motion: presentation and treatment options. J Voice 2000;14:99–103.
171. Maschka DA, Bauman NM, McCray PB, et al. A classification scheme for paradoxical vocal cord motion. Laryngoscope 1997;107:1429–35.
172. Hayes JP, Nolan MT, Brennan N, Fitzgerald MX. Three cases of paradoxical vocal cord adduction followed up over a 10 year period. Chest 1993;104: 678–80.
173. Lew MF, Shindo M, Moskowitz CB, et al. Adductor laryngeal breathing dystonia in a patient with lubag (X-linked dystonia-parkinsonism syndrome). Mov Disord 1994;9:318–20.
174. Grillone GA, Blitzer A, Brin MF, et al. Treatment of adductor laryngeal breathing dystonia with botulinum toxin type A. Laryngoscope 1994;104: 30–2.
175. Collett PW, Brancatisano T, Konno K. Changes in glottic aperture during bronchial asthma. Am Rev Respir Dis 1983;128:719–23.
176. Bahrainwala AH, Simon MR. Wheezing and vocal cord dysfunction mimicking asthma. Curr Opin Pulm Med 2001;7:8–13.
177. Morrison M, Rammage L, Nichol H, et al. The management of voice disorders. San Diego (CA): Singular Publishing Group; 1994.
178 Belisle G, Morrison MD. Anatomic correlation for muscle tension dysphonia. J Otolaryngol 1983;12: 319–21.
179. Roy N, Bless DM, Heisey D, Ford CN. Manual circumlaryngeal therapy for functional dysphonia: an evaluation of short- and long-term treatment outcomes. J Voice 1997;11:321–31.
180. Morrison M, Rammage L, Emami AJ. The irritable larynx syndrome. J Voice 1999;13:447–55.
181. Woodson GE, Murry T. Botulinum toxin in the treatment of recalcitrant mutational dysphonia. J Voice 1994;8:347–51.
182. Aronson AE, Brown JR, Litin EM, Pearson JS. Spastic dysphonia. I. Voice, neurologic, and psychiatric aspects. J Speech Hear Disord 1968;33:203–18.
183. Cannito MP. Emotional considerations in spasmodic dysphonia: psychometric quantification. J Commun Disord 1991;24:313–29.
184. Chevrie-Muller C, Arabia-Guidet C, Pfauwadel MC. Can one recover from spasmodic dysphonia? Br J Disord Commun 1987;22:117–28.
185. Elias A, Raven R, Butxher P, Littlejohns DW. Speech therapy for psychogenic voice disorder: a survey of current practice and training. Br J Disord Commun 1989;24:61–76.
186. Ginsberg BI, Wallach JJ, Srain JJ, Biller HF. Defining the psychiatric role in spastic dysphonia. Gen Hosp Psychiatry 1988;10:132–7.
187. Diagnostic and statistical manual of mental disorders—revised (DSM-III-R). 3rd ed. Washington (DC): American Psychiatric Association; 1987.
188. Kinzl J, Biebl W, Rauchegger H. Functional aphonia. A conversion symptom as defensive mechanism against anxiety. Psychother Psychosom 1988;49:31–6.
189. Koller WC, Huber SJ. Tremor disorders of aging: diagnosis and management. Geriatrics 1989;44:33–6.
190. Morrison MD, Nichol H, Rammage RA. Diagnostic criteria in functional dysphonia. Laryngoscope 1986;96:1–8.
191. Murry T, Cannito MP, Woodson GE. Spasmodic dysphonia: emotional status and botulinum toxin. Arch Otolaryngol Head Neck Surg 1994;120:310–6.

CHAPTER 54

Neoplasms of the Larynx and Laryngopharynx

Robert A. Weisman, MD, Kris S. Moe, MD, Lisa A. Orloff, MD

BENIGN TUMORS

Benign neoplasms are uncommon in the larynx and laryngopharynx, compared with their malignant counterparts. They are presented in the order of their frequency of occurrence. Other mass lesions, such as laryngoceles, nodules, polyps, cysts, and varices, may simulate a benign neoplasm and are discussed in Chapter 66.

Papillomas

Papillomas are verrucous lesions, often with multiple warty excrescences that occur primarily on the true vocal folds (Figure 54–1). They may extend beyond the true vocal folds to the supraglottis, the subglottis, and, rarely, the tracheobroncheal tree.

Neuroendocrine Tumors

Neuroendocrine cells, derived from the neural crest, have been demonstrated in the larynx.[1] Two types of tumors, paragangliomas and carcinoids, are thought to arise from these cells. Laryngeal paraganglia are associated with the superior and recurrent laryngeal nerves. Carcinoid tumors are discussed later in this chapter.

Paragangliomas These tumors arise from the paired paraganglia associated with the internal branch of the superior laryngeal nerve and the posterior branch of the recurrent laryngeal nerve (sensory branches of Galen's anastomosis). Rarely, they may be found anteriorly in the cricothyroid membrane. Laryngeal paragangliomas are histologically identical to carotid body tumors, glomus vagale tumors, and glomus jugulare tumors and may be part of a multifocal pattern seen in some patients. The lesion is most often associated with the superior laryngeal nerve, has a dumbbell shape, and may be partially within and outside the larynx, resembling a laryngocele. Differentiation from a laryngocele is best made by computed tomography (CT), magnetic resonance imaging (MRI), and angiography. As with other "glomus" tumors, treatment can involve observation, embolization, excision, or radiation therapy, depending on the individual patient characteristics.

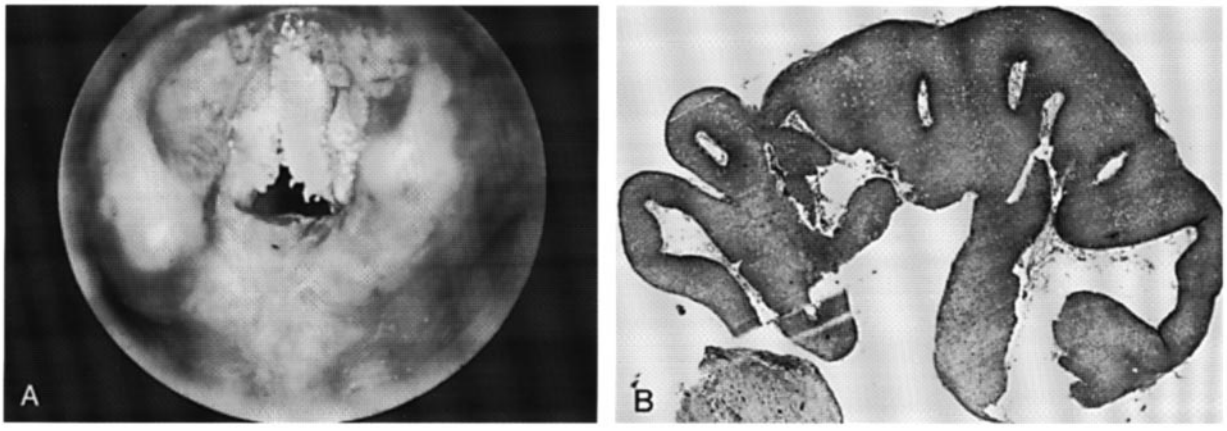


FIGURE 54–1. *A*, Papilloma involving the true vocal folds. The lesion is typically raised and verrucous. *B*, Histologic appearance of pedunculated papilloma (hematoxylin and eosin stain; original magnification × 40).

Schwannomas

This benign nerve sheath tumor arises most frequently in the head and neck region and has been reported in the larynx in over 100 patients. Most of the laryngeal tumors arise around the superior laryngeal nerve and may be found in the aryepiglottic fold. There is a slight female preponderence, and most occur in the fifth and sixth decades of life. Most are small enough to be removed endoscopically, whereas larger tumors may require a lateral pharyngotomy. Malignant degeneration is very rare but has been reported.[2]

Neurofibromas

Neurofibromas contain a mixture of neural axonal or dendritic fibers and Schwann cell elements.[3] They are rare in the larynx and usually appear in the extrinsic portion. They are frequently associated with von Recklinghausen's disease, as part of a generalized neurofibromatosis. They are most frequently encountered in children and young adults and frequently present with stridor. The lesion is submucosal, occurs most often in the supraglottis, and is not encapsulated. Surgical cure is usually associated with a neural deficit, either sensory or motor.

Granular Cell Tumors (Myoblastomas)

Although frequently referred to in the literature as a "myoblastoma," this neoplasm is actually of Schwann cell origin. Approximately 50% arise in the head and neck region, most commonly in the tongue. About 10% arise in the larynx, where they most commonly arise at the junction of the vocal ligament and the vocal process of the arytenoid cartilage as a submucosal mass. These lesions can be removed endoscopically or by thyrotomy and are not well encapsulated. They are sometimes multifocal[4] and are most frequently encountered in the fourth and fifth decades of life. Incomplete excision has been noted to result in apparent long-term cure in some cases.

Vascular Tumors

Hemangiomas Angiomas may be simple (capillary) or cavernous in the larynx and pharynx. The most common symptom is bleeding, which may be severe. On inspection, these lesions are bluish masses and, when exophytic, may resemble a "bag of blue worms." Biopsy of such lesions may be hazardous in an uncontrolled setting and is often nondiagnostic. Treatment consists of endoscopic excision for small lesions and external approaches for larger tumors, where proximal and distal vascular control is optimal.

Cavernous hemangiomas occur most frequently in the false vocal folds, and capillary hemangiomas may occur in both the supraglottis and the vocal folds. Congenital hemangiomas, usually detected in infancy, occur mostly in the subglottis. First described by Phillips and Rush in 1913,[5] subglottic hemangioma is considered a distinct clinical entity. It occurs twice as frequently in males as in females. There is a common association with multiple cutaneous or other capillary hemangiomas of the skin and oral mucosa (Figure 54–2). There is a 5% association with other congenital anomalies, and over 70% occur in the posterior part of the larynx or posterior tracheal wall. Management is discussed in Chapter 45.

Lymphangiomas Lymphangiomas are generally synonymous with cystic hygromas of infancy. In the laryngopharynx, they are most frequently found in

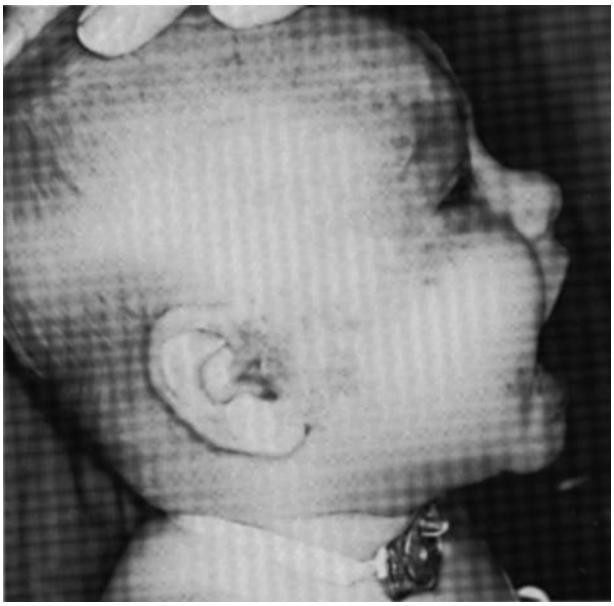


FIGURE 54–2. Male infant with subglottic congenital hemangioma and facial and mucosal capillary hemangiomas.

the supraglottis and the hypopharynx, and patients with them often present with airway obstruction. In the supraglottis, extensive infiltration throughout adjacent structures is common.[6]

Hemangiopericytomas Hemangiopericytomas arise from the pericytes lining the walls of vessels. The majority are benign and arise in the supraglottis. They are most often encountered in elderly patients and may bleed easily when manipulated.

Myogenic Tumors

Leiomyomas Leiomyomas of the larynx are rare. They occur most often in the false vocal folds, aryepiglottic folds, and ventricles. A vascular form called angioleiomyoma has been described in older men. They are most often submucosal and occasionally pedunculated, arising from the vocalis muscle. Lancinating pain is noted in some cases. Incomplete excision has been associated with significant bleeding.[7]

Myomas and Myoblastomas Myomas and myoblastomas are benign tumors that originate over the arytenoid areas. They appear as red swellings with smooth surfaces. Excision can be performed endoscopically but may require an open procedure if there is extension to deeper tissues or if they are broadly based in the larynx or pharynx.

MALIGNANT NEOPLASMS

Squamous Cell Carcinoma

Incidence The incidence of laryngeal cancer is relatively low compared with that of carcinomas of all sites, comprising about 2 to 5% of all cancers worldwide.[8] The relative incidence is similar to that for all sites in the oral cavity as a group and is similar to that of thyroid cancer (12,000 cases in the United States per year).[8] According to the American Cancer Society, 10,000 new cases of laryngeal cancer and 4,000 deaths related to laryngeal cancer were expected to occur in the United States in the year 2000.[8] In addition, 2,500 new cases of hypopharyngeal cancer were expected. Over the years, there has been a male predominance noted (about 5:1 in the United States). As more women have taken up smoking in recent decades, a shift in this ratio can be expected. In the United States, a study at Ben Taub General Hospital revealed that comparing the two 15-year periods 1959 to 1973 and 1974 to 1988, the ratio of male-to-female incidence of laryngeal cancer dropped from 5.6 to 1 to 4.5 to 1.[9] Another study revealed an increase in the incidence of supraglottic lesions in women.[10]

There is an increased reported incidence of laryngeal cancers in industrialized areas (up to 6 to 8% of all cancers). In the Connecticut Tumor Registry, the incidence of carcinoma of the larynx was recently reported to be 12% of all tumors diagnosed. About 3,700 deaths related to laryngeal cancer were reported in 1989, with an overall mortality probability of 32% for those newly diagnosed with the disease. At the time of diagnosis, 25% had regional metastases, and 8 to 10% had distant metastases. These statistics were similar in 2001.

The greatest incidence of laryngeal cancer occurs in the fifth, sixth, and seventh decades (more than 80%). The largest number of cases occurs in the sixth decade (approximately 40% of all cases), with the average age for the occurrence of laryngeal carcinoma appearing to be approximately 59 years. In Scandinavian countries, the peak age of occurrence appears to be older, with a peak between 70 and 74 seen in Norway. Worldwide, the peak incidence appears to be between 55 and 65. In age-corrected data, the incidence continues to rise until age 75. Laryngeal cancer in children and adolescents is rare and tends to be papillary-lymphocytic and nonkeratinizing.

Laryngeal and laryngopharyngeal cancers continue to be the most common malignant tumors of the head and neck in the US, despite relative increases in the incidence of oral cavity and oropharynx tumors in recent decades. In contrast, both incidence and mortality have doubled in Italy over the last few decades.[11] Similar findings have been reported in the Scandinavian countries, Uruguay, and Puerto Rico. Among the countries with the lowest reported incidences of these cancers are Japan, Singapore, Syria, Armenia, and Australia. Within individual countries, inhabitants of urban, highly industrialized areas have a laryngeal cancer risk two to three times that of rural inhabitants.[12] Certain religious groups, such as Mormons, Seventh-Day Adventists, and Parsis, which forbid drinking ethanol and smoking tobacco, have low rates of laryngeal cancer.

Etiologic Factors By far, smoking tobacco has been implicated as the prime factor in causing laryn-

geal cancer. Cigarette, pipe, and cigar smoking are all considered risk factors, with cigarette smoking posing the greatest risk. Only about 1% of larynx cancer occurs in nonsmokers. Nicotine has not been clearly shown to be carcinogenic, but there is abundant evidence that tars and polycyclic hydrocarbons found in tobacco smoke are strong carcinogens in the larynx. Epidemiologically, larynx cancer risk increases with younger age at onset of smoking behavior, number of years smoking, and average quantity of tobacco smoked per day.[13]

Other etiologic factors have been suggested. There are conflicting reports on the role of alcohol in the development of laryngeal cancer, although the preponderance of evidence suggests a promoting effect of alcohol in the pathogenesis of laryngeal cancer, especially in smokers who drink heavily.[14] Alcohol use appears to correlate better with the occurrence of supraglottic cancers rather than glottic tumors.[15] Alcoholics who smoke are thought to carry a 25- to 50-fold increase in risk of laryngeal cancer over nonsmokers. Other etiologic factors include gastroesophageal reflux[16] and exposure to wood dust, asbestos, volatile chemicals, nitrogen mustard, and ionizing radiation. Persons with various types of immunodeficiency states are also thought to be at higher risk for developing laryngeal cancer. An underlying genetic susceptibility to this type of cancer may also exist and has been theorized as a potential cofactor in oncogenesis. A growing body of evidence from molecular genetic studies supports this hypothesis.[17–23] Such a theory might explain why some individuals smoke heavily and never develop cancer, whereas others may smoke lightly or not at all and develop laryngeal cancer. The role of "passive smoking" in the development of laryngeal cancer remains unknown. Human papillomavirus (HPV) is also considered to be an important cofactor in the development of laryngeal cancers. The strain HPV 16 has been most frequently demonstrated in laryngeal squamous carcinomas by using probes for HPV deoxyribonucleic acid (DNA).[24] In addition, the capsid antigen for HPV has been identified in as many as 70% of all laryngeal cancers.[25] A multifactorial model for the development of laryngeal cancer has been proposed by Koufman and Burke and is shown in Figure 54–3.

Variations in the occurrence of laryngeal cancers are seen among countries and ethnic groups but generally reflect the prevalence of tobacco use in

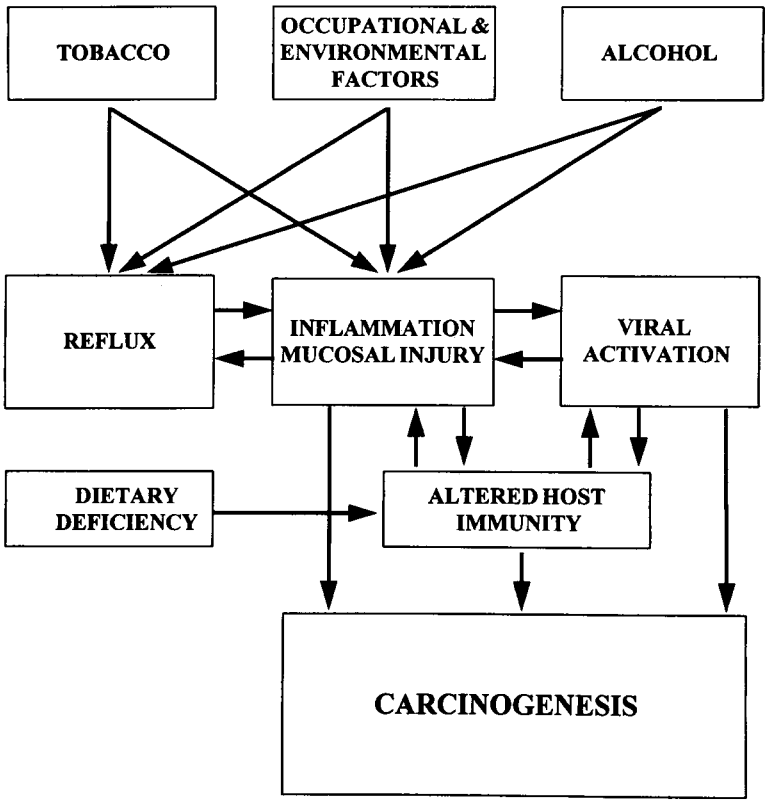


FIGURE 54–3. Schematic model of multifactorial development of squamous cell carcinoma of the larynx. Reproduced with permission from Koufman JA, Burke AJ. The etiology and pathogenesis of laryngeal carcinoma. Otolaryngol Clin North Am 1997;30:1–19.

those countries. There are also wide variations geographically in the distribution of laryngeal cancers among the various subsites. In some countries (eg, Finland), supraglottic cancers predominate, whereas in neighboring countries (Sweden), glottic tumors are in the majority. In the United States, African Americans have a 50% higher rate of laryngeal cancer compared with whites.[26] Clearly, genetic influences are important in determining the occurrence and site of involvement. However, this is a complex issue, and other factors, such as general nutrition, are important. Deficiencies in certain nutrients, particularly the B vitamins, vitamin A, beta-carotene, and retinoids, are considered to be important in the development of squamous cell cancer in general. Heavy alcohol use and lower socioeconomic class may predispose to these nutritional deficiencies or to a higher incidence of gastroesophageal reflux disease and could account for some of the differences noted between different countries, cultures, and ethnic groups.

Verrucous carcinoma of the larynx has been studied extensively with regard to its etiology. Tobacco use is not as highly correlated with this variant of squamous cell carcinoma, but there is a very strong association with HPV.[27] In situ hybridization studies have demonstrated the presence of DNA from the HPV 16 strain, almost exclusively.[24] Human papillomavirus DNA has not been demonstrated in normal mucosa adjacent to these cancers.[28] As in other aerodigestive tract squamous cell cancers, the virus is thought to act as a promoter in the process of carcinogenesis.[29]

Oral herpes simplex infections are also thought to be possible predisposing causes of laryngeal cancers. Elevated levels of circulating carcinoembryonic antigen have been demonstrated in this group of patients. Furthermore, affected patients are postulated to have depressed cellular immunity, which could lead to failure of immune surveillance and progression of a developing cancer.

An increased risk of laryngeal cancer has been noted with certain anatomic conditions. Laryngoceles, or cystic dilations of the laryngeal ventricles, are found in 2% of adult larynges but can be identified in as many as 18% of laryngeal cancers. The incidence of laryngoceles increases with age, especially in males. Whether the laryngocele precedes the cancer and is a causative factor or whether it is secondary to obstruction of the ventricle by tumor is debatable. Another anatomic condition, sulcus vocalis, an indentation along the free edge of the vocal fold, has been reported to occur with greater than expected frequency in patients with laryngeal cancer.[30]

Plummer-Vinson syndrome (sideropenic dysplasia with atrophic gastritis, achlorhydria, and glossitis) is reported to be associated with an increased incidence of postcricoid carcinoma. Most of these patients are older women living in northern latitudes. The exact mechanism by which the carcinomas are induced is not known, but iron deficiency, vitamin deficiencies, and genetic and nutritional factors may all play a role. Up to 30% of affected individuals will develop the carcinoma.[31]

Ionizing radiation is thought to induce some laryngeal cancers, just as it is thought to play a role in thyroid and salivary gland cancers. Workers exposed to radioactive sources were noted as early as the 1930s to have an increased rate of laryngeal cancer. Reports have also appeared linking therapeutic doses of radioactive iodine to subsequent laryngeal neoplasia. Patients who received external beam radiation in relatively low doses for such conditions as acne, scrofula, and skin cancer have been noted to have a slightly increased risk of laryngeal cancer. In juvenile laryngeal papillomatosis, reports of malignant degeneration after therapeutic external beam radiation suggest that the radiation may have induced the malignant change. However, many of the patients who were treated in this fashion had advanced, aggressive, recurrent, and disseminated papillomas that may have already begun to undergo malignant change. It has also been suggested that when second primary tumors appear many years (usually 15 to 20) after previous external beam radiation treatment for cancer of the head and neck, the second tumor was induced by the radiotherapy.[32] This association is difficult to prove because most of these patients have risk factors for head and neck cancer, and there is roughly a 4% per year risk of developing a second head and neck primary in this population.

Association with Second Primary Malignant Tumors Second primary tumors occur synchronously in about 1% of laryngeal cancers and metachronously in 5 to 10%. This occurrence seems to be greater with supraglottic compared with glottic primaries. Some series have indicated a three-fold

increase in supraglottic cancers.[33] Laryngeal cancer is one of the tumors with one of the highest rates of associated second primary cancers. The most common second primary tumor is bronchogenic carcinoma. Of these second primaries in the lung, 10% are synchronous, 30% occur in the first year, another 30% by 5 years, and the rest within the first 20 years. The risk of a second lung tumor is directly proportional to the duration and degree of smoking and continuation of smoking after definitive therapy for the laryngeal cancer.[34] Eighty percent of the patients developing second lung primary tumors have carcinoma in situ (Cis) or T1 laryngeal tumors.[35]

Although annual plain chest radiographs are used to screen for second primary tumors (and metastatic disease) and are useful in detecting asymptomatic lung primaries, the effect on outcome may be small. A study by Engelen and associates revealed a malignant lung primary in 12.4% of patients being followed for laryngeal cancer.[36] They noted 69% of the patients to be asymptomatic when the lung primary was detected. The survival time for these patients was longer than for those whose lung primary was detected after symptoms appeared, but overall survival was not significantly impacted. The role of chest CT in screening for second primaries and metastatic disease is not clear since the test is very sensitive and not very specific. Small nodules of indeterminate significance are often found on chest CT, with the minority of the nodules being malignant. Positron emission tomography (PET) may prove to be more effective as a screening procedure for secondary primary tumors in laryngeal cancer, and in head and neck cancer in general, because of early reports of its sensitivity and specificity for malignant deposits. Concurrent PET and CT or MRI may provide anatomic detail regarding areas of abnormal functional activity, allowing targeted biopsy of small areas to determine whether a new lesion is a second primary. Investigations are currently in progress to determine if this is a cost-effective modality to screen for second primary tumors and for metastatic disease.

Pathology Over 95% of all malignant laryngeal tumors are squamous cell carcinomas. These tumors arise from the surface epithelium, which is a stratified squamous lining, except at the free edge of the vocal fold, where the mucosa consists of a pseudostratified squamous layer. Areas of hyperkeratosis, dysplasia, and Cis are often found adjacent to the primary invasive lesion. With proper orientation, it is usually possible to define a transition from normal epithelium, to dysplastic epithelium, to invasive carcinoma (Figure 54–4). Islands, tongues, and clusters or cords of invasive atypical cells within the underlying stroma characterize the tumor. The cells show squamous differentiation, including keratinization of individual cells, keratin pearl formation, or intercellular bridging.

When hyperkeratosis is present, a characteristic white lesion is seen on the vocal fold. Generally,

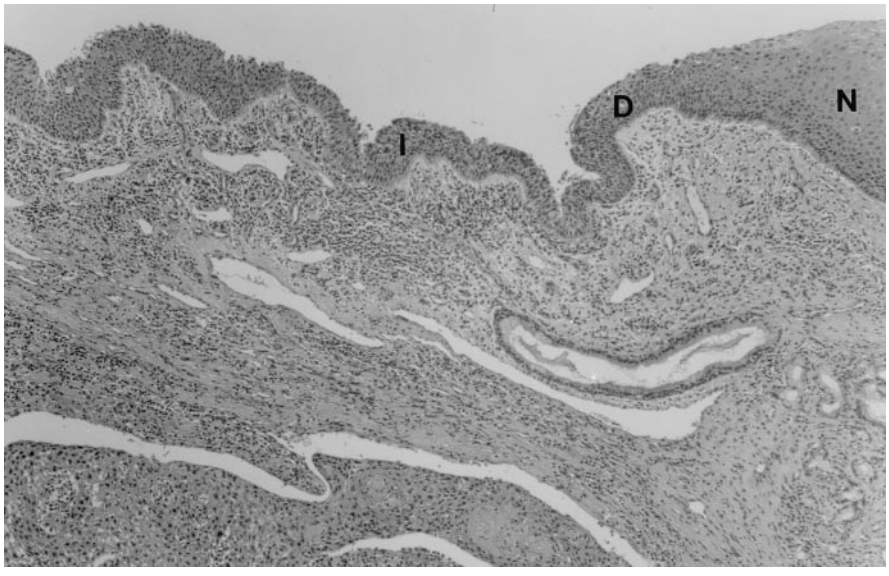


FIGURE 54–4. Photomicrograph depicting transition from normal epithelium (N) to dysplasia (D) to in situ carcinoma (I) (hematoxylin and eosin stain).

such lesions exhibit thickening of the epithelial layer and abnormal keratinization of the superficial layers. The generic term leukoplakia is often used to describe such a finding, although both benign and malignant lesions can have this appearance (Figure 54–5). This lesion may be solitary or multifocal. Histologically, most of these lesions are benign, but there is thought to be an approximately 3% risk of malignancy for leukoplakia of the vocal fold.[37]

Erythroplasia or erythroplakia refers to a red lesion of the vocal fold. Most of these lesions are raised and irregular, with increased vascularity and abnormally engorged U-shaped vessels on the surface. These lesions are thought to carry even greater risk of harboring malignancy than white lesions.[38] Both lesions require biopsy to establish their nature. Recently, contact endoscopy with vital staining has been introduced,[39] with reports indicating that benign and malignant lesions can be distinguished by this method without biopsy.[40]

It is important for clinicians to understand the meaning and significance of certain pathologic terms that are used to describe changes in the vocal fold epithelium that precede the development of true cancer. *Hyperplasia* is merely a thickening of the epithelial layer owing to an increased number of cells. This finding is often seen in areas of irritation or trauma. Metaplasia refers to a change in the histology of a tissue from one benign form to another. *Hyperkeratosis* is an increase in the keratin layer of squamous epithelium. *Keratosis* and *keratotic* are terms that more accurately describe this change in the vocal fold since the normal epithelial layer there is nonkeratinized. *Dysplasia* is a truly premalignant change and describes disorder in the usual progressive maturation of cells from the basal layer to the outer layer of the squamous epithelium, usually in association with changes in individual cells on a spectrum from benign to malignant. Dysplasia (some pathologists prefer the term *atypia*) is usually graded mild, moderate, or severe, depending on the degree of these changes. Most pathologists agree that severe dysplasia is synonymous with Cis. In Cis, the carcinoma is limited to the epithelial layer, and the basement membrane remains intact (Figure 54–6). The tumor can involve a seromucinous gland and still be considered in situ. Only when the basement membrane is crossed by the tumor is the term *invasive* carcinoma used. Since Cis and invasive carcinoma are often found on the vocal fold surrounded by areas of dysplasia of varying degrees, one school of thought suggests that carcinoma occurs after an orderly progression from mild to moderate to severe dysplasia (Cis).[41] Others have disputed this theory since some carcinomas appear to be surrounded by only mild dysplasia, and some patients have focal Cis in the basal layers without dysplasia of the surrounding epithelial cells. It is possible that more than one mechanism exists.

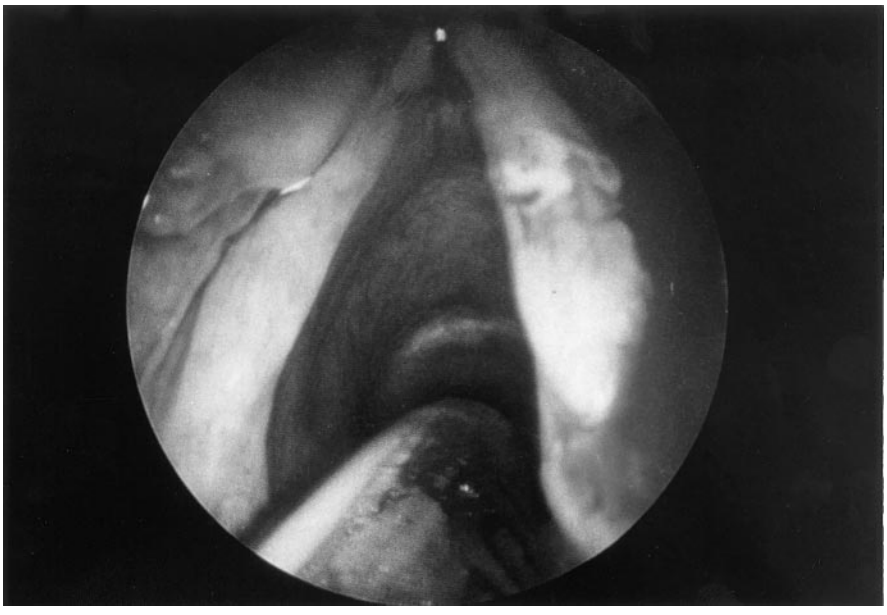


FIGURE 54–5. Leukoplakia of the true vocal fold. The malignant potential of these lesions warrants biopsy.

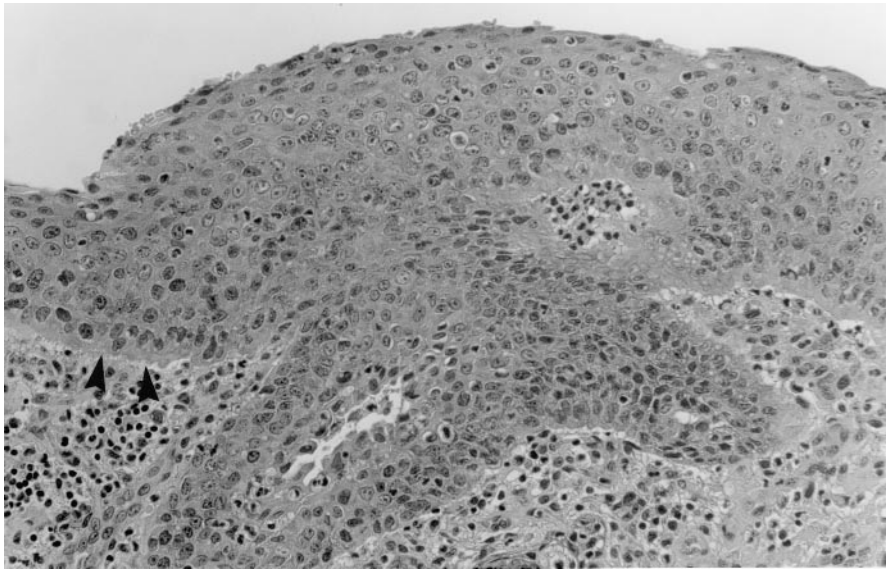


FIGURE 54–6. Photomicrograph of carcinoma in situ of the true vocal fold. Although the individual epithelial cells are clearly malignant, the basement membrane is intact (*arrows*) (hematoxylin and eosin stain).

These lesions have clinical relevance since 3% of hyperkeratoses without dysplasia, 7% of mild dysplasias, 18% of moderate dysplasias, and 24% of severe dysplasias of the vocal folds have been shown to ultimately develop invasive carcinoma.[37] Since these lesions begin on the vocal fold and cause hoarseness early, they are often detected early, in contrast to many other primary sites. It is important to understand the impact of these pathologic descriptions to counsel the patient regarding cancer risk and to institute appropriate plans for close follow-up.

Other pathologic features of laryngeal cancers are relevant to prognosis. Supraglottic cancers tend to be less differentiated than glottic cancers and exhibit earlier and more extensive local invasion as a group. Although grading of cellular differentiation does not correlate strongly as an independent variable with regard to outcome in most series, more poorly differentiated tumors have a greater likelihood of lymphatic metastasis,[42] which has a major impact on survival.[43] Other microscopic features, such as presence of a pushing versus an infiltrating border, the presence or absence of a local host inflammatory reaction to the tumor, and presence or absence of vascular or perineural invasion, have been noted as important prognostic indicators in laryngeal cancers and in head and neck cancers in general. The thickness of the lesion or depth of invasion may be an important prognostic indicator, as has been shown in other head and neck cancer primary sites (eg, floor of mouth), but this correlation has not been demonstrated in the larynx.

Besides the gross and microscopic appearance of vocal fold lesions, genetic and molecular biologic characteristics of these abnormal areas are being studied to determine their malignant potential. Cell cycle and flow cytometry techniques have been used to study the DNA content of laryngeal neoplasms. Dysplastic lesions that are aneuploid have a high risk of progressing to invasive carcinoma.

Variants of Squamous Cell Carcinoma. Verrucous Carcinoma. Verrucous carcinoma is a subtype of squamous cell carcinoma that comprises between 1 and 4% of all malignant tumors of the larynx.[44] Grossly, it is characterized by an exophytic, keratotic mass with a distinct edge. Lymph node metastases do not occur.[45] Typically, multiple biopsies and observation of the clinical behavior of the lesion are necessary before a definite diagnosis of malignancy is made. Histologically, these tumors are extremely well differentiated, invade the underlying stroma with a pushing border, and have very few mitoses (Figure 54–7). The differential diagnosis includes verruciform squamous cell carcinoma, verruca vulgaris, and verrucous hyperplasia.

Basaloid Squamous Cell Carcinoma. This variant of squamous cell cancer is seen more frequently in

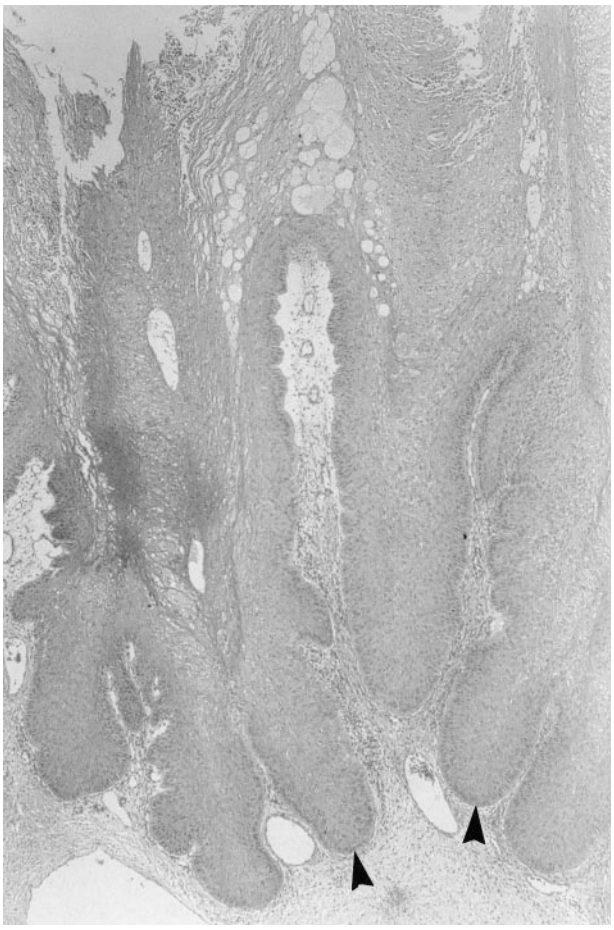


FIGURE 54–7. Photomicrograph showing a well-differentiated tumor with a "pushing" margin (*arrows*), typical of verrucous carcinoma (hematoxylin and eosin stain).

larger lesions that exhibit aggressive behavior.[46] It is more frequently encountered in the hypopharynx than the supraglottic or glottic larynx. Histologically, nests of small, basaloid cells that stain predominantly blue because of their high nuclear-to-cytoplasm ratio characterize the tumor. Comedonecrosis and stromal hyalinization are frequently encountered. Pseudopalisading of the tumor cells is occasionally seen (Figure 54–8). The differential diagnosis includes adenoid cystic carcinoma (solid variant). Whereas some studies have indicated a poorer prognosis for this subtype compared with other squamous cell cancer, others have shown no difference in survival when subjects are matched for stage, site, and type of treatment.[47] Interestingly, DNA analysis showing aneuploidy in basaloid variants predicts an improved survival.[48]

Spindle Cell Variant. Spindle cell squamous carcinoma is also called pseudosarcoma, carcinosarcoma, and sarcomatoid carcinoma of the larynx. Grossly, the lesion is usually polypoid, but about one-third of cases exhibit ulceration and endophytic growth.[49] Histologically, these tumors show a pleomorphic spindle cell component associated with invasive cancer or Cis. Occasionally, the associated squamous cancer cannot be identified. Rarely, but particularly after radiation therapy, the spindle cell portion of the tumor undergoes osteocartilaginous differentiation.[49] Using immunohistochemical stains and electron microscopy, the spindle cells can often be shown to be of epithelial origin, leading to speculation that the spindle cell portion represents an area of malignant metaplasia.[50] The differential diagnosis includes soft tissue sarcoma, malignant melanoma, and unusual reaction to radiation therapy in exuberant granulation tissue. Despite its aggressive appearance and worse prognosis than spindle cell tumors at most sites, glottic spindle cell tumors have an excellent (90%) survival rate. These tumors are radioresistant, and surgery is favored as the mainstay of treatment.

Nonsquamous Tumors. Adenoid Cystic Carcinoma. This tumor is rare in the larynx and probably arises from minor salivary glands or seromucinous glands. Reflecting this origin, most of these tumors are either subglottic or supraglottic.[51] As at other sites, the tumor can exhibit a cribriform, tubular, or solid growth pattern, with the latter being associated with a somewhat worse prognosis. As at other sites, the neoplasm has a strong proclivity for perineural invasion and spread and often presents with pain. Pulmonary and osseous metastases are fairly common, and lymph node metastasis is infrequent enough that elective treatment of the neck is not indicated.

Neuroendocrine Carcinoma. There are three distinct subtypes in this group of unusual neoplasms. Since they behave differently and have different treatments, correct histologic diagnosis is essential.

Typical carcinoid of the larynx is extremely rare and is seen almost exclusively in males. Histologically, the tumors contain islands, nests, and cords of cells that lack nuclear atypia, mitotic figures, or necrosis. Surgery is usually curative.[52]

Atypical carcinoid is more common than typical carcinoid in the larynx. Differing from typical

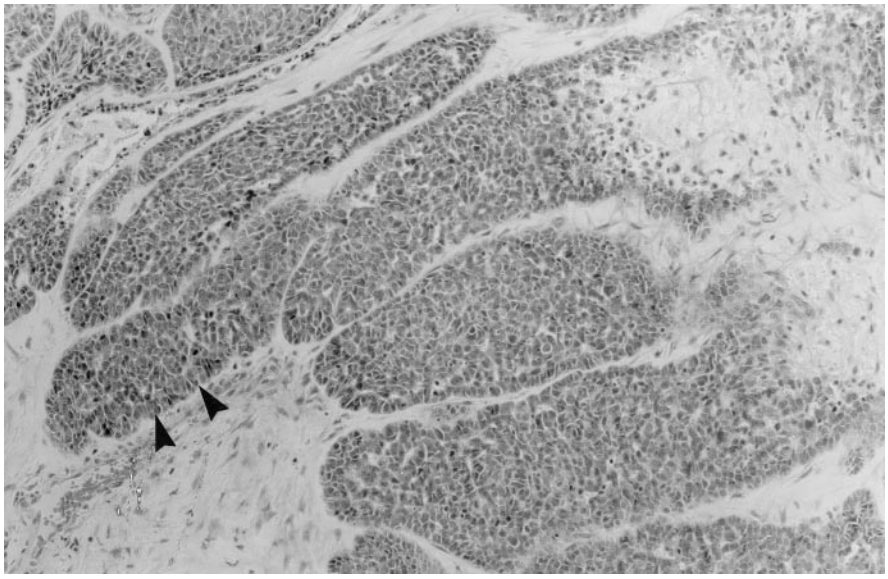


FIGURE 54–8. Photomicrograph of basaloid squamous cell carcinoma. The basaloid cells are small and dark, with pseudopalisading (*arrows*); this is a squamous cell cancer (hematoxylin and eosin stain).

carcinoid, the tumor cells exhibit mild to moderate nuclear atypia, occasional mitotic figures, and individual cell necrosis. Microscopically, these tumors strongly resemble medullary thyroid carcinoma and often stain positive for calcitonin, even though serum calcitonin levels are rarely elevated. Atypical carcinoids are more biologically aggressive than typical carcinoids, with regional and distant metastases being common. Surgery with or without adjuvant radiation and chemotherapy is the treatment of choice, yielding a survival rate of about 60%.[53]

Small cell carcinoma of the larynx is quite rare, comprising 0.5% of all laryngeal malignancies.[52] Histologically, it is identical to small cell carcinoma of the lung, and a metastasis from a lung primary should be considered when this lesion is encountered. Like small cell lung cancer, the prognosis is poor. Surgery is ineffective, and treatment consists of chemotherapy and radiation therapy.

CHONDROSARCOMA. Chondrosarcoma is rare but is the most common sarcoma arising in the larynx.[54] Clinically, the tumor is slow growing and submucosal and most commonly arises from the cricoid cartilage. Biopsy can be problematic because of the location of the lesion and its firm consistency and because a large tissue sample is needed to establish the grade of the neoplasm. Histologically, the tumor exhibits hypercellularity, nuclear atypia, and double nucleated cells, all of which distinguish the tumor from a benign chondroma. Surgery is the main treatment, and lymph node metastasis is rare.

Anatomic Subsites and Patterns of Spread of Laryngeal Tumors The spread of cancers arising in the larynx follows fairly predictable patterns. The pathway of spread is determined by several factors, among them the presence of anatomic barriers to spread (eg, perichondrium and cartilage) and the density of local vessels and lymphatic channels. Because the supraglottis arises from a different embryologic source than the glottis and subglottis, its lymphatic drainage patterns are significantly separate and different. This separation is often referred to as compartmentalization of the larynx, as described by Pressman.[55] A schematic diagram of the anatomic subsites used to classify laryngeal tumors is shown in Figure 54–9.

SUPRAGLOTTIC TUMORS. Supraglottic cancers involve the region bounded superiorly by the free border of the epiglottis and inferiorly by the false vocal folds and the laryngeal ventricles (Figure 54–10). Laterally, the medial aspect of the aryepiglottic folds defines the border. These neoplasms tend to spread by local extension. The majority start on the epiglottis and from there can extend to the aryepiglottic folds, false cords, and ventricle. The epiglottic perichondrium and cartilage are initial barriers to anterior spread. The thyroepiglottic ligament is the last

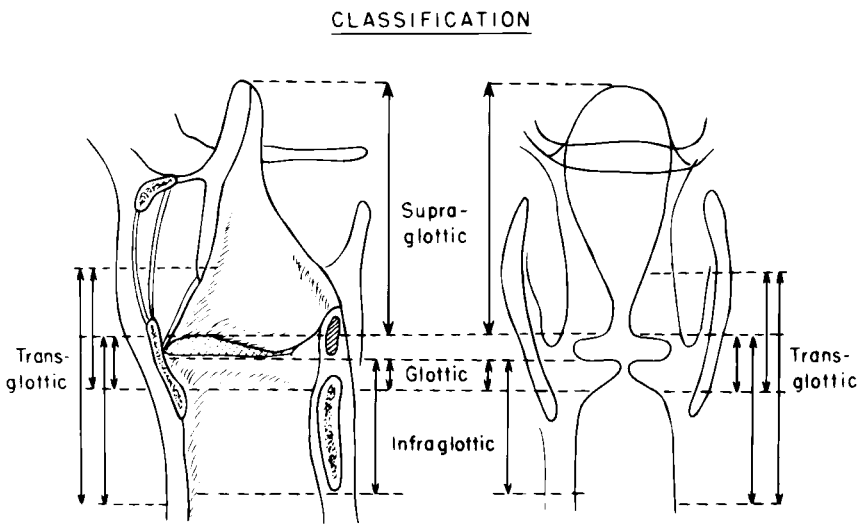


FIGURE 54–9. Schematic representation of the anatomic classification of laryngeal tumors.

barrier to spread anteriorly into the preepiglottic space. Anterior to the soft fibrofatty tissue of the preepiglottic space, the inner perichondrium of the thyroid cartilage and the cartilage itself and the outer perichondrium are the final barriers to spread into the neck. Extension through the vallecula allows tumor spread directly into the base of the tongue. In the aryepiglottic fold, the quadrangular membrane is a moderate barrier to spread of tumors superiorly and laterally from the epiglottis or the false vocal fold.

There is a strong tendency for supraglottic tumors to spread via lymphatics. Numerous reports estimate that 39 to 65% of patients with T2 to T4 supraglottic cancers present with overt lymph node metastasis, whereas 32 to 34% of such patients staged N0 have pathologically positive nodes.[56,57] This observation is consistent with the rich lym-

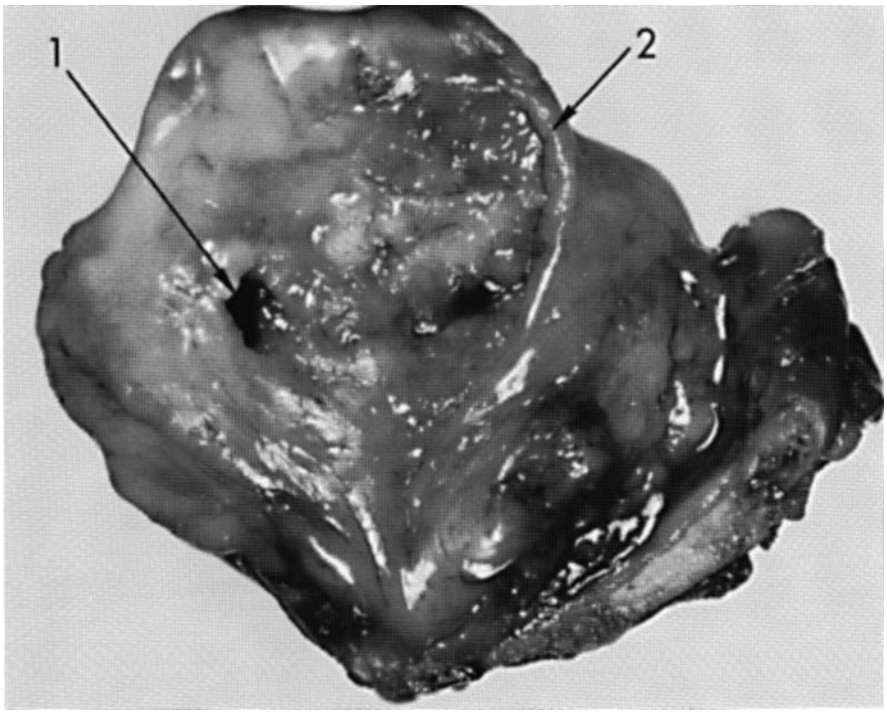


FIGURE 54–10. Specimen from a subtotal supraglottic laryngectomy demonstrating gross appearance of the tumor. Note infiltrating and spreading margins (2) and central area of necrosis with epiglottic foraminal invasion of the cartilage (1).

phatic network found in the supraglottic submucosa. From the epiglottis, lymph channels drain bilaterally to the false vocal folds. The drainage from the false vocal folds and the remainder of the supraglottic larynx is superior and lateral. Most of these lymph channels exit the larynx via the thyrohyoid membrane in parallel with the superior laryngeal veins. The primary nodal beds draining the supraglottis are in levels II and III in the neck (Figure 54–11).

Cancers arising on the edge or margin of the aryepiglottic fold are subclassified as "marginal tumors" and are considered by some to be separate from supraglottic lesions. Their biologic behavior resembles that of piriform sinus cancers. They therefore have a very high rate of submucosal spread, early lymphatic metastasis, and greater propensity for distant metastasis than supraglottic lesions.

Glottic Tumors. Glottic tumors involve the true vocal folds. The inferior limit of this region is generally thought to be 10 mm below the free edge of the true vocal fold, although there is some debate over this distance. The intrinsic muscles of the larynx extend about 10 mm below the true vocal folds. Some contend that the glottis extends up to 20 mm below the vocal fold, whereas others choose 5 mm as the limit. The anterior limit of the glottis is the anterior commissure, and the posterior limit is the vocal processes of the arytenoid cartilages.

The true vocal folds are unusual in that they have very sparse lymphatic drainage. Glottic neoplasms, therefore, tend not to spread via lymphatics until the tumor has extended beyond the vocal fold into an area richer in lymph vessels. The main means of spread of glottic tumors is by local extension (Figure 54–12). Primary vocal fold neoplasms tend to stay confined to the vocal fold for a relatively long period of time and are detected at that stage frequently because patients develop hoarseness early in the course of their disease and seek medical attention. When glottic tumors spread, they tend to

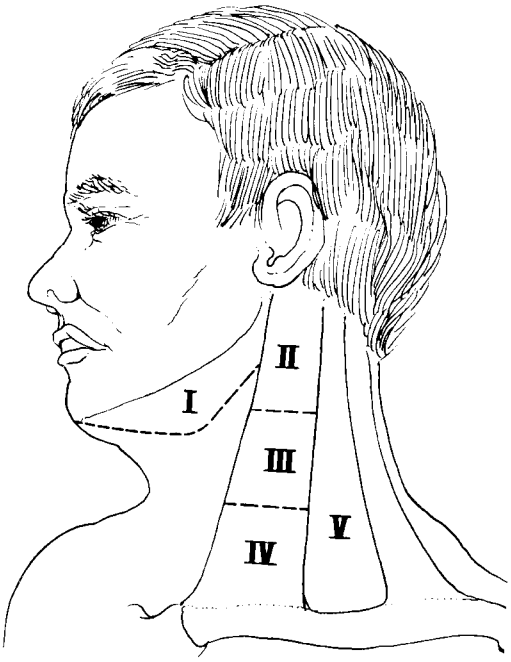


FIGURE 54–11. Levels of lymph node involvement that have been designated for prognostic significance. Reproduced with permission from Alvi A, Myers EN, Johnson JT. Cancer of the oral cavity. In: Myers EN, Suen JY, editors. Cancer of the head and neck. 3rd ed. Philadelphia: WB Saunders; 1996.

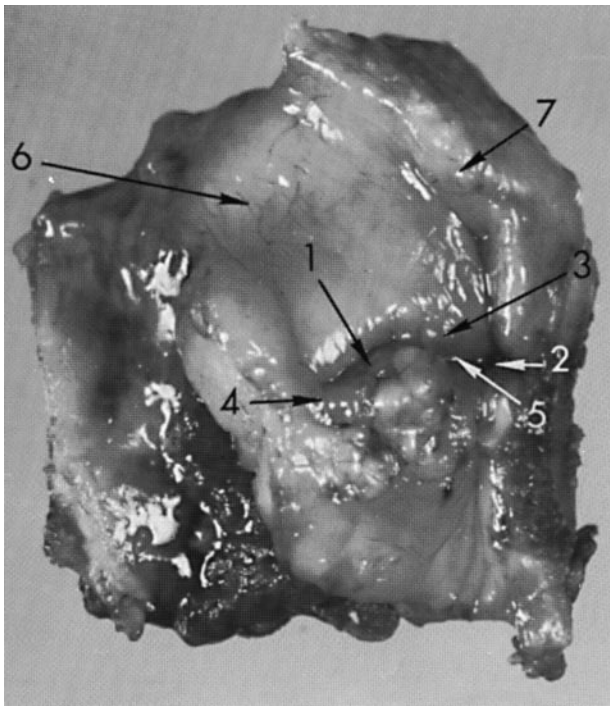


FIGURE 54–12. Left hemilaryngectomy specimen with glottic carcinoma (1) invading the ventricle and subglottis. Anterior commissure (2), false cord (3), vocal process of arytenoid cartilage (4), laryngeal ventricle (5), aryepiglottic fold (6), and epiglottic petiole (7). Note vascularity and heaped-up appearance of the tumor.

extend laterally and inferiorly, respecting the embryologic compartmentalization of the larynx. When vocal fold tumors extend laterally, they enter the paraglottic space and have access to spread either superiorly or inferiorly along the thyroid cartilage. Direct invasion of the thyroarytenoid muscle and spread to the cricoarytenoid area occur, resulting in vocal cord fixation or paralysis.

Superior spread in the paraglottic space results in tumor invasion of the preepiglottic space or even the thyrohyoid membrane. Inferior spread results in extension to and through the cricothyroid membrane into the thyroid gland or other neck structures. Such extension can also result in spread along the conus elasticus into the subglottic space. Extension of tumors onto the undersurface of the true vocal fold can also result in extension to the subglottis, either on the surface or submucosally. Another important pathway of spread from the glottis occurs at the anterior commissure. The anterior vocal ligament and associated vessels perforate the thyroid cartilage at that point, creating a preformed pathway for spread of cancers that involve the anterior commissure. This mode of spread has been used to explain treatment failures of relatively small lesions at the anterior commissure treated with either radiation or endoscopic resection. Direct invasion of the thyroid cartilage at other locations is usually seen only in tumors that are very advanced locally.

Subglottic Cancers. The subglottic region begins 10 mm below the vocal folds and extends inferiorly to the level of the inferior border of the cricoid cartilage. Primary tumors in the subglottis are rare, comprising only 1% of all laryngeal cancers. They tend to spread downward, either on the surface or submucosally in the upper trachea or anteriorly through the cricothyroid membrane. Exit through the cricothyroid membrane results in spread into adjacent neck structures, as noted above for glottic lesions. The risk of lymphatic spread in these tumors is probably intermediate between supraglottic and glottic tumors and ranges between 10 and 34% in the small series that have been reported.

Transglottic Tumors. These cancers cross the ventricle from either the true cord upward or the supraglottis downward. Structures of both the supraglottic and glottic larynx are involved. This term can also be applied to cancers arising on the true vocal folds that extend greater than 10 mm below the edge of the vocal fold to involve the subglottis or to tumors that involve all three major subdivisions of the larynx. Patterns of spread follow those for all of the affected subsites.

Superior Hypopharyngeal Tumors. Cancer that involves the lingual surface of the epiglottis, vallecula, or base of the tongue falls into this category. The lateral boundaries are the glossoepiglottic folds.

Inferior Hypopharyngeal Tumors. These are mainly piriform sinus tumors. They are found in an area bounded superiorly by the glossoepiglottic fold, medially by the aryepiglottic fold and medial wall of the piriform sinus, laterally by the mucosa over the thyroid lamina, and posteriorly by the lateral wall of the hypopharynx. Postcricoid tumors arise from the mucosa over the posterior cricoid lamina and are bounded superiorly by the arytenoids and inferiorly by the entrance to the esophagus. Posterior pharyngeal wall tumors arise from the posterior wall of the hypopharynx. All of the inferior hypopharyngeal tumors tend to spread by local infiltration on the surface and submucosally, spread via lymphatics early in their course, and have a high rate of distant metastasis. Posteriorly, the prevertebral fascia is the main important barrier to spread.

Distant Metastasis. Distant spread of primary laryngeal tumors is usually a late event, when tumors have recurred locally and/or regionally after treatment. At presentation, 25% will have regional nodal metastasis, and only 8 to 10% have distant metastasis. Overall, 50% of all laryngeal cancers remain locally confined, whereas 35% have regional spread, and 15% develop distant metastases. The lung, liver, and bone are the most frequently involved sites for metastatic disease.

Clinical Evaluation Assessment of the patient with laryngeal cancer begins with a thorough history and physical examination. Patients with supraglottic cancers often remain asymptomatic until a relatively large tumor bulk is present. Nodal metastasis is often the initial complaint. Patients with glottic tumors tend to present early, with hoarseness as their chief complaint. Subglottic tumors are rare and may present with stridor or hemoptysis. Several key symptoms should be noted in the history.

Hoarseness is the major presenting symptom in patients with glottic cancer. The examiner must determine the duration of this symptom, its progression, and associated symptoms. The degree of voice alteration is related to the extent of the lesion. On the true vocal fold, even a small tumor on the surface epithelium may alter the voice by interfering with the normal mucosal wave and by adding mass to the vocal fold. More extensive lesions that invade below the surface epithelium tether the surface, interfering with normal propagation of the mucosal wave of the vocal fold. This subtle disturbance is best appreciated on videostroboscopy. More extensive invasion involves the vocalis muscle, leading to diminished lateral movement of the vocal fold, even to the point of complete fixation.

Hoarseness is commonly present for 3 months or longer in patients who present with vocal fold cancer. As many as 30% have been hoarse for a year or more. In supraglottic tumors, hoarseness tends to occur later, when tumors have become bulky and overhang the glottis or spread through the paraglottic space to fix the vocal fold. The voice change seen with supraglottic lesions tends to be more of a muffled voice, somewhat similar to the "hot potato" voice of epiglottitis. Some patients with supraglottic tumors have no hoarseness at all and may be asymptomatic or have other complaints. Hypopharyngeal tumors rarely cause hoarseness until they become quite extensive, causing either vocal fold fixation or effects on the voice caused by sheer bulk.

Other symptoms that should be sought in the interview of a patient suspected of having laryngeal cancer include dyspnea, dysphagia, and pain, particularly ear pain. Dyspnea occurs in glottic tumors when the airway is compromised by tumor bulk or by unilateral or bilateral vocal fold fixation. Edema associated with tumor invasion of local lymphatics may also contribute to airway narrowing. In supraglottic tumors, dyspnea is a late finding, when tumor bulk is so great that the oropharyngeal or hypopharyngeal portion of the airway becomes obstructed. Dysphagia is not frequently encountered with glottic tumors until tumor size has increased enough to cause arytenoid dysfunction. Supraglottic tumors are more likely to cause dysphagia or odynophagia because the tumor is frequently on the epiglottis, where base of tongue invasion or movement of the tumor with swallowing may elicit pain and cause swallowing dysfunction. Dysphagia is most often seen in tumors that involve the hypopharynx, such as piriform sinus, posterior pharyngeal wall, or the postcricoid area.

Pain is a late finding in glottic cancers and usually denotes extension of the tumor through cartilage and into extralaryngeal structures. Supraglottic tumors may present with pain earlier, in the form of a sore or scratchy throat. As supraglottic lesions enlarge, pain increases with involvement of the base of tongue, cartilage invasion, and extension into the superior laryngeal nerves and the soft tissues of the neck. Ear pain is considered "referred pain" and is considered to be a sign of glossopharyngeal or vagus nerve involvement, usually by extension of tumors into the base of the tongue or the piriform sinus. Primary tumors of the piriform sinus commonly present with referred ear pain, often before any other symptoms are present.

Other symptoms less frequently reported by patients with laryngeal cancer include cough, hemoptysis, bad breath, and weight loss. Cough is generally a sign of low-grade aspiration owing to loss of laryngeal sphincteric function and is most commonly seen in glottic tumors. Hemoptysis is seen mainly in patients with large, fungating, friable supraglottic neoplasms. Occasionally, blood loss can be insidious, causing severe anemia. Foul odor is seen in larger tumors, mainly in the supraglottis, that exhibit significant necrosis. Weight loss is usually an omen of a poor prognosis, indicating an advanced lesion, possibly already exhibiting distant metastasis. Tenderness to touch over the larynx or loss of the external landmarks of the larynx on palpation is also worrisome and usually indicates a tumor that has extended through the laryngeal cartilages.

Other historical items are of importance. The general health status of the patient should be sought in detail, including comorbidities that may have an impact on choice of treatment. The patient's use of tobacco should be documented. Patients often stop smoking shortly before presentation to a physician and may answer "no" if asked if they smoke. Asking if the patient has ever smoked will usually reveal the smoking history. Alcohol consumption should be documented in a similar manner. Learning the patient's occupation through the social history can also be important and may have an impact on decision making later. As with other cancer patients, it is important for the caregiver to be aware of the

patient's social circumstances and to develop an idea of what kind of support the patient will have in the post-treatment period out of the hospital setting. Proper discharge planning will ensure a smooth transition from hospital to home and maximize the patient's psychological well-being.

Physical Examination. All patients with hoarseness of 2 weeks or longer without a clear-cut diagnosis should undergo evaluation by a head and neck specialist. The larynx should be assessed by all means available to the physician until he/she is satisfied that an adequate examination has been achieved. Indirect mirror laryngoscopy in a cooperative patient can be quite adequate and gives an undistorted view of the larynx. Flexible or rigid fiber-optic laryngoscopy should be used in most cases in which cancer is suspected. Video documentation is useful for patient teaching, for review by the examiner, for presentation to other consultants not present at the endoscopy, and for comparison to the treated larynx after radiation or organ-conserving surgery.

Videostroboscopy adds another dimension to the evaluation of early glottic cancers. This technology focuses on the movement of the epithelial layer of the vocal fold relative to the underlying stroma, vocal ligament, and musculature. During phonation, the vocal fold epithelium moves freely over the fibroelastic cushion of the lamina propria in Reinke's space. A surface cancer on the vocal fold that invades the basement membrane, by definition, tethers the mucosa to the underlying stroma, causing dyskinesis of the mucosal wave on that vocal fold and loss of synchrony between the vocal folds. Deeper invasion, with tethering of the mucosa to the intrinsic muscles of the larynx, produces even greater distortion of the mucosal wave. These subtle changes are not apparent on naked eye or routine fiber-optic examination. However, such findings may be of critical importance when deciding whether to radiate a patient or use microsurgical excision for early vocal fold cancer.[58,59] Others have challenged this concept, pointing out that thick keratoses and scarring from previous vocal fold trauma can produce alterations similar to invasive tumors.[60] These authors suggest that stroboscopy may be more important in evaluating the vocal fold not involved with cancer since its status may be the best predictor of voice outcome after treatment.

Neck examination is always performed when a diagnosis of laryngeal cancer is suspected. Lymph nodes are rarely detected on physical examination with most glottic cancers, even when the primary is relatively advanced. Supraglottic cancers and tumors of the laryngopharynx are much more likely to present with nodal metastasis. When lymph nodes are detected, their size, character, number, and position in the neck should be recorded. In supraglottic cancers, levels II and III are most likely to harbor metastases, whereas in glottic cancers, levels II, III, and IV are the likely sites (see Figure 54–11). For cancers in the upper laryngopharynx, lymphatic drainage is similar to that for supraglottic cancers, and for those in the inferior region, levels III and IV are most often involved.

Imaging Studies. Several modalities are currently available to the clinician for evaluating the larynx and neck. Earlier methods such as plain radiographs, laryngograms, xeroradiography, and laminography have been completely replaced by CT and MRI in the United States. The newer modalities show fine anatomic detail and provide useful information on cartilage invasion (Figure 54–13), subglottic extension, extralaryngeal spread, base of tongue invasion, preepiglottic space invasion, and the status of lymph nodes, usually with greater accuracy than physical examination. Both CT and MRI should be considered complementary to physical examination, office endoscopy, and operative endoscopy. Neither modality delineates small or superficial primary lesions well, for example.

Neck staging is enhanced with both modalities, which have similar sensitivities and specificities for detecting nodal metastases. Both types of imaging still miss a significant percentage of occult metastatic disease in necks staged N0 clinically. Both modalities have sensitivities in the 60 to 80% range, whereas specificity ranges from 72 to 89%. Features such as size greater than 1 cm (1.5 cm for level II lymph nodes), evidence of central necrosis, irregular lymph node margins, loss of fat planes around lymph nodes, round rather than oval shape, and clusters of lymph nodes that do not individually meet size criteria are considered positive for metastatic disease radiographically.[61]

Positron emission tomography scanning is a relatively new development that images cancers differently than CT and MR. Positron emission tomography is a functional study based on the differential uptake of radioactive [18^{F}]fluorodeoxyglucose by

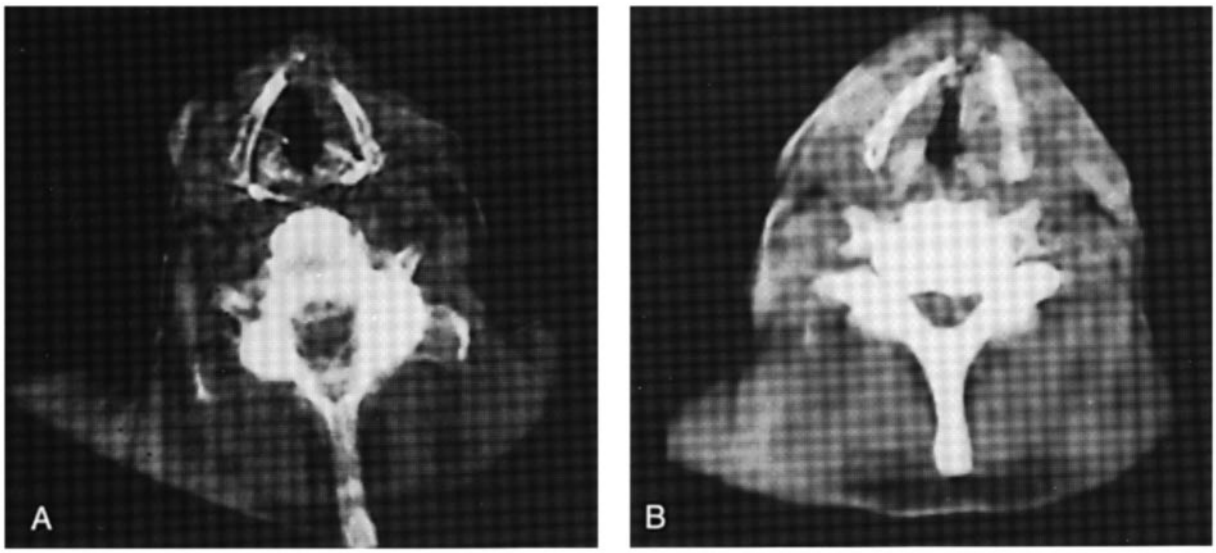


FIGURE 54–13. Computed tomographic scans of two vocal fold tumors. *A*, Anterior commissure and thyroid cartilage invasion. *B*, Left vocal fold carcinoma with vocalis muscle, anterior commissure, and thyroid cartilage invasion.

different tissues. Tumors tend to have a greater metabolic demand than normal tissues and take up the radionuclide more rapidly. This property of malignant tissues shows up as a "hot spot" on a PET scan (Figure 54–14). Positron emission tomography has been shown to be highly specific, sensitive, and accurate (up to 100%) in detecting occult nodal disease in the N0 neck and promises to be more helpful in staging the neck than the current standard of CT or MRI.[62–64] Assessment of primary tumors of the larynx with PET alone is not likely to be helpful, but newer technology combining PET and CT or MRI gives both functional and anatomic information about a tumor's size and location and may prove to be very valuable in the precise delineation of the extent of laryngeal cancers.[65] Positron emission tomography does appear to have promise in detecting recurrent tumors and in finding second primary lesions in the head and neck.[66,67] Caution may be necessary in interpreting the results of PET scans of the larynx since false positives are thought to occur in patients who talk extensively just prior to examination.

Panendoscopy. The next step in diagnostic evaluation of patients with laryngeal cancer is panendoscopy. This diagnostic procedure is sometimes referred to as triple endoscopy and includes direct laryngoscopy, esophagoscopy, and bronchoscopy. The purpose of the operation is to assess the extent of the laryngeal tumor by direct laryngoscopy, obtain representative tumor tissue by biopsy, and assess the respiratory tract and upper digestive tract for synchronous primary tumors. The procedure is performed under deep general anesthesia with paralysis to allow a thorough, safe examination of the larynx, esophagus, and tracheobronchial tree. The relative merits of esophagoscopy and bronchoscopy compared to imaging studies for the detection of multiple primary tumors are a subject of debate. However, panendoscopy remains the standard by which other modalities are measured. Esophagoscopy and bronchoscopy should be performed prior to laryngoscopy so that bleeding from the more proximal laryngeal lesion will not distort the examiner's view of the esophagus or tracheobronchial tree.

Direct laryngoscopy, although appearing to be a relatively simple procedure, requires technical skill and experience to be performed effectively, with minimal trauma to the patient. A variety of laryngoscopes are available for this procedure, although the technology has not changed greatly in the last 100 years. The older laryngoscopes, such as those introduced by Jackson and Hollinger, provide adequate lighting and access and permit monocular viewing. Biopsy is performed through the open tube of the laryngoscope. The larger laryngoscopes introduced

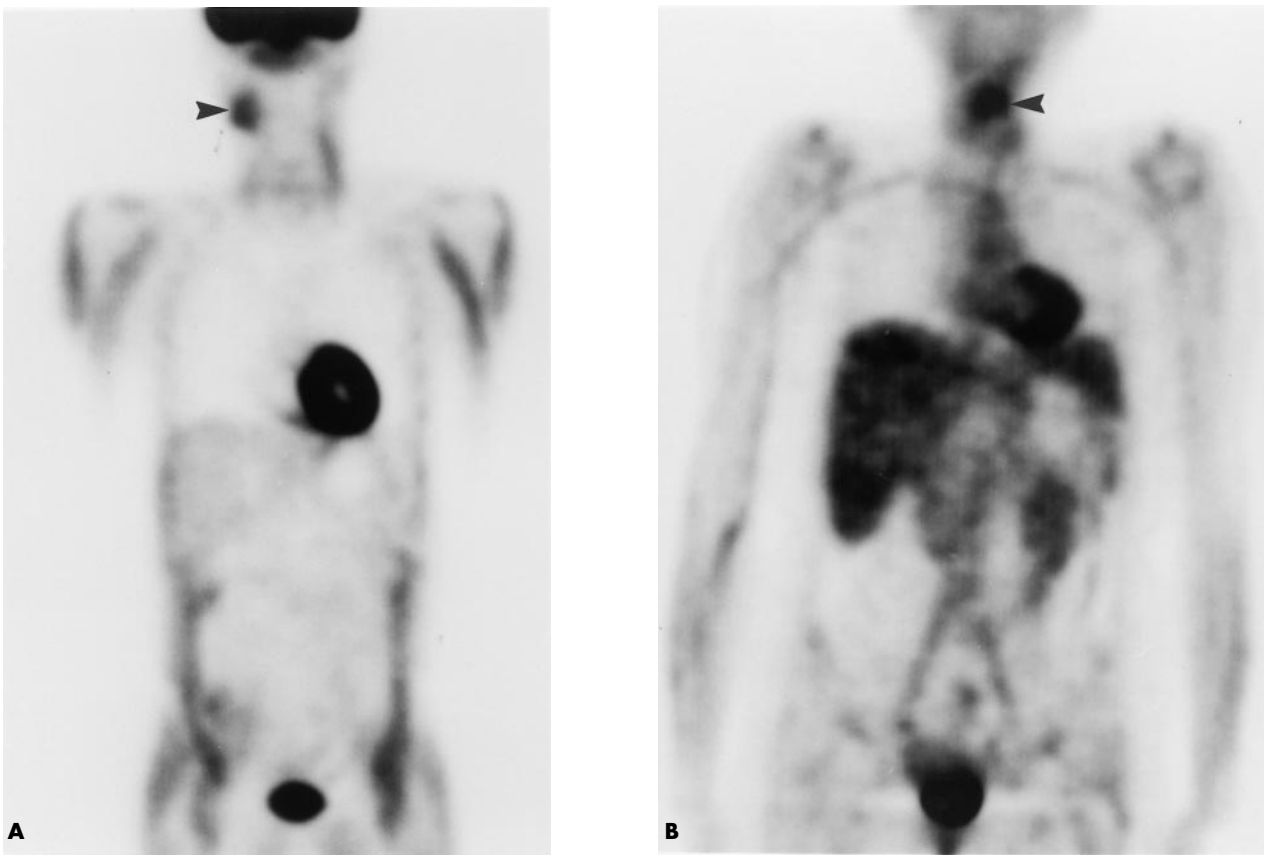


FIGURE 54–14. *A*, Positron emission tomographic scan showing uptake in the neck in a mass (*arrowhead*) thought to be radiation fibrosis. Fine-needle aspiration was negative on multiple passes. Open biopsy showed viable cancer cells. *B*, Positron emission tomographic scan showing uptake in the larynx (*arrowhead*) in a patient treated with radiation therapy for piriform sinus cancer. Office endoscopy and computed tomography were negative, but biopsy was positive.

by Jako, Dedo, and others permit binocular viewing and can be coupled with the operating microscope to provide a magnified view of early glottic lesions. Bivalved endoscopes were introduced in Europe in the last decade, provide wide exposure of supraglottic lesions in particular, and are useful in providing the wide access needed for endoscopic resection of small tumors. Zeitels recently introduced a modified laryngoscope with a triangular rather than oval tip, which is thought to improve exposure of the anterior commissure.[68] All of these laryngoscopes can be suspended to free both of the surgeon's hands for manipulation of the microscope and use of surgical instruments.

It is important to perform direct laryngoscopy gently. Great care should be taken to avoid injury to the anterior maxillary teeth, preferably by using a standard or custom-made tooth guard. The laryngoscope should be inserted gently, using the examiner's nondominant hand, so that instruments can be introduced with the dominant hand. When suspension is used, both hands can control instruments. The laryngoscope should be advanced slowly, following landmarks such as the uvula, epiglottis, and endotracheal tube to find the glottis. When examining supraglottic tumors, a broad, flat-tipped laryngoscope, modeled after the standard Jackson laryngoscope, should be used. Since supraglottic tumors bleed easily, the laryngoscope should be placed in the vallecula first to allow viewing of the epiglottis. Even with the best equipment, supraglottic tumors can be difficult to assess. They arise from and extend to structures that are relatively mobile and compressible. Glottic lesions remain more stable during examination and can be more accurately evaluated.

Microlaryngoscopy, in which the operating microscope is coupled with direct laryngoscopy, is highly useful in examining early cancers of the true vocal fold. The magnification provided by the microscope allows accurate determination of the extent of most small tumors. Biopsy using microinstrumentation can be very precise, limiting damage to adjacent normal portions of the vocal fold. Subglottic extension can be assessed with the microscope, although this is sometimes better evaluated by inserting a 0-degree rigid telescope between the vocal folds. Spatulas and vocal cord retractors used in laser surgery can be used to displace a vocal fold to examine its undersurface or to allow better visualization of the subglottic space or the laryngeal ventricle.

Adequate specimens are obtained for biopsy using cup forceps. Sufficient viable-appearing tissue should be obtained for pathologic examination. Excessive biopsies are not warranted and should be avoided. In the case of bulky tumors that intrude on the airway, an argument can be made for debulking at the time of biopsy to decrease the need for tracheostomy at the end of the endoscopy. Such patients need to be observed closely in the hospital since delayed bleeding and swelling can lead to airway obstruction, usually several hours after the procedure.

In patients with a marginal airway, a decision has to be made whether endotracheal intubation will be attempted or whether a tracheostomy under local anesthesia will be performed prior to direct laryngoscopy. Consultation between the surgeon and the anesthesiologist is critical to avoiding airway emergencies in such patients. The decision will be influenced partly by the skill and experience of the anesthesiologist in intubating the partially obstructed airway. Such patients should be intubated awake, with fiber-optic guidance. It is safer to perform a tracheostomy under general anesthesia with the airway already secured; if there is any doubt that the anesthesiologist will be able to pass an endotracheal tube into the airway without precipitating obstruction, a tracheostomy should be performed under local anesthesia. In some cases, patients with partial airway obstruction from a laryngeal tumor will not tolerate the supine position, and the tracheostomy has to be performed with the patient in the sitting position.

In certain patients, the risk of general anesthesia or other factors may preclude operative direct laryngoscopy. It is possible in some patients to obtain biopsy specimens without general anesthesia, usually to facilitate nonsurgical management or in the palliative setting. Fine-needle aspiration of cervical metastases is often sufficient. In the absence of clinically positive lymph nodes, some tumors can be sampled by performing a biopsy through a flexible bronchoscope introduced under topical anesthesia. Supraglottic tumors are fairly easy to access in this fashion. Glottic lesions on the side of a fixed vocal fold can also be accessed for biopsy in this manner. This technique should be avoided in patients with marginal airways. In certain cases, CT-guided needle biopsy can be employed to sample a laryngeal primary. Hypopharyngeal lesions can usually be sampled safely by either method.

Tumor Staging. The staging of laryngeal cancer in the United States follows the guidelines set forth in the tumor, node, metastasis (TNM) classification system created by the American Joint Committee on Cancer. The current classification system is presented in Table 54–1. It is important to remember that this is a clinical staging system, based on all types of clinical information available (physical

TABLE 54–1. TNM Staging for Carcinoma of the Larynx

Stage 0	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1	N1	M0
	T2	N1	M0
	T3	N1	M0
Stage IVa	T4	N0	M0
	T4	N0	M0
	Any T	N2	M0
Stage IVb	Any T	N3	M0
Stage IVc	Any T	Any N	M1

examination, endoscopy, biopsy, imaging) prior to institution of therapy. Like all staging systems, the purpose is to stratify patients according to extent of disease in such a way that groups are defined that can be predicted to have a range of outcomes from excellent to poor. The TNM system does this based on morphologic and functional findings and does a fairly good job of separating patients into four groups (stage I to IV) such that increased stage is associated with decreased survival. Nevertheless, some patients in stage I die, and some in stage IV survive, and prediction of outcome based on tumor extent alone lacks precision.

A modified staging system for cancer of the larynx was proposed by Piccirillo et al,[69] in which patient characteristics were combined with the description of tumor extent. Severity of symptoms related to cancer was combined with comorbidity factors to create a functional severity index. Three functional severity classes were defined based on this index, with survival rates of 83%, 58%, and 15%. These functional severity indices were combined with the TNM stage to derive a clinical severity index. Survival rates for the four stages of the clinical severity index were 88%, 80%, 63%, and 28%. This system proved to be significantly better than the TNM system alone at predicting outcome. The author have subsequently proposed a similar staging system for all head and neck cancer sites that includes comorbidity as a major component.[70]

Although these clinical staging systems are useful and are frequently employed to make decisions when recommending treatment to patients and for comparing outcomes in clinical trials, precise prediction of outcome remains elusive. In the future, analysis of molecular markers and gene expression is likely to provide a better prediction of prognosis than purely clinical staging. Furthermore, even though staging systems may predict outcome in terms of survival, quality of life issues are of great importance to head and neck cancer patients and especially to patients with laryngeal cancer. Measurement of treatment outcomes has already started to focus on quality of life after treatment as well as quantity of life (survival), and quality of life measures are included in most clinical trials of head and neck cancer today. Future staging may also include the likelihood of achieving a good quality of life after treatment as well as survival.

TREATMENT OF CANCERS OF THE LARYNX AND LARYNGOPHARYNX

General Principles

Optimal treatment of cancers of the larynx and laryngopharynx should result in the highest probability possible for cure of a given stage of disease, while either preserving normal functions as much as possible or allowing for post-treatment rehabilitation of the patient to as high a level of function as possible. For patients at every stage of disease, it is no longer sufficient to simply design treatment with survival as the only end point. To this end, there has been a trend over the last 20 years to improve all treatment modalities for laryngeal cancer, such that the impact of treatment on the patient's social and economic circumstances is minimized without compromising survival. The following sections of this chapter review the various treatment options for patients with laryngeal cancer, ranging from small, early lesions to large, advanced lesions. Rehabilitation of speech after treatment will also be discussed.

Management of Early Glottic Cancer

The vocal fold is the most common primary site of laryngeal cancer in the United States. Since minor alterations in its function cause symptoms very early in the course of the disease, early detection occurs often and results in a high rate of cure. Effective management of the disease depends on a thorough understanding of the anatomy, histology, and lymphatic drainage of the glottis.

As stated, the larynx is lined by respiratory epithelium except at the vocal cords. Here the tissue is pseudostratified squamous epithelium situated over the lamina propria, a gelatinous fibroelastic layer, which, in turn, covers the vocalis muscle. The lamina propria allows for a sliding mucosal wave over the muscle. Any invasive process through this layer will cause an abnormal wave and hoarseness. The area between the vocalis muscle and mucosa is Reinke's space. Deep to vocalis muscle is the paraglottic space, and after invasion into this space, the disease process is no longer "early," and vertical growth may proceed over the entire lateral aspect of the larynx.

The anatomic compartmentalization of the larynx adds to the favorable prognosis of early glot-

tic carcinoma. The conus elasticus may inhibit progression of a tumor. If the lesion remains superficial and caudal to the free edge, it will grow onto the undersurface of the vocal fold and then to the subglottis. If it progresses cephalically over the conus elasticus into the vocalis muscle and laryngeal ventricle, it is free to spread into the paraglottic space.

"Early" carcinoma refers to a heterogeneous group of tumors, including Cis and T1 and T2 lesions. These tumors are grouped together and designated as early because the largest drop in patient survival occurs with tumor progression from stage II to stage III.

Overall, early glottic carcinoma is understaged clinically in 40% of cases. True Cis is rather rare; most of these tumors have an invasive component and are actually T1 lesions. It is important to consider that although an anterior commissure lesion may appear quite innocuous, it is located in an area that may be difficult to visualize, where the barrier to spread through the perichondrium is violated by Broyle's tendon. Thus, a tumor that grossly appears to be T1 might have microscopic invasion of the cartilage, making it in reality a T4 lesion.

The management of early glottic cancer focuses on the primary site since lymph node metastasis is uncommon with T1 or T2 lesions. There is some debate regarding the best management for early glottic cancers. There are numerous reports regarding both radiotherapy (XRT) and surgery as highly successful modalities for treating early glottic cancers. Primary XRT is delivered as photons in an average dose of 6,600 cGy to a field limited to the larynx, usually 6×6 cm. The treatment is delivered in 6 to 7 weeks in single fractions in the range of 180 to 200 cGy, 5 days a week. With this approach, cure rates for T1 cancers range from 80 to 91%.[75] Many studies include surgical salvage in their results, whereas others do not. Similar cure rates are reported for surgical treatment of early-stage glottic cancers. The choice of which modality to use is frequently influenced by the expected voice outcome. Until recently, it was generally accepted that XRT produced a better voice result than surgery for all early-stage glottic cancers, and in the United States and most of the world, XRT was the preferred initial treatment for early vocal fold cancers.

Radiation Therapy Virtually all of the data on radiation treatment of early glottic cancer are based on retrospective analyses or on single-institution reports of consecutive patients treated. One extensive review of 38 reported series for T1 vocal fold cancer found an overall cure rate of 83.7%.[71] Others have reported a range from 80 to 91%. Local control rates as high as 95% and as low as 67% have been reported.[72] Male gender and bilateral vocal fold involvement portend a worse outcome.

T2 glottic carcinomas are reported to have a lower cure rate than T1 cancers. A study evaluating the results of treatment of T2 glottic cancer with primary XRT reported in 25 series indicated an overall cure rate of 64%.[73] There is wide variation in the reported success rate for treating T2 lesions, which may reflect selection of more favorable lesions at some centers compared with others. T2 lesions are often described as heterogeneous, and some propose dividing them into T2a (normal mobility) and T2b (impaired mobility).[74] Wang reported local control rates of 86% for T2a lesions and 63% for T2b lesions treated with primary XRT.[75] To improve the accuracy of staging of lesions that may involve the anterior commissure, Zeitels recommended resection of the anterior vestibular folds and infrapetiole region to allow inspection of Broyle's tendon. This can be done without disturbing the voice and will demonstrate if there is cartilage invasion in this potentially dangerous area.

Surgical Treatment of Early Glottic Cancer Surgeons have devised many operations for early glottic cancer over the last two centuries. Brauers introduced partial laryngectomy in the form of cordectomy via vertical laryngofissure (thyrotomy) in 1834. Early results with this approach were poor, probably because of a lack of adequate lighting, instrumentation, general anesthesia, antibiotics, and blood replacement. It was long held that partial resection of the larynx was dangerous and that only a total laryngectomy could be considered as surgical treatment for cancer. In the 1940s and 1950s, procedures were introduced that provided adequate tumor removal while preserving some degree of vocal communication, swallowing function, and breathing without a tracheostomy. In the modern era, hemilaryngectomy techniques have replaced cordectomy. Patient selection is critical to success in these procedures. The procedures are based on the knowledge that early glottic tumors remain confined to the vocal fold, tend to be well differentiated with a well-

demarcated leading edge, and rarely spread to lymphatics.

The following criteria should be met for patients to be candidates for partial laryngectomy of any type:

1. The patient should be in good general health. Elderly, debilitated, or mentally impaired patients who would not be able to tolerate a temporary tracheostomy and/or temporary aspiration with eating should not be candidates for conservation procedures.
2. The patient should have a good understanding of the procedure and the possibility that prolonged rehabilitation of speech and swallowing may be required.
3. Adequate pulmonary function is most critical for supraglottic laryngectomy. Poor exercise tolerance, either by history or demonstrated on pulmonary function tests, excludes most patients from open partial laryngeal surgery, particularly supraglottic laryngectomy. If doubt exists, a formal consultation with a pulmonary specialist should be obtained.

For glottic lesions, vertical, frontolateral, and anterior approaches are commonly used. The vertical hemilaryngectomy is most commonly used and is illustrated in Figure 54–15. Absolute control of the vocal fold margins by frozen section must be obtained to achieve maximum success rates. A margin of 1 mm on the vocal fold is considered adequate. Numerous variations on this basic surgical approach have been described. Up to one entire arytenoid, one-third of the contralateral vocal fold, and up to one half of the vertical height of the anterior part of the cricoid cartilage can be removed. If more than one-third of the contralateral vocal fold must be removed, a keel should be placed through the thyrotomy and secured to the skin with button fixation for 3 to 6 weeks (Figure 54–16). Others have described rotation of an epiglottic flap to reconstruct the loss of the anterior parts of the vocal folds[76] (Figure 54–17). Another option for extensive involvement of both true vocal folds is the CHEP procedure (cricohyoidoepiglottopexy), described later in this chapter and shown in Figure 54–18.

Endoscopic Excision of Early Vocal Fold Cancers

With the introduction of the operating microscope and its coupling to the carbon dioxide laser by Strong and Jako in 1972,[77] interest in endoscopic treatment of vocal fold cancer increased. In Western Europe over the last 10 years, a number of investigators have published relatively large series of patients treated by either laser or cold steel excision of early vocal fold cancers. Like partial laryngectomy techniques in the past, acceptance was initially slow in the United States and elsewhere. Doubts were raised concerning the oncologic safety of such techniques and about the voice outcomes compared to XRT. However, detailed reports from centers in Europe and the United States have documented local control and survival outcomes as good as or better than those achieved for XRT for Cis, T1, and T2 vocal fold cancers,[78–83] with local control rates ranging from 88 to 93% and laryngeal preservation rates ranging from 93 to 100% for T1 lesions. Voice results are quite good for early lesions with minimal invasion of

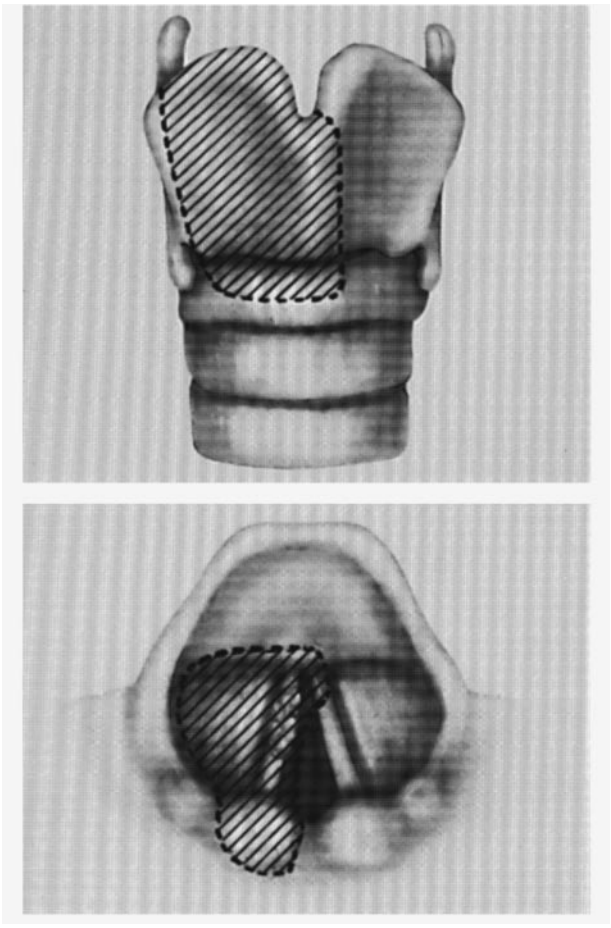


FIGURE 54–15. Schematic diagram of vertical hemilaryngectomy. The *outlined shaded areas* indicate the resected tissue.

Reinke's space and approach those observed after XRT. Furthermore, 40% of Tis to T2 patients whose cancer recurs can be salvaged with laryngeal conservation surgery.[83]

Patients with no or minimal disturbance of the mucosal wave on videostroboscopy are the best candidates for this procedure. Some debate exists over whether carbon dioxide laser or cold steel techniques should be used. Both methods allow precise microscopic control of the excision, but some argue that the heat from the laser–tissue interface results in more scarring after healing than in cold steel techniques, resulting in a worse voice outcome. Distortion of the margins by the laser is also noted as a potential disadvantage. Excision of the false vocal fold on the side of tumor involvement greatly aids visualization during surgeries for larger early lesions and also provides a wider field of view for following the patient endoscopically after treatment.[84]

There are several advantages proposed by advocates of endoscopic treatment of vocal fold cancers. First, cure rates are at least as good as seen with XRT. Second, voice results in selected patients are probably similar to those achieved with XRT. Third, since XRT is not used, it will still be available should the patient develop a second primary tumor or a

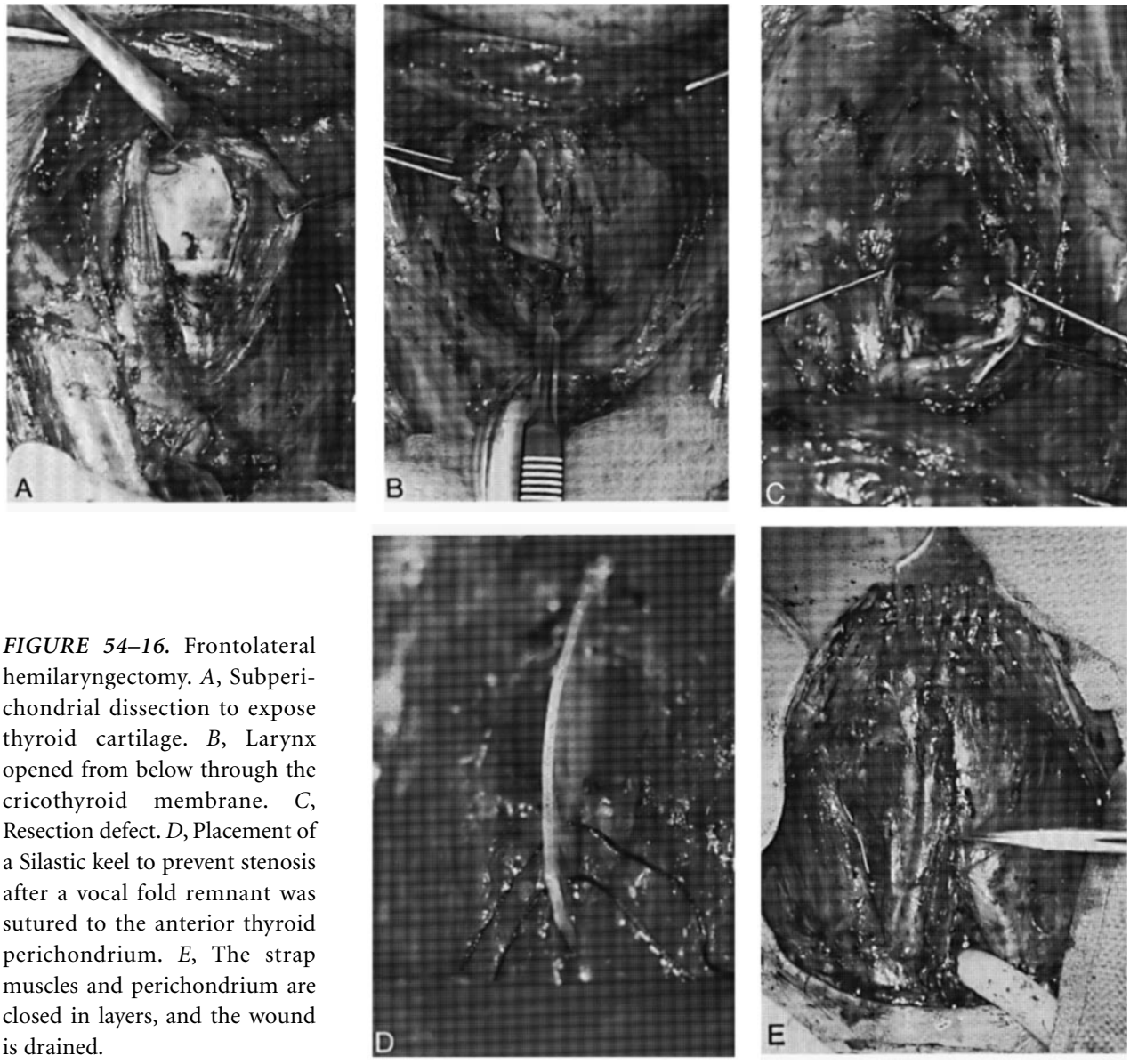


FIGURE 54–16. Frontolateral hemilaryngectomy. *A*, Subperichondrial dissection to expose thyroid cartilage. *B*, Larynx opened from below through the cricothyroid membrane. *C*, Resection defect. *D*, Placement of a Silastic keel to prevent stenosis after a vocal fold remnant was sutured to the anterior thyroid perichondrium. *E*, The strap muscles and perichondrium are closed in layers, and the wound is drained.

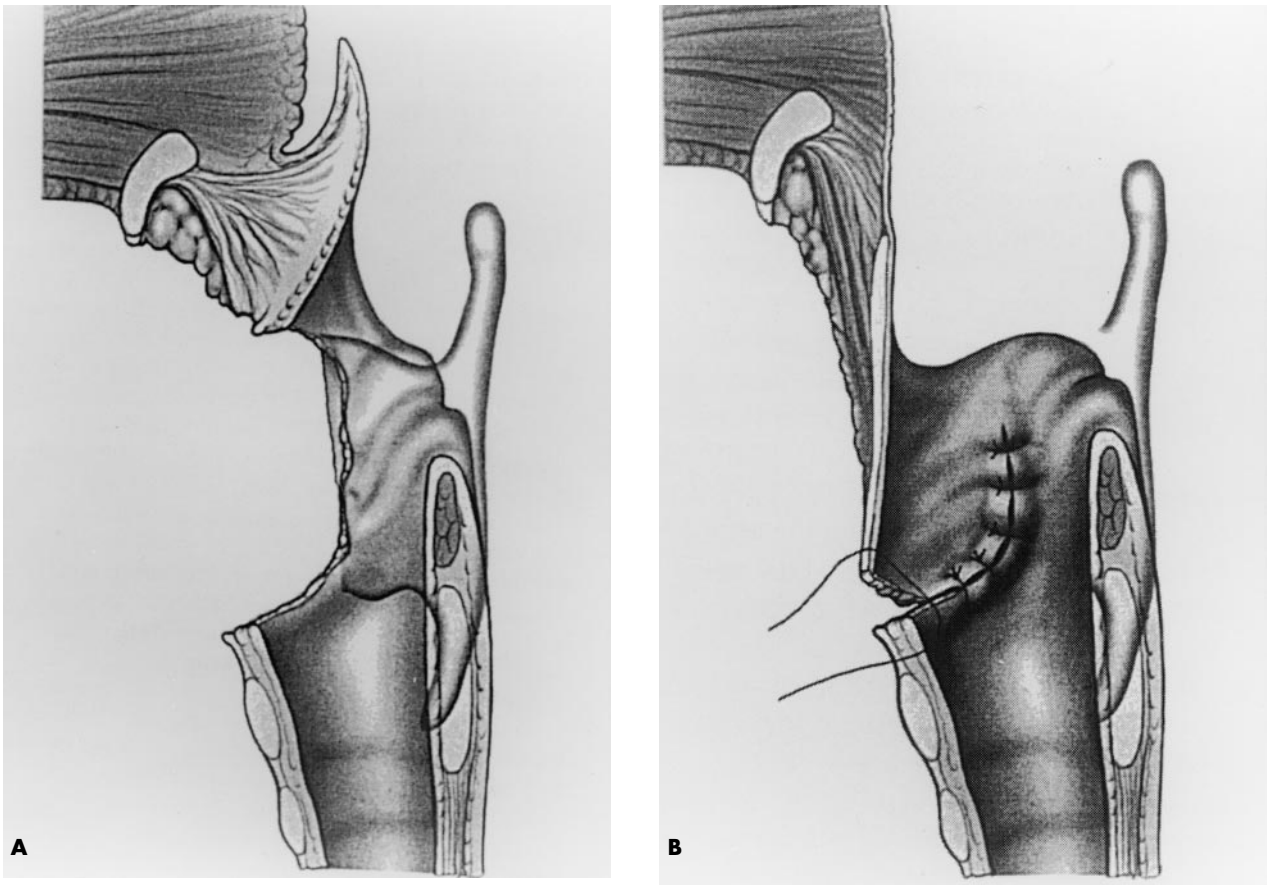


FIGURE 54–17. A, Schematic diagram of resected anterior commissure and anterior parts of the vocal folds, seen in the midsagittal plane. *B*, The epiglottis is rotated downward to fill the defect. Reproduced with permission from Silver CE, Ferlito A, editors. Surgery for cancer of the larynx and related structures. 2nd ed. Philadelphia: WB Saunders; 1996.

recurrence at the primary site that cannot be salvaged with endoscopic surgery. The high rate of laryngeal preservation noted with this approach supports this assertion. Fourth, the procedure is cost effective. A report by Myers et al confirmed that endoscopic management of T1 glottic cancer was less expensive and equally as effective as XRT.[85] Finally, treatment is accomplished quickly, with little morbidity, and may be suitable for those patients who wish to avoid daily treatment with XRT over a protracted course of several weeks. A typical endoscopic resection of an early vocal fold cancer is shown in Figure 54–19.

Based on available information, it seems reasonable to employ XRT as initial treatment for most T1 vocal fold lesions, except for those that appear to have minimal invasion of the subepithelial layer of the vocal fold, as determined by videostroboscopy or by examination at the time of direct microlaryngoscopy. T2a lesions are probably best treated by XRT, whereas T2b lesions should be considered for primary surgery, with XRT as the second option. Lesions defined as Cis based on biopsy are best managed by excision until they involve areas that cannot be followed adequately by office endoscopy, at which time XRT should be considered.

Management of Early Supraglottic Cancer

Early supraglottic cancer is rarely encountered. Even T1 and T2 primaries have a significant rate of lymph node metastasis. Management of the neck is critical since uncontrolled neck disease is the major reason for failure to control supraglottic carcinomas. For the purposes of this discussion, early supraglottic cancers are defined as T1 or T2 primary lesions, and discussion of neck management is similar to that

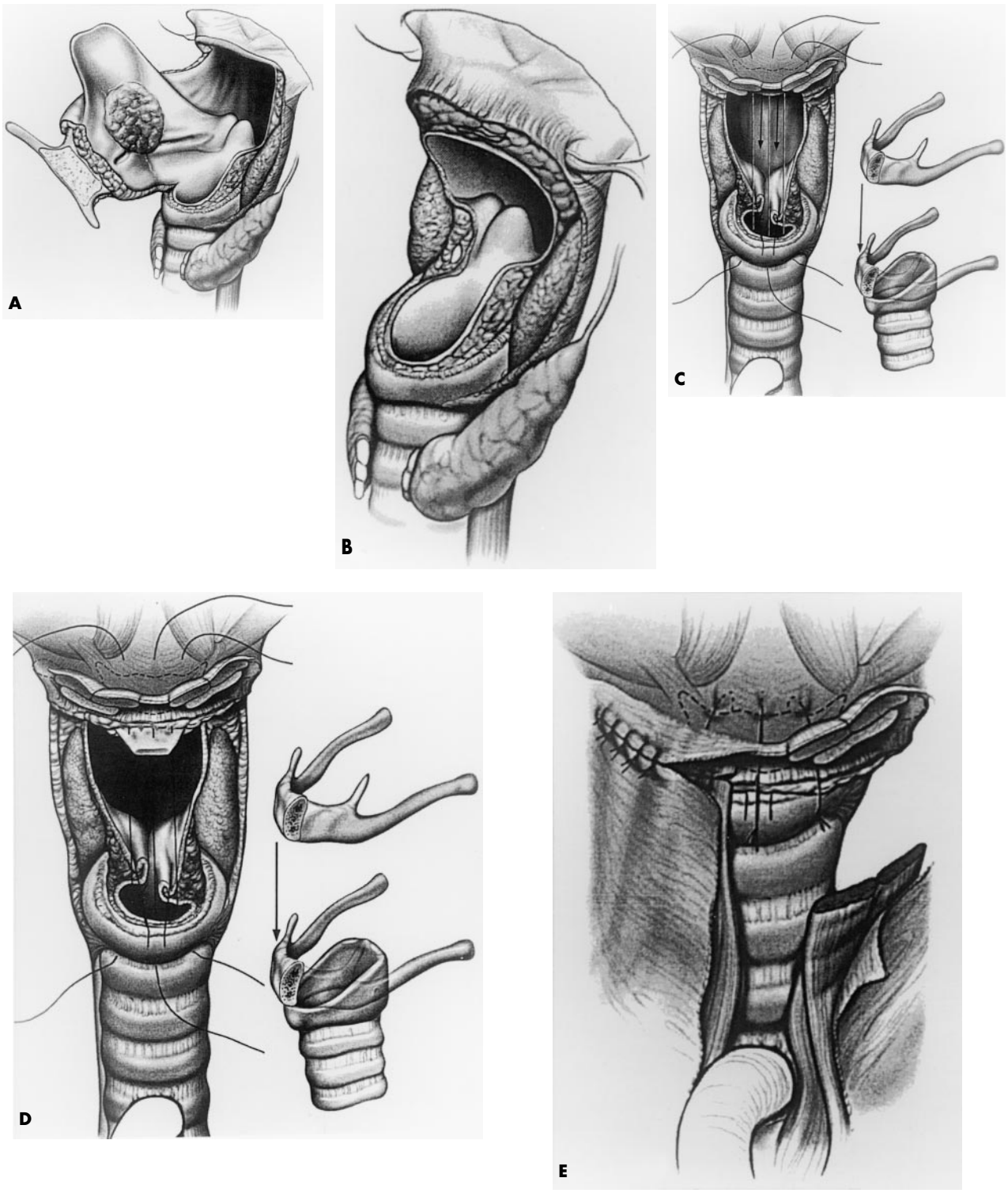


FIGURE 54–18. Schematic diagram of cricohyoidoepiglottopexy (CHEP) and cricohyoidopexy (CHP) procedures. *A*, Exposure and tumor cuts are shown for CHP. *B,* The defect after resection is shown for CHP. *C*, The vocal fold remnants are sutured to the anterior part of the cricoid cartilage to maintain their position and tension. Heavy sutures (eg, 2-0 polyglactin) are used to approximate the cricoid cartilage to the hyoid bone. *D*, In CHEP, sutures are passed through the epiglottis and base of the tongue to prevent epiglottic prolapse, but the closure is otherwise similar. *E*, The wound after closure of the larynx is complete.

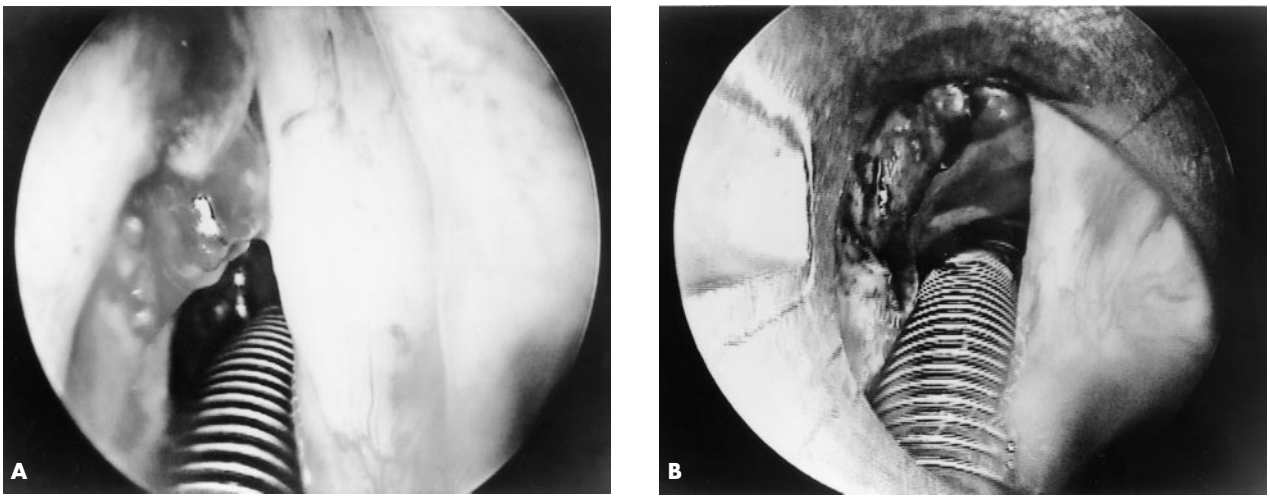


FIGURE 54–19. A, T1 carcinoma of the anterior part of the vocal fold viewed through direct microlaryngoscopy. *B*, Appearance of the same larynx immediately after excision of the lesion with carbon-dioxide laser.

discussed in the section on advanced supraglottic cancers later in this chapter.

Early supraglottic cancers can be managed either with surgery alone or radiation alone, although neck node involvement may mandate combined modality therapy. T1 primary tumors are rare since they usually remain silent as isolated lesions on the epiglottis. The limited reported experience with these lesions indicates that local control can be achieved with either surgery or radiation in over 90% of patients. T2 lesions involve two subsites of the supraglottis, without vocal fold fixation. Such tumors often involve the epiglottis and one aryepiglottic fold or other subsite combinations. Most reports, even though retrospective, indicate excellent local control with surgery that compares very favorably to conventional XRT. A study comparing primary surgery with primary XRT for T1 to 4N0 supraglottic carcinomas yielded identical 5-year survival rates for T categories 1 to 3, with an improved survival for T4 tumors treated surgically (56 versus 14%).[73] Local control rates for T2 lesions treated with XRT vary widely, from 53 to 89%,[86,87] but are generally similar to or slightly less than local control rates for surgery.

Surgery for Supraglottic Cancer Partial supraglottic surgery was performed as early as 1920 via transhyoid pharyngotomy. Alonso described a two-stage procedure for supraglottic resection that was performed through a lateral pharyngotomy. Ogura combined this approach into a one-stage procedure that has been widely used since the late 1950s. This supraglottic laryngectomy or horizontal partial laryngectomy encompasses the epiglottis, aryepiglottic folds, and false vocal folds and is illustrated in Figures 54–20 and 54–21. It can be extended to include one arytenoid cartilage, part of the tongue base, and the vallecula. If the arytenoid cartilage is excised, the true vocal fold is pexed in the midline to the posterior part of the cricoid cartilage to facilitate glottic closure for speech and swallowing. Closure of the defect created is almost always under tension. The preserved outer perichondrium of the thyroid cartilage and strap muscles are sutured to the base of the tongue anteriorly, and the cut edge of the aryepiglottic fold is sutured to the lateral part of the base of the tongue and pharynx. Some surgeons keep the patient's neck flexed postoperatively for 10 to 14 days, in some instances suturing the chin to the chest to prevent neck extension. Contraindications to the procedure have been cited as vocal fold fixation, tumor extension to 5 mm or less from the anterior commissure (often difficult to evaluate by preoperative endoscopy or imaging), and thyroid cartilage invasion. Patients must have excellent pulmonary function to be candidates for the operation.

For tumors that extend to the posterior parts of the aryepiglottic folds and the glottis, an extended or "three-quarter" laryngectomy has been described. It combines a supraglottic laryngectomy with a ver-

tical hemilaryngectomy. Reconstruction requires rotation of either the saved superior cornu of the thyroid cartilage or the posterior edge of the thyroid ala to reconstruct a stable pseudocord pedicled on the inferior constrictor muscle and covered with mucosa advanced from the hypopharynx (Figure 54–22).

If a cancer involves the piriform sinus by extension from the supraglottis, an oblique extension of the standard supraglottic laryngectomy can be made to encompass the lesion (Figure 54–23). Involvement of the piriform apex, the thyroid or cricoid cartilage, the glottis or subglottis, or an anatomic site beyond the laryngopharynx is a contraindication to this procedure. The defect created is usually too large for primary closure, and a pedicled myocutaneous flap (pectoralis major, latissimus dorsi) or a free flap (radial forearm) is the best means of reconstruction. The free flap usually retains vibratory properties owing to its pliability and minimal thickness, resulting in better voice results than myocutaneous flaps.[88]

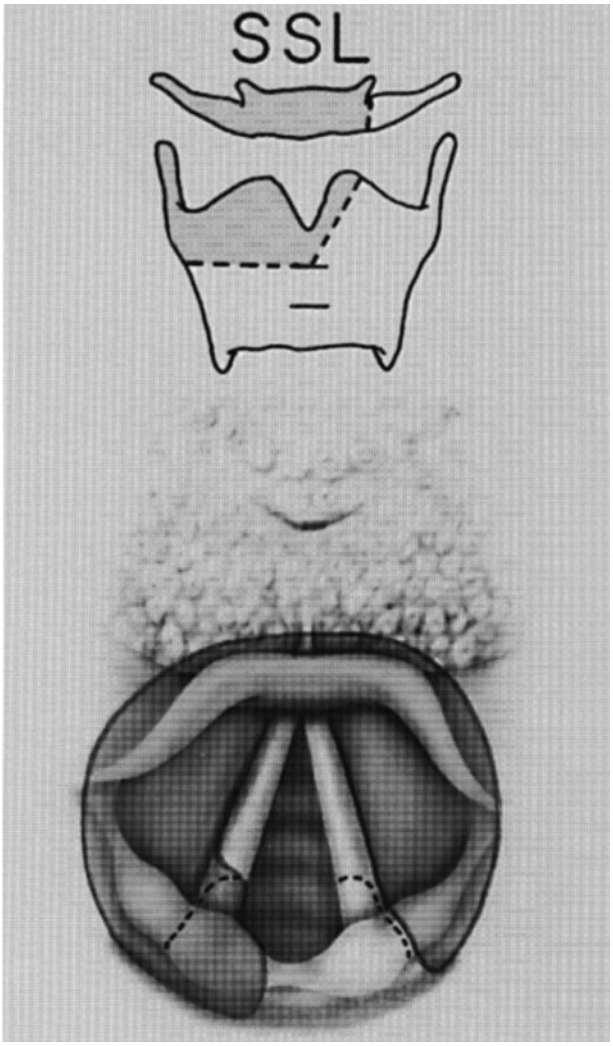


FIGURE 54–20. Schematic view of a subtotal supraglottic laryngectomy (SSL). The cartilage and bone cuts are shown above. The true vocal folds and at least one arytenoid cartilage are spared.

Cricohyoidoepiglottopexy and cricohyoidopexy should be kept in mind as alternatives to traditional open supraglottic laryngectomy techniques. These approaches are applicable to selected T2 supraglottic cancers thought not to be amenable to supraglottic partial laryngectomy because of ventricular invasion, glottic extension, and/or impaired motion of the vocal fold.

Transoral Carbon Dioxide Laser Resection of Supraglottic Cancers

The slow acceptance of endoscopic techniques for supraglottic cancers bears similarities to the delayed acceptance of open partial laryngectomy techniques 40 to 50 years ago. New developments in instrumentation, particularly the introduction of "split" or bivalved laryngoscopes, have enhanced exposure that is needed for this kind of surgery (Figure 54–24). Reports of relatively large series from Europe[89–91] and the United States[92,93] indicate excellent local control rates for T1 and T2 supraglottic lesions. For T1 lesions of the suprahyoid epiglottis, there is little controversy that this technique is effective. Proponents of transoral carbon dioxide resection cite low local recurrence rates for T2 lesions, similar to those for open supraglottic laryngectomy approaches, as support for their method.[94] Even T3 lesions based on preepiglottic space invasion can be managed endoscopically since the entire inner perichondrium of the thyroid cartilage and preepiglottic fat can be removed (Figure 54–25). If limited invasion of the thyroid cartilage is found, the procedure can be converted to an open approach, or the area can even be resected endoscopically with the laser or cold steel instrumentation.

The technique for larger lesions that cross the midline is quite different than those used in open techniques. The tumor is split vertically in the midline and is removed in at least two segments. Splitting the tumor actually allows the surgeon to see with the operating microscope and feel with instru-

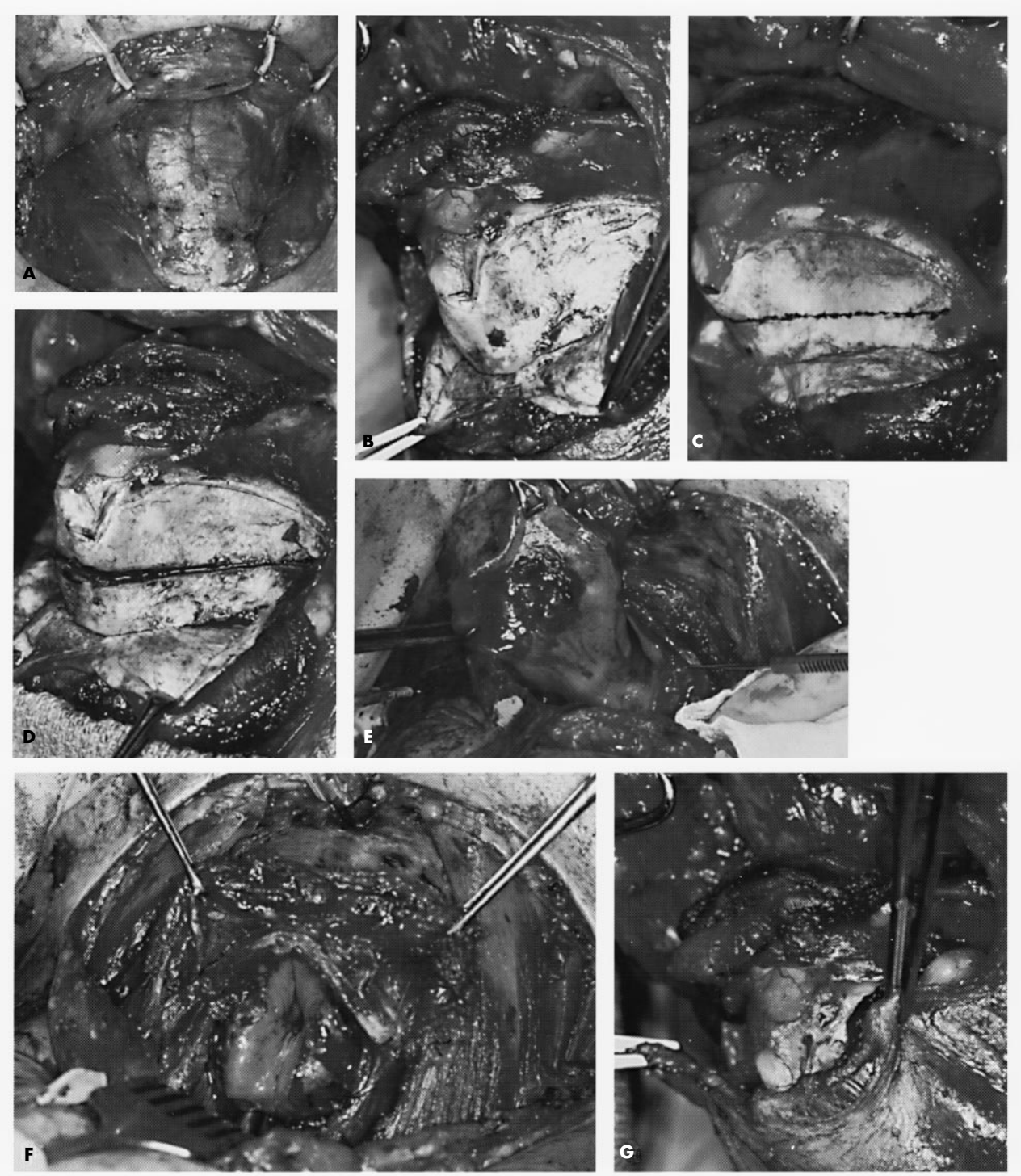


FIGURE 54–21. Subtotal supraglottic laryngectomy. *A*, Anterior neck exposure. *B*, Perichondrial elevation. *C*, Horizontal subperichondrial incision outlined. *D*, Incision completed with oscillating saw. *E*, Resection of the supraglottis with direct vocal fold visualization. *F*, View of defect after resection. *G*, Strap muscles and perichondrium are elevated; closure of perichondrium and strap muscles to base of tongue.

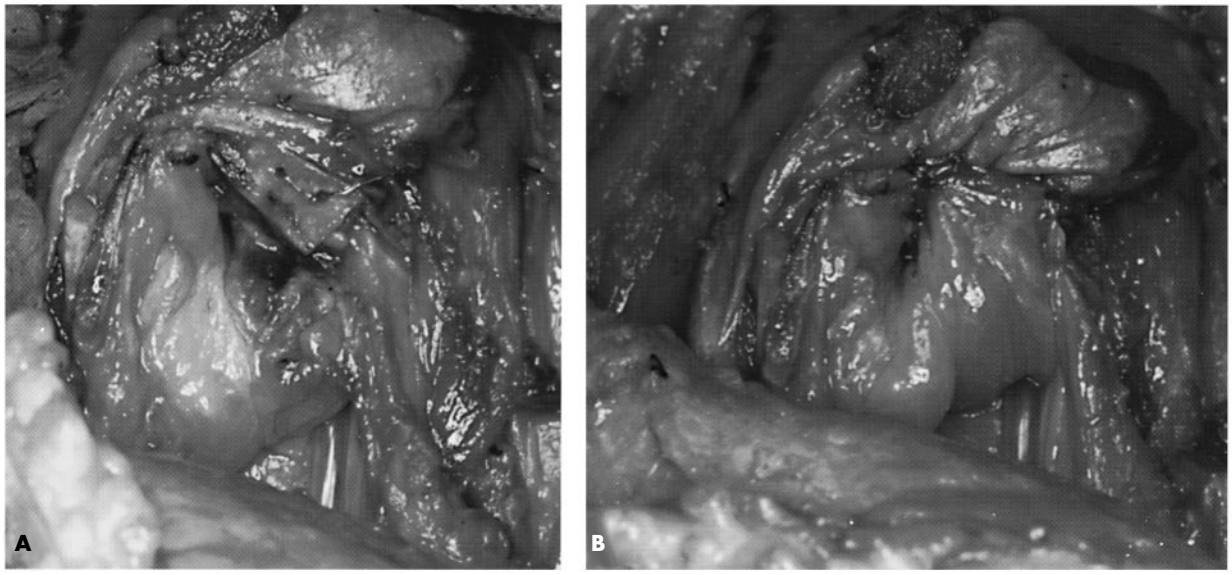


FIGURE 54–22. Three-quarter laryngectomy reconstruction. *A*, In-fractured thyroid ala is sutured in place. *B*, In-fractured thyroid ala is covered with postcricoid mucosa.

ments where the interface is between tumor and normal tissue. Debulking the tumor can also facilitate exposure without compromising the resection. Frozen sections are obtained after tumor removal is completed. Proponents of the endoscopic laser resection point out that their patients have less morbidity than those treated by open techniques in that tracheostomy is usually avoided, hospital stay is dramatically shorter, and time to resumption of speech and swallowing is considerably shortened.[92] Many patients require nasogastric tube feedings for only a few days, and some can swallow shortly after surgery and do not require a feeding tube. It seems likely that this technique will gain popularity in the United States in the future and will probably be used more frequently than open supraglottic laryngectomy procedures.

It would appear that early supraglottic cancers can be managed well with either radiation or surgery. In patients who are N0, single-modality therapy that includes both sides of the neck, as well as the primary site, is often sufficient. Surgery has the advantage of providing information on the adequacy of primary treatment (margins) and true pathologic staging of the neck. Whereas XRT may be less morbid, voice results are similar with surgery, and newer, less invasive surgical techniques may make the cost and morbidity of the two approaches equivalent. Patients with N+ necks almost always require combined surgery and radiation, particularly for advanced (N2 and N3) neck disease. In such cases, XRT for the primary site and combined treatment for the neck seems rational since XRT tends to adversely impact the functional results of conservation surgery of the supraglottic larynx.

Other Methods for Treating Early Laryngeal Cancers

Photodynamic therapy has been used to treat early vocal cord cancers, particularly recurrences after radiation failure. This method employs a photoactive drug that is administered systemically to the patient up to several days in advance of treatment, depending on the kinetics of the drug. The lesion is viewed endoscopically and is treated with a laser of appropriate wavelength to activate the drug to a toxic form in tumor cells that have been exposed to the laser light. Although a number of photoactive drugs have been investigated, most patients have been treated with hematoporphyrin derivative (HPD). The advantage of this approach is that there is little damage to nontumor tissue if precautions are taken to limit laser exposure to the tumor only. Another advantage is that unlike radiation, the treatment can be repeated as many times as needed, either for incomplete response

or for recurrence. For Cis and T1 glotttic cancers, complete responses approached 90% with a single treatment, and a cure rate of 95% was seen with a 79-month follow-up period.[95] The effect on tissue is superficial, but that is all that is required on the true vocal fold in most cases. The disadvantage is that, after the photoactive drug is administered (4 days before photoactivation for HPD), the patient must avoid exposure of the skin, eyes, and mucous membranes to light for an extended period of time (about 1 month for HPD). Severe burns have occurred in patients when adequate precautions were not taken. The photoactive drugs are also very expensive, and the laser used is proprietary and does not employ a wavelength found in other commonly used lasers. If a photoactive drug with a short half-life can be developed at a competitive cost, this technique will probably gain wider acceptance.

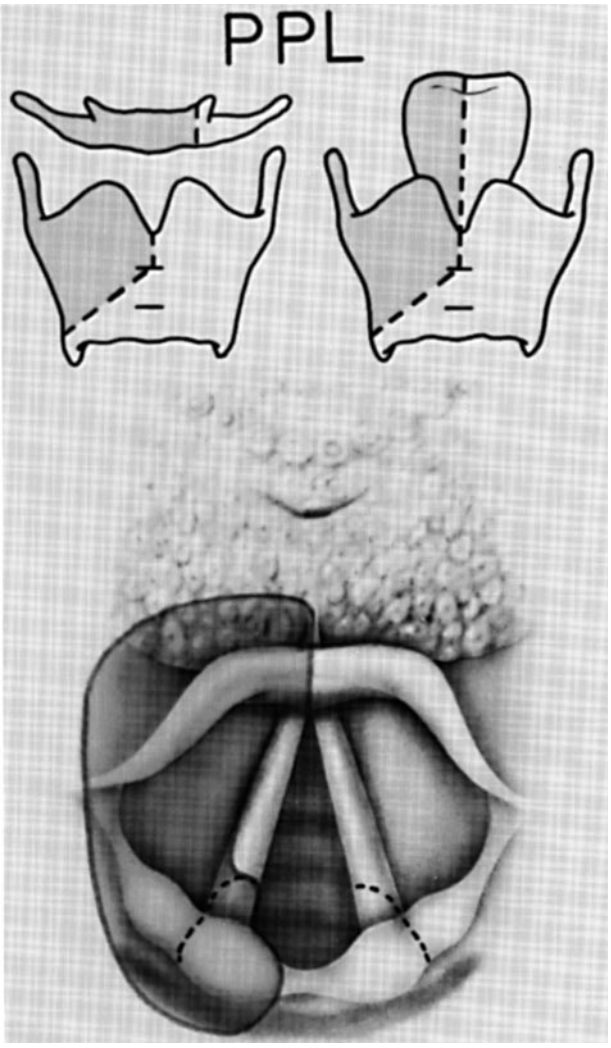


FIGURE 54–23. Schematic drawing showing the extent of cartilage and soft tissue resection for a partial laryngopharyngectomy. An oblique cartilage resection is performed to include the piriform sinus. PPL = partial pharyngolaryngectomy.

Chemotherapy for Early Glottic Cancer

There is little experience with chemotherapy as the sole treatment for early laryngeal cancer. One intriguing report of 21 patients with T1 to 3N0 glottic cancers who exhibited a complete response to *cis*-platinum/5-fluorouracil chemotherapy had a 5-year survival rate of 95%, local control rate of 71%, and metachronous second primary incidence of 14%. None of these patients developed regional or distant metastases. All local recurrences were salvaged with either partial laryngectomy or XRT.[96] Such an approach may merit further investigation, but the potential toxicity of treatment must be weighed against the excellent results that are already available with the many approaches described above.

Management of Advanced Laryngeal Cancer

Advanced cancer of the larynx includes stage III and IV tumors. Most such tumors are T3 or T4 at the primary site. Any size primary tumor that presents with a neck metastasis is considered advanced. It is fairly uncommon to see neck metastasis at presentation with a T1 or T2 glottic tumor or a T1 supraglottic tumor, but in such patients, management of the primary tumor is usually straightforward, with the neck metastasis being of greater concern and having the greatest impact on survival. For practical purposes, advanced cancer will be defined here as T3 or T4 disease, with or without cervical lymph node metastases.

Surgical management for laryngeal cancer is suggested in reports in the medical literature as early as 1801. It is not clear whether these early operations were for cancer or for benign lesions. In 1874, Billroth reported the first total laryngectomy. Although there was initial resistance to this procedure, other surgeons subsequently confirmed the safety and efficacy of this procedure. Numerous operations less radical than a total laryngectomy were described before and after its introduction. Most of these "conservation" procedures were designed for smaller (T1 and T2) primaries.

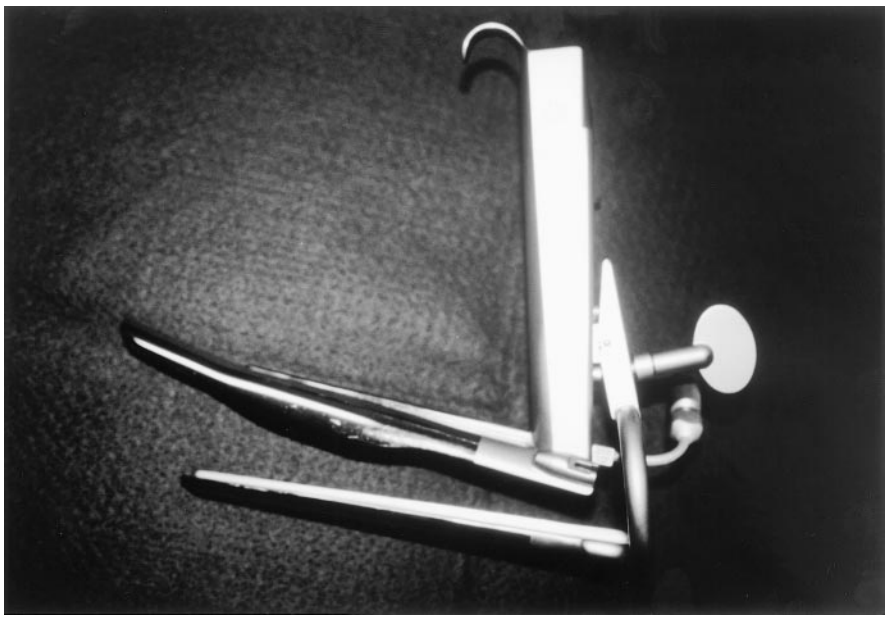


FIGURE 54–24. Bivalved laryngoscope essential for exposure needed to perform endoscopic carbon-dioxide laser resection of supraglottic cancers.

Other surgical procedures have been developed for larger laryngeal tumors with the goal of preserving and reconstructing enough of the larynx to allow retention of speech, swallowing, and breathing functions while providing tumor control rates comparable to total laryngectomy. Several types of "subtotal" laryngectomy have been described, most notably that of Pearson in 1981.[97] Piquet and associates[98,99] and Weinstein and Laccourreye[100] published extensive experience with supracricoid laryngectomy techniques that allow resection of selected advanced laryngeal and laryngopharyngeal cancers with primary closure of the defect. The cricoid cartilage is sutured to the retained hyoid bone or epiglottis, with preservation of one functioning cricoarytenoid unit being a prerequisite.

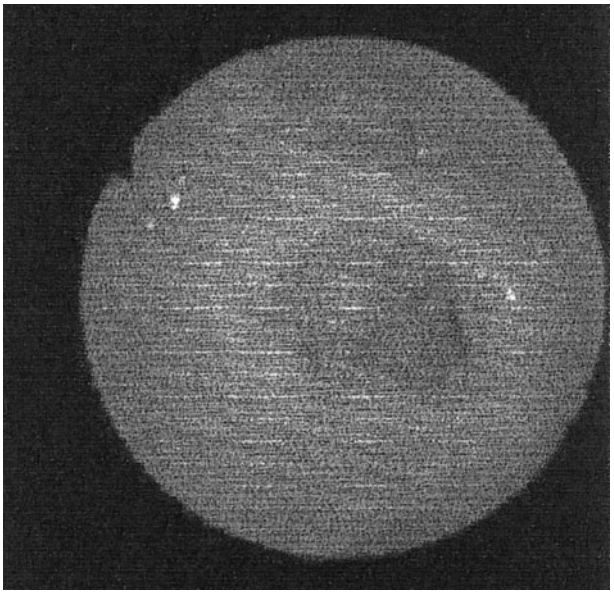


FIGURE 54–25. Postoperative view showing surgical defect after transoral endoscopic carbon-dioxide laser resection of a T2 supraglottic cancer. Part of the preepiglottic space was included in the specimen.

Primary radiotherapy has been used extensively to treat advanced laryngeal cancers, particularly in centers outside the United States. When radiation is used as single-modality therapy, surgery is kept in reserve for salvage of treatment failures. In general, when this approach has been analyzed critically, radiation alone has not fared as well as surgery for resectable lesions.[101–104] Vocal fold fixation is considered a relative contraindication for radiation because it portends a high failure rate. As newer radiation technology and hyperfractionated treatment schedules have been introduced, some have challenged this concept.[105]

Management of Specific Subtypes of Advanced Cancers

Advanced Supraglottic Cancer T3 supraglottic cancer, by definition, involves the preepiglottic space, invades cartilage, extends to the medial wall of the

piriform sinus, or causes vocal fold fixation. There is also a very high probability of lymphatic metastasis. Treatment options include surgery alone, surgery combined with preoperative or postoperative radiation therapy, primary radiotherapy, or a combination of chemotherapy and radiotherapy, with surgical salvage for failures of the latter two options. Controversy exists over the best treatment, and planning tends to be individualized according to tumor size and location, age and general health status of the patient, presence or absence of vocal fold fixation, estimated pulmonary function, other patient parameters, and the wishes of the patient. Preservation of laryngeal function without compromising survival should be attempted whenever possible.

T3N0 supraglottic cancers can be treated by supraglottic laryngectomy as long as patient selection criteria are strictly followed. The operation should not be attempted if there is vocal fold fixation, extensive involvement of the piriform sinus, thyroid or cricoid cartilage invasion, or extensive involvement of the base of the tongue (to or beyond the circumvallate papillae). Preepiglottic space invasion is not an absolute contraindication to supraglottic laryngectomy. Surgical management of properly selected patients results in a 3-year disease-free survival of 75%.[106–108] Local recurrences occur in only 10% of patients at this stage undergoing this procedure. Because of the high rate of lymphatic metastasis, elective treatment of the neck is indicated. Bilateral supraomohyoid neck dissection is advocated as therapeutic (in pathologically N0 necks) and as a staging procedure in patients who are pathologically lymph node positive. Patients with more than one positive lymph node or with any lymph node exhibiting extracapsular spread should probably receive postoperative radiotherapy.

A major requirement for selection for supraglottic laryngectomy is adequate pulmonary function. Almost all patients exhibit some degree of aspiration after supraglottic laryngectomy, even those performed for T1 and T2 lesions endoscopically by carbon dioxide laser excision. Patients with poor exercise tolerance can be expected to be at high risk for aspiration pneumonia after supraglottic laryngectomy. Pulmonary function studies should be performed preoperatively in all patients being considered for this operation. Forced expiratory volume in 1 second of at least 75% is considered by many to be a requirement.

Even in properly selected patients, approximately 5 to 10% will eventually require a completion laryngectomy for intractable aspiration.[109] About 20% will require a prolonged tracheostomy (more than 3 months), most of whom are patients who have received postoperative radiation therapy. Rates of prolonged dysphagia after surgery are low but are directly proportional to the extent of surgical resection, with the highest rates observed in patients who have had part of the base of the tongue or piriform sinus removed. Reports from Europe on carbon dioxide laser endoscopic resection of large supraglottic cancers describe lower rates of dysphagia and aspiration than those observed for open procedures for similar-size lesions. Less tissue destruction and removal, absence of tension associated with wound closure in open procedures, and less damage to the superior laryngeal nerves are probable explanations for these observations. A high degree of technical training, skill, and experience is necessary to achieve good results with such endoscopic techniques.

Voice results for patients treated with supraglottic laryngectomy without postoperative radiation are equivalent to those for similar patients treated with primary radiotherapy.[110] Voice results after transoral laser excision of larger supraglottic cancers have not been systematically studied but are thought to be at least as good, if not better than, those achieved by open supraglottic laryngectomy. It is noteworthy that patients with supraglottic cancers amenable to supraglottic laryngectomy who are treated with total laryngectomy have survival rates that are no higher than similar patients treated with supraglottic laryngectomy.

Advanced supraglottic cancers that exhibit vocal fold fixation, postcricoid extension, cartilage invasion, or extension outside the larynx are usually treated with total laryngectomy and postoperative radiation. However, several organ-preserving strategies are applicable to such lesions and are discussed later in this section. Most advanced supraglottic cancers have positive neck lymph nodes or have a high risk of neck metastasis in the N0 neck. It is fair to say that the neck should be treated in all cases of advanced supraglottic cancer. Because most of the supraglottic subsites have bilateral lymphatic drainage, both sides of the neck require treatment. For the N0 neck, bilateral selective neck dissection encompassing levels II, III, and IV is considered suf-

ficient by most surgical oncologists. Although some argue that such surgery does not impact survival,[111] valuable pathologic staging information is gained, which certainly impacts the decision whether to use postoperative radiation. Patients with greater than one positive lymph node and those with any lymph node exhibiting extracapsular spread should receive postoperative radiation. Patients with N+ neck disease should have a more extensive neck operation, usually a modified radical or radical neck dissection on the involved side(s). Neck structures spared in such operations depend on the size and location of the involved lymph nodes. All such patients should receive postoperative radiation.

Radiation Therapy. Standard, once-daily radiation therapy has been investigated as single-modality treatment for advanced supraglottic cancers that are considered resectable by supraglottic laryngectomy. Higher local failure rates, lower survival rates, and lower rates of laryngeal preservation have been reported for patients treated with radiation for cure, with surgery reserved for salvage of treatment failures.[101–104] In most of those patients, surgical salvage consists of total laryngectomy, and salvage surgery is usually possible in only 50% of such patients.[112] Twice-daily or hyperfractionated radiation therapy schedules have been reported to produce improved local control rates for advanced supraglottic cancers compared with standard, once-daily fractionation schedules.[113] This radiation approach will be discussed in greater detail in the section on advanced glottic cancer.

Advanced Glottic Cancers Until the last decade, advanced glottic cancers were treated either by total laryngectomy or by radiation therapy for cure, with total laryngectomy in reserve for surgical salvage. Although the merits of the two approaches were debated extensively in the medical literature, a prospective, randomized, controlled study was never performed. Proponents of surgery as primary therapy claimed higher survival rates, whereas those espousing radiation first pointed out the advantage of voice preservation in those individuals who did not relapse after radiation. Many studies were retrospective, and survival rates for T3 lesions ranged from 49 to 80%[73,114,115] and for T4 lesions from 32 to 63%.[73,115]

The technique of total laryngectomy has changed very little in the last century. A radical, en bloc resection of the larynx is performed. The hyoid bone is included in the resection, and the trachea is transected at the highest level that is oncologically feasible to facilitate creation of a tracheostoma of sufficient size that can be sutured to the skin with as little tension as possible. The operation can be extended to include adjacent involved areas of the base of the tongue and hypopharynx. When a partial or total pharyngectomy is added to a standard total laryngectomy, some sort of flap reconstruction is usually required. The main flaps used are the pectoralis major myocutaneous flap and the radial forearm free flap, especially when a strip of intact normal hypopharyngeal mucosa can be maintained between the oropharynx and the esophagus. When total pharyngectomy and/or esophagectomy are performed, visceral reconstruction is usually achieved with either a jejunal free flap or, more commonly, gastric pull-up. For piriform sinus tumors and for tumors with greater than 10 mm of subglottic extension, a thyroid lobectomy or subtotal thyroidectomy should be performed, depending on the extent of spread in relationship to the midline.

Radiation treatment for laryngeal cancer dates to 1903. As the technology involved in radiation treatment has improved and as our understanding of radiobiology has increased, so has the methodology advanced in delivering radiotherapy to patients. Today, there are numerous radiation fractionation strategies and schedules that are used for treating advanced laryngeal cancer.

It was initially felt that radiation was effective against cancer because tumor cells were more sensitive to ionizing radiation than normal cells of the same tissue type. Although this has been found not to be the case from tissue culture studies, it is clear that tissues with a large portion of proliferating cells (high growth fraction) are more susceptible to the lethal effects of radiation. The percentage of cells in the proliferative phases of the cell cycle is one determinant of radiation sensitivity. The ability of cells to repair DNA damage caused by radiation is now recognized as an important factor in determining the effect of radiation and chemotherapy on cancer cells. Depending on the degree of damage and the cell's ability to repair the damage, the cell will either go into growth arrest (to repair damaged DNA) or go on to apoptosis (programmed cell death). Certain

strategies, such as twice-daily radiation (hyperfractionation), are designed to kill more cancer cells by treating in more fractions, over a shorter interval, to confer lethal damage to cells that might otherwise repair themselves. Chemotherapy concurrent with radiotherapy may act as a radiation sensitizer in the same manner or by forcing more of the tumor cell population into phases of the cell cycle in which they are more susceptible to the effects of radiation.

A prospective trial concerning XRT only for laryngeal cancer was conducted at the Princess Margaret Hospital in Toronto.[113] Patients with advanced head and neck cancer were randomized to either once-daily XRT to a total dose of 51 Gy in 20 fractions delivered in 4 weeks at 2.55 Gy per day or twice-daily XRT to a total dose of 58 Gy delivered in 40 fractions in 4 weeks at 1.45 Gy twice a day at least 6 hours apart. Of the 336 patients entered, 116 had T3 or T4 laryngeal cancers, of which 30 were glottic and 86 supraglottic. Fifty-one were T3 and 65 were T4. Local control at 3 years was achieved in 50% of the once-a-day treatment arm and 54% in the twice-daily treatment group. Overall survival was 47% for once-daily XRT and 69% for twice-daily XRT, a significant difference. A number of other trials comparing once-a-day with twice-a-day radiation schedules have been reported, most indicating improved locoregional control with twice-daily XRT but mixed results in terms of any impact on survival.[116-121] Most of these studies have involved large numbers of patients from all head and neck subsites and have not focused specifically on the larynx.

Organ Preservation Approaches Today the emphasis has shifted to organ-preserving strategies, both surgical and nonsurgical, for managing advanced glottic (and supraglottic) cancers. Perhaps the major impetus for this trend was the Department of Veterans Affairs (VA) Laryngeal Cancer Study Group, which initiated a multicenter, randomized, prospective trial to evaluate a laryngeal preservation strategy in patients with locally advanced, resectable laryngeal cancer.[122] The majority of the patients entered had advanced glottic cancers, but some supraglottic cancers were also included. All patients entered would normally have been candidates for total laryngectomy. Patients were randomized to one of two treatment arms. One arm was standard therapy with total laryngectomy and postoperative radiation. The other arm was induction chemotherapy with *cis*-platinum and 5-fluorouracil for two to three cycles. Patients exhibiting at least a partial response were then treated with radiation therapy. Patients who did not exhibit a major response to chemotherapy were treated with laryngectomy followed by radiation therapy. At 60 months median follow-up, there was a similar survival rate in the two treatment arms (68%). Patients in the chemotherapy arm had a higher rate of local recurrence, whereas those in the standard therapy arm had a higher rate of distant metastasis. Survival was equivalent for chemotherapy responders and nonresponders.

The results from this trial are important because they showed in a randomized prospective study that survival was equivalent between an organ preservation strategy and standard therapy that included laryngectomy. Laryngeal preservation occurred in 64% of the survivors in the chemotherapy arm without compromising survival. It should be noted that 65 of the 166 patients entered in the chemotherapy arm (40%) actually were alive with a larynx at the end of the observation period.

There have been a number of criticisms aimed at the VA study. First, there was no treatment arm in which patients were randomized to radiation only. A number of centers have reported survival rates and laryngeal preservation rates with radiation alone that are similar to those achieved in the two arms of the VA study. The role of chemotherapy in this study has also been questioned. Was the chemotherapy beneficial as a therapeutic agent, or did it just select those patients who would be likely to do well with radiation alone? Some studies, including meta-analyses, have even shown a negative effect of induction chemotherapy on survival in head and neck cancer.[123] Another criticism of the study is that the quality of speech and swallowing in the patients treated in the two arms was not compared. Even though many patients retained their larynx in the chemotherapy arm, the functionality of the preserved organ was not addressed, nor was it compared to the results after total laryngectomy with a tracheo-esophageal puncture (TEP) or other means of speech rehabilitation.

Despite these criticisms, it is clear that laryngeal preservation strategies are worth considering since some larynges can be saved without compromising overall survival. Even though the original VA trial is open to criticism, many community practitioners

have adopted the sequential chemotherapy and radiation strategy of the VA study in their management of advanced laryngeal cancer. Guidelines from the National Comprehensive Cancer Network list induction chemotherapy as an option for T2 and T3 hypopharyngeal cancer and for laryngeal cancers that would require total laryngectomy.[124] This is of concern since standard and hyperfractionated radiation-only approaches result in reported cure rates similar to those in the VA study.[125,126] Furthermore, there is a definite trend emerging from the results of a number of trials of concurrent chemotherapy and radiation therapy that indicate results superior to sequential chemotherapy and radiation. So far, there have been no trials published evaluating concurrent chemoradiation exclusively in advanced laryngeal cancers. However, a second VA trial is under way in which there are three arms: the previously used induction chemotherapy arm, a radiation-only arm, and a radiation and concurrent single-agent chemotherapy arm. The results of this trial will soon be available and should provide important information on the relative merits of these organ preservation strategies.

Surgeons have also been active in the area of organ preservation for advanced laryngeal cancers. Pearson described the subtotal laryngectomy in the early 1980s.[97] Although some natural phonatory function without an artificial device could be achieved, a permanent tracheostomy was required. This procedure did not seem to offer a great advantage over a total laryngectomy with adequate voice rehabilitation. Supracricoid laryngectomy was introduced by Piquet and associates[98,99] and has been popularized by others.[100] The two main procedures, CHEP and cricohyoidopexy, overcome the need for a permanent tracheostomy and are illustrated in Figure 54–18. In these procedures, at least one functioning cricoarytenoid unit, the hyoid bone, and an intact cricoid ring are required. Survival rates of 80 to 85% have been reported for patients with T3N0 cancers using these techniques.[127,128] Proper surgical technique and patient selection are important for success with these procedures.

Others have used induction chemotherapy to downstage laryngeal primary lesions so that partial laryngectomy can be performed rather than total laryngectomy. Brasnu et al reported a series of 60 patients with T3 and T4 supraglottic and transglottic carcinomas managed this way.[129] They reported 5-year survival of 72.7%, with local and regional failure rates of 8.3 and 9.2%, respectively. Another study reported on a series of 69 patients who were initially candidates for total laryngectomy who were treated with induction chemotherapy. Thirty-three patients were selected for partial laryngectomy (n = 14) or radiation (n = 19) and had their larynx preserved. Rudert and others have even reported successful application of transoral carbon dioxide laser resection to selected T3 laryngeal cancers, particularly supraglottic cancers with extension to the central portion of the preepiglottic space.[130]

In summary, great controversy continues to surround the choice of treatment for advanced cancer of the larynx. Insufficient objective data exist to differentiate between standard therapy (laryngectomy with postoperative XRT), altered fractionation XRT, chemoradiation protocols, and the newer, function-preserving operations for selected advanced lesions. Treatment continues to be individualized, with such factors as the patient's age and general health, occupation, ability to use alternative means of communication, and the patient's own wishes all coming into play. Quality of life concerns, in addition to survival data, will increasingly be used to aid in the decision-making process.

Voice Results after Treatment for Laryngeal Carcinoma

When survival outcomes for different treatment regimens for the same stage of disease are similar, voice quality after treatment becomes an important factor in choice of therapy. Measurement of voice can be done using objective mechanical or physical parameters (shimmer, jitter, signal-to-noise ratio, fundamental frequency, intensity), aerodynamic properties (subglottic pressure, maximal airflow), and general phonatory measures (eg, maximum phonation time). Such measures are quantifiable, but voice quality is perceptual and therefore subjective. Most studies comparing voice outcomes employ assessments by trained listeners.

Early Lesions For Cis and microinvasive carcinoma, videostroboscopic examination is critical. If there is any propagation of the mucosal wave through the lesion, then microsurgical excision with margins can usually be performed with an excellent voice result.

The voice results reported after endoscopic carbon dioxide partial cordectomy for early cancers not invading the anterior commissure, vocal process, ventricle, or subglottis are surprisingly good. McGuirt et al reported that 83% of 15 patients so treated achieved normal glottic closure, 61% had a demonstrable mucosal wave, and all were rated as either normal or mildly abnormal by a panel of expert listeners.[131] A more recent study resulted in lower rates of mucosal wave preservation, but voice results were generally very acceptable.[132] Based on the few reported studies comparing radiation to laser resection for T1 lesions limited to the middle third of the vocal fold, subjective assessment reveals similar results for the two modalities. Objective measures tend to be better in the patients receiving radiation. The quality of the voice after laser excision appears to decline in proportion to the amount of vocalis muscle resected.[132]

The voice results reported after endoscopic carbon dioxide cordectomy for early cancer not involving the anterior commissure, vocal process, ventricle, or subglottis are surprisingly good. The entire vocal fold is removed from the anterior commissure to the vocal process and laterally to the inner perichondrium of the thyroid cartilage. The postoperative voice in patients treated this way is breathy, weak, and harsh, and complete glottic closure rarely occurs.[133] The voice results when radiation is used for similar lesions are almost always superior.

Hemilaryngectomy produces variable voice results, but almost all patients experience harsh, breathy voices and incomplete glottic closure. Voicing often is accomplished by recruitment of supraglottic structures to achieve laryngeal closure.[134] Although there are no data comparing the voice results of the many reconstruction methods used for the hemilaryngectomy defect, it appears that the postoperative voice is better in those patients who undergo reconstruction than in those who are left to heal by secondary intention. There are no studies comparing voice results between patients having hemilaryngectomy and those receiving XRT for similar lesions. Since the majority of patients who undergo hemilaryngectomy are radiation failures, it would be difficult to conduct such a trial. Subtotal or near-total laryngectomy techniques result in a voice that is variable and unpredictable, with a higher fundamental frequency than that of normal speakers.[135]

Alaryngeal Speech There are several rehabilitative options available to individuals who have undergone total laryngectomy. The three major categories are electronic devices, esophageal speech, and TEP. These techniques are complementary and are not mutually exclusive. Most patients ultimately use only one of these forms of communication.

Electrolarynx. Electronic devices can be used shortly after laryngectomy and are often the initial method of oral communication. The instrument is battery powered and handheld and produces vibratory energy that is transmitted to the mouth and/or pharynx by placing the vibrating end against the submandibular area or cheek or via a plastic tube placed directly in the mouth. Although this option has the advantage of allowing early voicing after laryngectomy, the mechanical voice quality is a disadvantage. All consonants are voiced when the instrument is used, making intelligibility difficult at times.

Esophageal Speech. This form of alaryngeal speech does not involve artificial devices. The speaker swallows or "injects" air into the stomach and esophagus and expels it through the pharynx, speaking while the pharyngoesophageal segment vibrates. A minority of all laryngectomees actually master this technique, and maximum phonatory time (MFT) is short, requiring frequent gulping of air. Fundamental frequency is lower than in normal speakers. Nevertheless, motivated individuals who achieve esophageal speech can be excellent speakers who do not depend on mechanical devices and who can speak with both hands free.

Tracheoesophageal Puncture. This surgical/prosthetic type of voice restoration was introduced in 1980 by Singer and Blom[136] and rapidly became accepted as a means of voice restoration after laryngectomy. The technique consists of creating a tract between the superior aspect of the tracheal stoma and the neopharynx. The procedure is usually performed under general anesthesia, using a rigid esophagoscope to visualize the puncturing instrument as it enters the lumen of the pharynx, protecting the posterior wall of the pharynx from injury. A stent (often a 16 Fr. feeding tube) is placed in the puncture tract, is secured to the skin with sutures, and is left in place for 5 to 10 days after the procedure to allow the tract to mature. The stent is removed,

and a prosthesis is inserted (Figure 54–26, A). Several companies make excellent prostheses.

When the stoma is occluded, air is forced into the prosthesis, and a vibratory sound is made in the pharynx, allowing the patient to voice during exhalation. In principle, this approach is similar to esophageal speech since expelled air creates a vibrating sound in the pharyngoesophageal segment. However, the volume of air that can be expelled from the lungs is much greater than that expelled from the esophagus, allowing greater volume and much longer MFT and providing some ability to modulate intensity.

Experienced listeners consistently rate the voices of TEP recipients superior to those of esophageal speakers.[137,138] Objective parameters of TEP more closely resemble normal speech than do those of either esophageal speakers or those using an electrolarynx. Maximum phonatory time for TEP speakers averages 12 seconds, compared with 1.9 seconds for esophageal speakers and 25 seconds for individuals with a normal larynx.[139]

Tracheoesophageal puncture can be performed as a primary procedure during laryngectomy or as a secondary procedure after the tracheal stoma is well healed. Voice results and complication rates are similar for primary and secondary TEP. Success and complication rates appear to be similar in radiated and nonradiated patients. Tracheoesophageal puncture can also be performed successfully in patients requiring free flap reconstruction of the pharynx in conjunction with laryngectomy. Voice quality is reduced compared with standard TEP but is still rated as good by the patients themselves.[140]

Some patients with TEP can wear an external valve over the stoma that allows them to phonate with their hands free. The flutter valve is secured to the skin with an adhesive (Figure 54–26, B). Patients with a flat neck contour around their stoma seem to have the most success with this device since adhesion to the skin is facilitated. It has been noted that some TEP users who do not use the external valve and have to use one hand to cover their stoma will return to esophageal speech because they prefer the hands-free advantage over the improved voice quality of the TEP.

Complications of TEP include bleeding, infection, aspiration, and failure to achieve speech. Many of the voice failures after TEP are thought to be owing to spasm in the pharyngoesophageal segment. Pharyngeal plexus lysis or cricopharyngeal myotomy has been shown to overcome this type of failure in a high percentage of patients.[141] Performing only a two-layer closure and avoiding closure of the constrictor muscles during the repair of the pharynx after laryngectomy is thought to reduce the incidence of pharyngoesophageal spasm and improve the chance for successful TEP after laryngectomy.[142]

Management of Cancer Recurrence

The means of managing recurrence of cancer of the larynx are determined by the extent of the original lesion and its treatment. Recurrence of early lesions after XRT can often be salvaged by surgery, either conservation procedures, where indicated, or by total laryngectomy. Recurrence after initial conser-

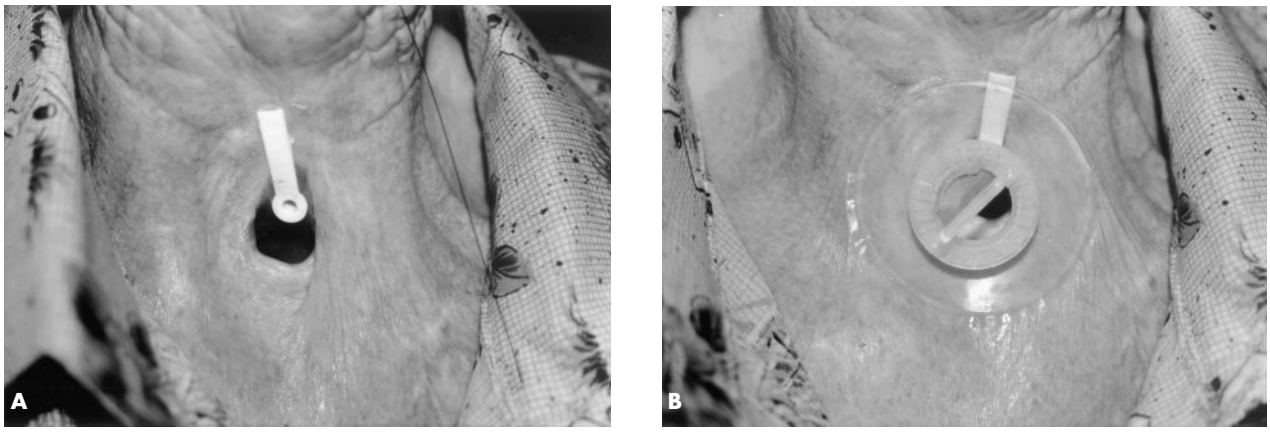


FIGURE 54–26. *A*, Blom-Singer voice prosthesis in place in tracheal stoma. *B*, External valve covering stoma, allowing hands-free phonation.

vation laryngeal surgery can be salvaged either by surgery (usually completion laryngectomy) or by radical radiation therapy if radiation was not used previously. Failure after radiation for larger primary tumors can be salvaged by total laryngectomy in 30 to 60% of patients. Failure in the neck can sometimes be salvaged by further surgery and/or XRT but usually portends a poor outcome, often with distant metastases.

Parastomal recurrence is a specific clinical entity in which tumor recurrence takes place near the tracheal stoma (Figure 54–27). Risk factors for this type of recurrence are thought to be pretreatment tracheostomy, subglottic extension greater than 10 mm, large tumor volume, and paratracheal lymph node involvement. Some have reported that survival is 0% with paratracheal lymph node involvement and have recommended elective unilateral or bilateral dissection of these lymph nodes in patients at risk for stomal recurrence.[143] Although some have devised aggressive surgical procedures, including en bloc resection of the stoma with mediastinal lymph node dissection, operative morbidity is high, and the salvage rate is very low (17 to 45%).[144] Recurrences confined to the area above the equator of the stoma, without involvement of the esophagus, appear to be the group amenable to salvage surgery. Computed tomography and MRI are very helpful in determining tumor extent prior to attempted salvage.

Complications of Treatment of Laryngeal Cancer

Radiotherapy Complications Complications similar to other sites are encountered, such as skin desquamation, tanning, or slough; mucosal ulceration and dryness; loss of taste function; hoarseness; dysphagia; and esophageal stricture. Another important complication is chondroradionecrosis of the laryngeal cartilages, which is similar to osteoradionecrosis of bone. Ulceration of mucosa over cartilage, either secondary to radiation itself or after biopsies or other instrumentation of the airway, leads to bacterial infection of radiated cartilage. Pain is the hallmark of this complication, often associated with increased swelling over the laryngeal cartilage landmarks. Low-grade fever may be present, and odynophagia is common. Differentiating this condition from recurrent tumor may be difficult, and biopsies and imaging studies are often inconclusive. Diagnosis is clinical, after attempts have been made to determine if there is tumor recurrence. Treatment consists of high-dose parenteral antibiotics and hyperbaric oxygen, usually 30 to 40 treatments. Failure to respond to treatment usually results in laryngectomy, and it is not unusual to find occult cancer in such specimens.

Surgical Complications The complications seen after surgery for laryngeal cancer are similar to those for most major head and neck cancer operations.

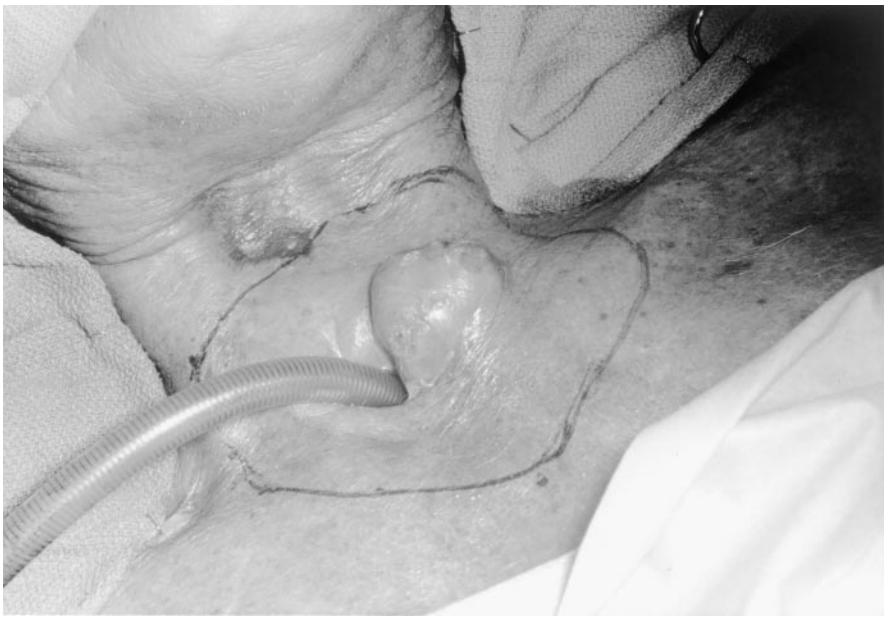


FIGURE 54–27. Parastomal recurrence after laryngectomy. Even with radical surgery in selected patients, prognosis is poor.

Bleeding, infection, fistula formation, aspiration pneumonia, stomal stenosis, and pharyngeal and esophageal stenosis are known to occur with total laryngectomy. Similar complications, plus glottic or supraglottic stenosis, can occur with hemilaryngectomy or supraglottic laryngectomy operations. Fistula formation occurs in 6 to 66%, with a lower incidence occurring in glottic tumor resections (6 to 15%). Higher rates of fistula formation occur with large supraglottic primaries, in all patients who have undergone previous XRT with or without chemotherapy for cure, and when the surgery is not done at a planned interval (usually 3 to 6 weeks) after completion of radiation. Fistulae after laryngectomy are especially dangerous since the carotid artery lies adjacent to the resection bed and is subject to rupture. Fistulae that leak around the stoma are also dangerous since direct aspiration via the tracheostoma is likely.

Distant Metastases

Patients with glottic primaries are the least likely to develop distant metastases, whereas supraglottic tumors that extend to the hypopharynx, especially the piriform sinus, are most likely to spread distantly. Lung, liver, and bone are the most common distant metastatic sites. Distant metastatic disease is also most likely to occur in patients with advanced (N2 or N3) neck disease and in patients who have recurred after initial treatment.

A dilemma arises with a solitary pulmonary lesion that is suspected to be metastatic in a patient with a laryngeal primary cancer since a second primary in the lung is a possibility. If there are no other metastatic deposits detected and the pulmonary lesion is resectable, the laryngeal tumor is treated first, followed by thoracotomy for the lung lesion. If the lung lesion is not resectable, the laryngeal tumor is treated with palliative XRT as needed.

The lung is the most commonly detected site for metastasis from laryngeal cancer, with 50 to 80% occurring there. Lung metastasis is usually associated with a mean survival of 12 months. Bone metastases make up 10 to 35% of laryngeal cancer metastases and portend a poor prognosis, with a mean survival of less than 4 months. The lesions are often lytic and painful and occur most often in the thoracolumbar vertebrae and the ribs. Liver metastases are quite common at autopsy, occurring in 80% of patients with distant metastases. They are often asymptomatic and undetected during life. Skin metastases, often to the neck, are common in preterminal patients.

Palliative Care of Laryngeal Cancer

Occasionally, patients present with such advanced disease and in such poor general health that curative treatment is either not possible or is too dangerous to undertake. About 7 to 12% of patients with laryngeal cancer will present with distant metastases. In such cases, palliative treatment with XRT may be useful. Symptoms of airway obstruction, pain, and dysphagia can often be improved by a course of XRT characterized by larger fraction size, fewer treatments per week, and shorter overall treatment time. In other patients, airway obstruction can be relieved by tracheostomy, and dysphagia and malnutrition can be ameliorated by placement of a gastric feeding tube.

Recurrent disease leading to death occurs in about 25 to 35% of patients with laryngeal cancer. Local and regional recurrence can be palliated with appropriate chemotherapy. A variety of drugs possess significant activity against squamous cell cancer and include *cis*-platinum, carboplatinum, the taxanes, 5-fluorouracil, etoposide, and methotrexate, to name a few. Palliative treatment regimens should be designed to produce as little toxicity as possible, while holding out a reasonable chance for amelioration of symptoms. Such treatment should not be offered or given to patients who are not having major symptoms (pain, dysphagia, dyspnea) and should be held in reserve until needed. As with all cancer patients, the patient and his/her family need to be included in the decision-making process when palliative care is considered. Entry into an accredited hospice program is appropriate when medical interventions have been exhausted or have been refused by the patient and survival beyond 6 months is unlikely.

REFERENCES

1. Pesce C, Tobis-Gallelli F, Toncini C. AUPD cells of the larynx. Acta Otolaryngol (Stockh) 1984;98: 158–62.
2. Delozier HG. Intrinsic malignant schwannoma of the larynx: a case report. Ann Otol Rhinol Laryngol 1982;91:336.
3. Maisel RH, Ogura JH. Neurofibromatosis with laryngeal involvement. Laryngoscope 1974;89:132–40.
4. Ivatury RR, Shah D, Ascer E, et al. Granular cell tumor of the larynx and bronchus. Ann Thorac Surg

1982;33:69–73.
5. Phillips JH, Rush HO. Angioma of the larynx, especially its relationship to chronic laryngitis. Am J Dis Child 1913;5:123.
6. Myer CM, Bratcher GO. Laryngeal cystic hygroma. Head Neck Surg 1983;6:706–9.
7. Nuutinen J, Syrjanen K. Angioleiomyoma of the larynx. Report of a case and review of the literature. Laryngoscope 1983;93:941–3.
8. American Cancer Society. Cancer facts and figures—2000. Atlanta: American Cancer Society; 2000.
9. DeRienzo DP, Greenberg SD, Fraire AE. Carcinoma of the larynx. Changing incidence in women. Arch Otolaryngol Head Neck Surg 1991;117:681–4.
10. Wynder EL, Covey LS, Mabucki K, Mushinski M. Environmental factors in cancer of the larynx: a second look. Cancer 1976;38:1591–601.
11. Terracini B, Anglesio E, Panero M, Vineis P. Descriptive epidemiology of cancer of the larynx in the province of Torino, Italy. Tumori 1978;64: 445–56.
12. Doll R. The epidemiology of cancer. Cancer 1980;45:2475–85.
13. Wynder EL, Stellman SD. Comparative epidemiology of tobacco-related cancers. Cancer Res 1977; 37:4608–22.
14. Spitz M, Fueger JJ, Goepfert H, Neell GR. Squamous cell carcinoma of the upper aerodigestive tract. A case comparison analysis. Cancer 1988;61:203–8.
15. Ramadan MF, Morton RP, Stell PM, Pharoah PO. Epidemiology of laryngeal cancer. Clin Otolaryngol 1982;7:417–28.
16. El-Serag HB, Hepworth EJ, Lee P, Sonnenberg A. Gastroesophageal reflux disease is a risk factor for laryngeal and pharyngeal cancer. Am J Gastroenterol 2001;96:2013–8.
17. Schantz SP, Hsu TC. Mutagen-induced chromosone fragility within peripheral blood lymphocytes of head and neck cancer patients. Head Neck 1989; 11:337–42.
18. Cloos J, Spitz MR, Schantz SP, et al. Genetic susceptibility to head and neck squamous cell carcinoma. J Natl Cancer Inst 1996;88:530–5.
19. Schantz SP, Hsu TC, Ainslie N, et al. Young adults with head and neck cancer express increased susceptibility to mutagen-induced chromosome damage. JAMA 1989;262:3313–5.
20. Bondy ML, Spitz MR, Halbi S, et al. Association between family history of cancer and mutagen sensitivity in upper aerodigestive tract cancer patients. Cancer Epidemiol Biomarkers Prev 1993;2: 103–6.
21. Trizna Z, Clayman Gl, Spitz MR, et al. Glutathione S-transferase genotypes as risk factors for head and neck cancer. Am J Surg 1995;170:499–501.
22. Jahnke V, Matthias C, Fryer A, et al. Glutathione S-transferase and cytochrome P-450 polymorphism as risk factors for squamous cell carcinoma of the larynx. Am J Surg 1996;172:671–3.
23. Sturgis EM, Castillo EJ, Li L, et al. Polymorphisms of the DNA repair gene XRCC1 in squamous cell carcinoma of the head and neck. Carcinogenesis 1999; 20:2125–9.
24. Brandsma JL, Steinberg BM, Abramson AL, Winkler B. Presence of human papilloma virus-16 related sequences in verrucous carcinoma of the larynx. Cancer Res 1986;46:2185–8.
25. Kashima H, Mounts P, Kuhajda F, Loury M. Demonstration of human papilloma virus capsid antigen in carcinoma in situ of the larynx. Ann Otol Rhinol Laryngol 1986;95:603–7.
26. Austen D. Larynx. In: Schottenfeld P, Fraumeni J, editors. Cancer epidemiology and prevention. Philadelphia: WB Saunders; 1982. p. 554.
27. Abramson AL, Brandsma JL, Steinberg BM, Winkler B. Verrucous carcinoma of the larynx: possible human papillomavirus etiology. Acta Otolaryngol (Stockh) 1985;111:709–15.
28. Loning T, Meichsner M, Milde-Langosch K, et al. HPV DNA detection in tumours of the head and neck: a comparative light microscopy and DNA hybridization study. ORL J Otorhinolaryngol Relat Spec 1987;49:257–69.
29. Koufman JA, Burke AJ. The etiology and pathogenesis of laryngeal carcinoma. Otolaryngol Clin North Am 1997;30:14–9.
30. Nakayama M, Ford CN, Brandenburg JH, Bless DM. Sulcus vocalis in laryngeal cancer: a histopathologic study. Laryngoscope 1994;104:16–24.
31. Christholm M. The association between webs, iron and post cricoid carcinoma. Postgrad Med 1974;50:215.
32. Aanesen JP, Olofsson J. Irradiation-induced tumours of the head and neck. Acta Otolaryngol Suppl (Stockh) 1979;360:178–81.
33. Wagenfeld DJ, Harwood AR, Bryce DP, et al. Second primary respiratory tract malignancies in glottic carcinoma. Cancer 1980;46:1883–6.
34. Moore C. Cigarette smoking and cancer of the mouth, pharynx, and larynx: a continuing study. JAMA 1971;218:553–8.

35. Savary M, Monnier P, Pasche R, et al. Multiple primary malignancies: bearing of basic research on clinical otolaryngology. Adv Otorhinolaryngol 1991;46:165–75.
36. Engelen AM, Stalpers LJ, Manni JJ, et al. Yearly chest radiography in the early detection of lung cancer following laryngeal cancer. Eur Arch Otorhinolaryngol 1992;249:364–9.
37. Sllamniku B, Bauer W, Painter C, Sessions D. The transformation of laryngeal keratosis into invasive carcinoma. Am J Otolaryngol 1989;10:42–54.
38. Hellquist H, Lundgren J, Olofsson J. Hyperplasia, keratosis, dysplasia and carcinoma in situ of the vocal cords: a follow-up study. Clin Otolaryngol 1982;7:11–27.
39. Andrea M, Dias O, Santos A. Contact endoscopy during microlaryngeal surgery: a new technique for endoscopic examination of the larynx. Ann Otol Rhinol Laryngol 1995;104:333–9.
40. Wardrop PJ, Sim S, McLaren K. Contact endoscopy of the larynx: a quantitative study. J Laryngol Otol 2000;114:437–40.
41. Bauer W. Concomitant carcinoma in situ and invasive carcinoma of the larynx. In: Alberti P, Bryce D, editors. Workshops from the Centennial Conference on Laryngeal Cancer. East Norwalk (CT): Appleton-Century-Crofts; 1976. p. 127.
42. Reid A, Robin P, Powell J, et al. Staging carcinoma: its value in cancer of the larynx. J Laryngol Otol 1991;105:456–8.
43. Molinari R. Clinical classification of laryngeal carcinoma. Critique of existing classifications and proposals of a new working classification. Acta Otorhinolaryngol Ital 1990;10:579–91.
44. Hagen P, Lyons GD, Haindel C. Verrucous carcinoma of the larynx: role of human papillomavirus, radiation, and surgery. Laryngoscope 1993;103: 253–7.
45. Ferlito A. Diagnosis and treatment of verrucous squamous cell carcinoma of the larynx: a critical review. Ann Otol Rhinol Laryngol 1985;94:575–9.
46. Wain SL, Kier R, Vollmer RT, Bossen EH. Basaloid-squamous cell carcinoma of tongue, hypopharynx, and larynx: report of 10 cases. Hum Pathol 1986; 17:1158–66.
47. Ereno C, Lopez JI, Sanchez JM, Toledo JD. Basaloid-squamous cell carcinoma of the larynx and hypopharynx. A clinicopathologic study of 7 cases. Pathol Res Pract 1994;190:186–93.
48. Luna MA, el-Naggar A, Parichatikanond P, et al. Basaloid-squamous cell carcinoma of the upper aerodigestive tract. Clinicopathologic and DNA flow cytometric analysis. Cancer 1990;66:537–42.
49. Weidner N. Sarcomatoid carcinoma of the upper aerodigestive tract. Semin Diagn Pathol 1987;4: 157–68.
50. Ellis GL, Langloss JM, Heffner DK, Hyams VJ. Spindle-cell carcinoma of the aerodigestive tract. An immunohistochemical analysis of 21 cases. Am J Surg Pathol 1987;11:335–42.
51. Olofsson J, van Nostrand AW. Adenoid cystic carcinoma of the larynx: a report of four cases and a review of the literature. Cancer 1977;40:1307–13.
52. Batsakis JG, el-Naggar AK, Luna MA. Neuroendocrine tumors of larynx. Ann Otol Rhinol Laryngol 1992;101:710–4.
53. Wenig BM, Gnepp DR. The spectrum of neuroendocrine carcinomas of the larynx. Semin Diagn Pathol 1989;6:329–50.
54. Nicolai P, Ferlito A, Sasaki CT, Kirchner JA. Laryngeal chondrosarcoma: incidence, pathology, biological behavior, and treatment. Ann Otol Rhinol Laryngol 1990;99:515–23.
55. Pressman JJ. Submucosal compartmentation of the larynx. Ann Otol Rhinol Laryngol 1956;65:766–71.
56. Hiranandani L. Panel on epidemiology and etiology of laryngeal carcinoma. Laryngoscope 1987;85: 1197.
57. De Stefani E, Correa P, Oreggia F, et al. Risk factors for laryngeal cancer. Cancer 1987;60:3087–91.
58. Sessions RB, Miller SD, Martin GF, et al. Videolaryngostroboscopic analysis of minimal glottic cancer. Ann Otol Rhinol Laryngol 1989;96:56–9.
59. Zhao R, Hirano M, Tanaka S, Sato K. Vocal fold epithelial hyperplasia: vibratory behavior versus extent of lesion. Arch Otol Head Neck Surg 1991; 117:1015–8.
60. Colden D, Zeitels SM, Hillman RE, et al. Stroboscopic assessment of vocal fold keratosis and glottic cancer. Ann Otol Rhinol Laryngol 2001;11: 293–8.
61. Curtin HD, Ishwaran H, Mancuso A, et al. Comparison of CT and MR imaging in staging of neck metastases. Radiology 1998;207:123–30.
62. Stokkel MP, ten Broek FW, Hordijk GJ, et al. Preoperative evaluation of patients with primary head and neck cancer using dual-head 18fluorodeoxyglucose positron emission tomography. Ann Surg 2000; 231:229–39.
63. Strauss LG. Sensitivity and specificity of positron emission tomography (PET) for the diagnosis of

lymph node metastases. Recent Results Cancer Res 2000;157:21–9.
64. Kau RJ, Alexiou C, Stimmer H, Arnold W. Diagnostic procedures for detection of lymph node metastases in cancer of the larynx. ORL J Otorhinolaryngol Relat Spec 2000;62:199–203.
65. Beyer T, Townsend DW, Brun T, et al. A combined PET/CT scanner for clinical oncology. J Nucl Med 2000;41:1369–79.
66. Jungehulsing M, Scheidhauer K, Damm M, et al. 2[F]-fluoro-2-deoxy-D-glucose positron emission tomography is sensitive tool for the detection of occult primary cancer (carcinoma of unknown primary syndrome) with head and neck lymph node manifestation. Otolaryngol Head Neck Surg 2000;123:294–301.
67. Bohulslavizki KH, Klutmann S, Kroger S, et al. FDG PET detection of unknown primary tumors. J Nucl Med 2000;41:816–22.
68. Zeitels SM. A universal modular glottiscope system: the evolution of a century of design and technique for direct laryngoscopy. Ann Otol Rhinol Laryngol 1999;108 Suppl 179:1–24.
69. Piccirillo JF, Wells CK, Sasaki CT, Feinstein AR. New clinical severity staging system for cancer of the larynx. Five-year survival rates. Ann Otol Rhinol Laryngol 1994;103:83–92.
70. Piccirillo JF. Importance of comorbidity in head and neck cancer. Laryngoscope 2000;110:593–602.
71. Chacko DC, Hendrickson FR, Fisher A. Definitive irradiation of T1-T4N0 larynx cancer. Cancer 1983; 51:994–1000.
72. Jorgensen K, Hjelm-Hansen M, Andersen AP, Lund C. Laryngeal carcinoma. I. Treatment results. Acta Radiol Oncol Radiat Phys Biol 1979; Suppl 18:282–94.
73. Vermund H. Role of radiotherapy in cancer of the larynx as related to TNM systems of staging. A review. Cancer 1970;25:485–504.
74. Dickens WJ, Cassisi NJ, Million RR, Bova FJ. Treatment of early vocal cord carcinoma: a comparison of apples and apples. Laryngoscope 1983; 93:216–9.
75. Wang CC. Treatment of glottic carcinoma by megavoltage radiation therapy and results. Am J Roentgenol Ther Radium Nucl Med 1974;120:157–63.
76. Cummings C. Malignant glottic carcinomas. In: Cummings C, Sessions D, Weymuller E, Wood P, editors. Atlas of laryngeal surgery St. Louis: CV Mosby; 1984. p. 209.
77. Strong MS, Jako GJ. Laser surgery in the larynx. Ann Otol Rhinol Laryngol 1972;81:791–8.
78. Eckel HE, Thumfart WF. Laser surgery for the treatment of larynx carcinomas: indications, techniques, and preliminary results. Ann Otol Rhinol Laryngol 1992;101:113–8.
79. Leirens J, Vidts G, Schmelzer B, et al. Premalignant lesions of the vocal cords. A retrospective study of 62 cases treated with CO_2 laser. Acta Otolaryngol (Stockh) 1997;117:903–8.
80. Lindholm CE, Elner A. Transoral laser surgery of laryngeal carcinomas. State of the art in Sweden. Adv Otorhinolaryngol 1995;49:250–3.
81. Mahieu HF, Patel P, Annyas AA, van der Laan T. Carbon dioxide laser vaporization in early glottic carcinoma. Arch Otolaryngol Head Neck Surg 1994; 120:383–7.
82. Rudert HH, Werner JA. Endoscopic resections of glottic and supraglottic carcinomas with the CO_2 laser. Eur Arch Otorhinolaryngol 1995;252:146–8.
83. Steiner W. Results of curative laser microsurgery of laryngeal carcinoma. Am J Otolaryngol 1993;14: 116–21.
84. Kashimi HK, Lee DJ, Zinreich SJ. Vestibulectomy in early vocal fold carcinoma. In: Johnson JT, Didolkar MS, editors. Head and neck cancer. Vol III. Amsterdam: Elsevier Science BV; 1993. p. 181–189.
85. Myers EN, Wagner RL, Johnson JT. Microlaryngoscopic surgery for T1 glottic lesions: a cost-effective option. Ann Otol Rhinol Laryngol 1992;103:236–41.
86. Lederman M. Radiotherapy of cancer of the larynx. J Laryngol Otol 1970;84:867–96.
87. Fletcher GH, Lindberg RD, Hamberger A, Horiot JC. Reasons for irradiation failure in squamous cell carcinoma of the larynx. Laryngoscope 1975;85:987–1003.
88. Haughey BH, Fredrickson JM, Sessions DG, Fuller D. Vibratory segment function after free flap reconstruction of the pharyngoesophagus. Laryngoscope 1995;105:487–90.
89. Ambrosch P, Kron M, Steiner W. Carbon dioxide laser microsurgery for early supraglottic carcinoma. Ann Otol Rhinol Laryngol 1998;107:680–8.
90. Rudert HH, Werner JA, Hoft ST. Transoral carbon dioxide laser resection of supraglottic carcinoma. Ann Otol Rhinol Laryngol 1999;108:819–27.
91. Eckel HE. Endoscopic laser resection of supraglottic carcinoma. Otolaryngol Head Neck Surg 1998; 117:681–7.
92. Davis RK, Hayes JK. Management of supraglottic cancer: selected endoscopic laser resection and postoperative irradiation. Adv Otorhinolaryngol 1995; 49:231–6.

93. Zeitels SM, Davis RK. Endoscopic laser management of supraglottic cancer. Am J Otolaryngol 1995;16:2–11.
94. Rudert HH. Laser surgery in the management of carcinoma of the supraglottic larynx: evolution, current concepts, and future trends. Laryngoscope 1996;106:559–67.
95. Kambic V, Radsel Z, Smid L. Laryngeal reconstruction with epiglottis after vertical hemilaryngectomy. J Laryngol Otol 1976;90:467–73.
96. Laccourreye O, Brasnu D, Bassot V, et al. Cisplatin-fluorouracil exclusive chemotherapy for T1-T3N0 glottic squamous cell carcinoma complete clinical responders: five-year results. J Clin Oncol 1996; 14:2331–6.
97. Pearson BW. Subtotal laryngectomy. Laryngoscope 1981;91:1904–12.
98. Piquet JJ, Chevalier D. Subtotal laryngectomy with crico-hyoido-epiglotto-pexy for the treatment of extended glottic carcinomas. Am J Surg 1991; 162:357–61.
99. Piquet JJ, Darras J, Berrier A, et al. Les laryngectomies subtotales fonctionnelles avec cricohyoidopexie: technique, indiciation, resultats. Ann Otolaryngol Chir Cervicofac 1986;103: 411–5.
100. Weinstein GS, Laccourreye O. Supracricoid laryngectomy with cricohyoidoepiglottopexy. Otolaryngol Head Neck Surg 1994;111:684–5.
101. Coates HL, DeSanto LW, Devine KD, Eiveback LR. Carcinoma of the supraglottic larynx. A review of 221 cases. Arch Otolaryngol 1976;102:686–9.
102. Marks JE, Freeman RB, Lee F, Ogua JH. Carcinoma of the supraglottic larynx. Am J Surg 1979;156:301.
103. Maceri DR, Laupe HB, Makielski KH, et al. Conservation laryngeal surgery. A critical analysis. Arch Otolaryngol 1985;111:361–5.
104. DeSanto LW. Cancer of the supraglottic larynx: a review of 260 patients. Otolaryngol Head Neck Surg 1977;93:705.
105. Mendenhall WM, Parsons JT, Mancuso AA, et al. Radiotherapy for squamous cell carcinoma of the supraglottic larynx: an alternative to surgery. Head Neck 1996;18:24–35.
106. DeSanto LW. T3 glottic cancer: options and consequences of the options. Laryngoscope 1984;94: 1311–5.
107. Soo KC, Shah JP, Gopinath KS, et al. Analysis of prognostic variables and results after supraglottic partial laryngectomy. Am J Surg 1988;156:301–5.
108. Lee NK, Goepfert H, Awendt CD. Supraglottic laryngectomy for intermediate-stage cancer: U.T.M.D. Anderson Cancer Center experience with combined therapy. Laryngoscope 1990;100:831–6.
109. Bocca E, Pignataro O, Oldini C. Supraglottic laryngectomy: 30 years of experience. Ann Otol Rhinol Laryngol 1983;92:14–8.
110. Strijbos M, Van Den Broek P, Manni JJ, Huygen PL. Supraglottic laryngectomy: short- and long-term functional results. Clin Otolaryngol 1987;12:265–70.
111. Shah JP, Toliefsen HR. Epidermoid carcinoma of the supraglottic larynx. Am J Surg 1974;128:494–9.
112. Wiernik G, Bates TD, Bleehen NM, et al. Final report of the general clinical results of the British Institute of Radiology fractionation study of the 3F/wk versus 5F/wk in radiotherapy of carcinoma of the larygo-pharynx. Br J Radiol 1990;63:169–80.
113. Cummings B, O'Sullivan B, Keane T, et al. Preservation of the larynx in a prospective trial of hyperfractionated (HFRT) versus conventional radiation (CRT) in locally advanced larynx cancer. Clin Invest Med 1997;20 Suppl 4:S88.
114. Skolnik EM, Yee KF, Wheatley MA, Martin LO. Carcinoma of the laryngeal glottis therapy and end results. Laryngoscope 1975;85:1453–66.
115. Yuen A, Medina JE, Goepfert H, Fletcher G. Management of stage T3 and T4 glottic carcinomas. Am J Surg 1984;148:467–72.
116. Horiot JC, Bontemps P, Van Den Bogaert W, et al. Accelerated fractionation (AF) compared to conventional fractionation (CF) improves locoregional control in radiotherapy of advanced head and neck cancers: results of the EORTC 22851 randomized trial. Radiother Oncol 1997;44:111–21.
117. Jackson SM, Weir LM, Hay JH, et al. A randomized trial of accelerated versus conventional radiotherapy in head and neck cancer. Radiother Oncol 1997;43:39–46.
118. Dische S, Saunders M, Barrett A, et al. A randomized multicenter trial of CHART versus conventional radiotherapy in head and neck cancer. Radiother Oncol 1997;44:123–36.
119. Dische S, Saunders MI. Randomized multicenter trial of CHART versus conventional radiotherapy in head and neck cancer—interim results and implications [abstract]. Int J Radiat Oncol Biol Phys 1996;36:234.
120. Fu KK, Pajak TF, Trotti A, et al. A Radiation Therapy Oncology Group (RTOG) phase III randomized study to compare hyperfractionation and two

variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: preliminary results of RTOG 9003 [abstract]. Int J Radiat Oncol Biol Phys 1999;45:145.
121. Stuschke M, Thames HD. Hyperfractionated radiotherapy of human tumors: overview of the randomized clinical trails. Int J Radiat Oncol Biol Phys 1997;37:259–67.
122. Department of Veterans Affairs Laryngeal Cancer Center Study Group. Induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. N Engl J Med 1991;324:1685–90.
123. Pignon JP, Bourhis J, Domenge C, Designe L, on behalf of the MACH-NC Collaborative Group. Chemotherapy added to locoregional treatment for head and neck squamous cell carcinoma: three meta-analyses of updated individual data. Lancet 2000;355:949–55.
124. NCCN practice guidelines for head and neck cancer. Oncology 1998;12:39–147.
125. Mendenhall WM. T3-4 squamous cell carcinoma of the larynx treated with radiation therapy alone. Semin Radiat Oncol 1998;8:262–9.
126. Tannock IF, Cummings BJ. Neoadjuvant chemotherapy in head and neck cancer: no way to preserve a larynx. J Clin Oncol 1992;10:343–5.
127. Chevalier D, Laccourreye O, Brasnu D, et al. Cricohyoidoepiglottopexy for glottic carcinoma with fixation or impaired motion of the true vocal cord: 5-year oncologic results with 112 patients. Ann Otol Rhinol Laryngol 1997;106:364–9.
128. Laccourreye O, Brasnu D, Biacabe B, et al. Neoadjuvant chemotherapy and supracricoid partial laryngectomy with cricohyoidopexy for advanced endolaryngeal carcinoma classified as T3-T4: 5-year oncologic results. Head Neck 1998;20:595–9.
129. Brasnu D, Laccourreye O, Chabardes E, et al. Conservative surgery of the larynx as an alternative to total laryngectomy after neoadjuvant chemotherapy. Initial results. Ann Otolaryngol Chir Cervicofac 1993;110:129–33.
130. Lecanu JB, Monceaux G, Perie S, et al. Conservative surgery in T3-4 pharyngolaryngeal squamous cell carcinoma: an alternative to radiation therapy and to total laryngectomy for good responders to induction chemotherapy. Laryngoscope 2000;110:412–6.
131. McGuirt WF, Blalock D, Koufman JA, Feehs RX. Voice analysis of patients with endoscopically treated early laryngeal carcinoma. Ann Otol Rhinol Laryngol 1992;101:(2 Pt 1):142–6.
132. McGuirt WF, Blalock D, Koufman JA, et al. Comparative voice results after laser resection or irradiation of T1 vocal cord carcinoma. Arch Otolaryngol Head Neck Surg 1994;120:951–5.
133. Casiano RR, Cooper JD, Lundy DS, Chandler JR. Laser cordectomy for T1 glottic carcinoma: a 10-year experience and videostroboscopic findings. Otolaryngol Head Neck Surg 1991;104:831–7.
134. Leeper HA, Heeneman H, Reynolds C. Vocal function following vertical hemilaryngectomy: a preliminary investigation. J Otolaryngol 1990;19:62–7.
135. Keith RL, Leeper HA, Doyle PC. Microanalytic acoustical voice characteristics of near-total laryngectomy. Otolaryngol Head Neck Surg 1995;113:689–94.
136. Singer MI, Blom ED. Selective myotomy for voice restoration after total laryngectomy. Arch Otolaryngol 1981;107:670–3.
137. Miralles JL, Cervera T. Voice intelligibility in patients who have undergone laryngectomies. J Speech Hear Res 1995;38:564–71.
138. St Guily JL, Angelard B, El-Bez M, et al. Post laryngectomy voice restoration. A prospective study of 83 patients. Arch Otolaryngol Head Neck Surg 1992;118:252–5.
139. Robbins J. Acoustic differentiation of laryngeal, esophageal, and tracheoesophageal speech. J Speech Hear Res 1984;27:577–85.
140. Mendelsohn M, Morris M, Gallagher R. A comparative study of speech after total laryngectomy and total laryngopharyngectomy. Arch Otolaryngol Head Neck Surg 1993;119:508–10.
141. Henley J, Souliere C. Tracheoesophageal speech failure in the laryngectomee: the role of constrictor myotomy. Laryngoscope 1986;96:1016–20.
142. Clevens RA, Esclamado RM, Hartshorn DO, Lewin JS. Voice rehabilitation after total laryngectomy and tracheoesophageal puncture using non-muscle closure. Ann Otol Rhinol Laryngol 1993;102:792–6.
143. Weber RS, Marvel J, Smith P, et al. Paratracheal lymph node dissection for carcinoma of the larynx, hypopharynx, and cervical esophagus. Otolaryngol Head Neck Surg 1993;108:11–7.
144. Sisson GA. 1989 Ogura memorial lecture: mediastinal dissection. Laryngoscope 1989;99:1262–6.

CHAPTER 55

Cancer Immunology

Scott E. Strome, MD, Ray O. Gustafson, MD, H. Bryan Neel III, MD, PhD

At the beginning of the twentieth century, microbiologists established the foundation for the science of immunology and described the process of host resistance to infectious disease. The field of tumor immunology grew out of speculation that immune mechanisms might be involved in host defense against malignancy.

In the laboratory, when biologists harvested tumor tissue from one animal and attempted to propagate it in a second animal, the tumor transplant was rejected quickly. This phenomenon stimulated intense interest and investigation, but the nascent field of tumor immunology was arrested in the 1930s by the demonstration that rejection of transplanted tumors was simply a manifestation of the general phenomenon of rejection of foreign tissues. Although this observation led to the study of major histocompatibility complexes (MHCs), formed the basis for organ transplantation, and resulted in the discovery of cell-mediated cytotoxicity, it impeded the search for an immune basis of tumor resistance for many years.

Renewed interest in the field of tumor immunology in the 1940s and 1950s resulted from the work of Gross,[1] Foley,[2] Prehn and Main,[3] and Klein et al[4] after the development of highly inbred (syngeneic) strains of mice. In experiments using the chemical carcinogen methylcholanthrene to induce murine sarcomas, these investigators conclusively demonstrated the ability of tumors to elicit a tumor-specific immune response.

As outlined in a review by Ritts and Neel,[5] resistance to subsequent transplantation of grafts of the same tumor could be induced by tumor excision or amputation, use of subtake tumor dose or cells that had been damaged irreversibly in their reproductive capacity by irradiation or chemicals, or necrosis in situ by ligation release,[6,7] cryosurgery,[8–11] or electrocoagulation. In a series of classic experiments, mice from which tumors had been excised showed a reduced rate of tumor growth at the challenge site compared with that of controls, indicating existence of immunity after excision. Mice from which tumors had been eradicated in situ by cryosurgery, ligation release, or electrocoagulation showed a reduced rate of tumor take and growth at the challenge site compared with controls and with mice subjected to tumor excision, indicating even greater immunity after necrosis of tumor in situ. The pattern of differential immunity is seen consistently after challenge with graded doses of tumor cells and in other tumor-host systems.[12]

All of these experiments strongly suggested that tumor cells must express unique surface determinants that can be recognized by the immune system as nonself, and these determinants were termed "tumor-specific antigens." Tumor immunologists subsequently have been concerned with the identification and characterization of tumor-associated antigens, processing and presentation of tumor-associated antigens, subsequent humoral and cellular immune responses, and manipulation of the immune system in the diagnosis of and therapy for malignant disease. A thorough understanding of the basic science of immunology is a prerequisite for the study of cancer immunology (see Chapter 28.)

SIGNIFICANCE OF THE IMMUNE RESPONSE

Although tumor immunology is a young field, it has stimulated a great deal of research; however, the actual clinical significance of the immune response to malignant disease remains unclear. In 1959, Thomas suggested that the host possessed a normal homeostatic cellular mechanism that was active in preventing establishment of "nonself,"[13] and Burnet later termed this homeostatic mechanism "immunosurveillance."[14]

Among the more impressive indirect data establishing the phenomenon of surveillance in cancer are the observations that the frequency of neoplasia ranges from 10 to 15% in patients who have cellular immunodeficiency diseases—10,000 times that of an actuarially matched population[15]—and is 0.7% in patients who have received selective immunosuppressive therapy, primarily at the cellular level, for kidney transplantation. The neoplasia observed in children with T-cell deficiency syndromes is most often (more than 60%) of the reticuloendothelial system. Spontaneous, complete regression of established tumors is uncommon, but cases are well documented. On the other hand, most physicians have observed patients who were apparently "cured" for many years only to discover metastases as many as 20 years after ablation of the primary neoplasm. Lymphocyte infiltration of solid tumors, including squamous cell carcinoma of the head and neck, is associated with a better prognosis.[16] Surgeons, particularly those working with diseases of the head and neck, know that identification of viable tumor cells in wound washings or in circulation does not reliably predict local recurrence or survival.[17] Finally, immune mechanisms are weak at the extremes of age, early childhood and late adulthood, and these are the times when the incidence of malignancy is greatest.

According to the theory of immunosurveillance, response against malignancy in the host has been attributed primarily to mature, specifically immune T cells. Accordingly, it was anticipated that T-cell–deficient laboratory animals, such as the congenitally athymic nude mouse, would lack resistance against malignancy and would be subject to a significantly higher incidence and more rapid progression of malignancy than its immunocompetent counterpart. Several reports have taken the failure to demonstrate such a significantly higher incidence of malignant change in the nude mouse as convincing evidence against the theory of immunosurveillance.[18–21] Studies of cell-mediated immunity, however, have provided additional information that diminishes the controversy surrounding immunosurveillance.

First, the nude mouse possesses high natural killer (NK) reactivity, which offers a potential explanation for its low incidence of spontaneous malignancy and its resistance to growth of malignant tissue transplants.[22] Second, Talmadge and associates have shown that malignant cells that are highly susceptible to natural killing in vitro grow poorly in normal mice but disseminate rapidly in the NK-deficient beige mouse.[23] In humans, depressed to absent NK reactivity has been found in patients with Chédiak-Higashi syndrome and a hereditary form of malignant melanoma.[24] Furthermore, Takasugi et al[25] and Pross and Baines[26] compared in vitro NK reactivity of a large number of cancer patients with various controls and demonstrated a correlation between increased tumor burden and depressed NK reactivity. Thus, there are several lines of evidence to associate the role of immunosurveillance with the NK phenomenon. Although NK activity may not be a critical factor in host defense against established tumors, evidence is accumulating to support the role of the NK cell in the control of metastatic disease.[27]

Lymphocytic blastogenic transformation induced by phytohemagglutinin and by specific antigens, as an in vitro correlate of delayed cutaneous hypersensitivity, is impaired in some patients with cancer. Because the end point of such studies is unrelated to the inflammatory response, it is concluded that there is a cellular immune defect or diminution in cell-mediated immunity. At least on an operational level, the failure to manifest cellular immunity in cancer is caused by a deficit or a malfunction, or both, of T lymphocytes.

Immunologic testing, performed at diagnosis and clinical staging and before treatment has been instituted, can provide information that predicts or assists in predicting patients in whom cancer is most likely to recur. Testing could also serve as a basis on which adjuvant forms of therapy can be instituted. In a prospective study of squamous cell carcinomas of the head and neck,[28] patients were categorized into two groups: those who remained tumor free and those who had cancer recurrence. Pretreatment phytohemagglutinin and concanavalin A–induced blastogenesis were more often depressed in patients in whom clinically apparent recurrence eventually developed than in patients without recurrence. This finding is consistent with the hypothesis that cellular immune function, as measured in the laboratory by nonspecific blastogenesis, can identify, within the staging system, subpopulations of patients who are less likely to survive after conventional therapy (Table 55–1). In this study, T-lymphocyte and B-lymphocyte enumerations had no predictive value. However, some studies have demonstrated consistently reduced T-cell numbers and activity in patients with advanced head and neck cancers.[29,30]

TABLE 55–1. Pretreatment Test Results (Mean) by Stage of Disease in Patients Who Remained Tumor Free and in Patients in Whom Cancer Recurred

	Tumor Free			*Cancer Recurrence*		
Variable	*Stage I*	*Stage II*	*Stage III*	*Stage I*	*Stage II*	*Stage III*
T cells						
N	2,213	1,537	2,342	2,876	1,881	1,824
%	79	75	74	65	77	71
B cells						
N	458	413	537	720	448	592
%	18	20	20	16	17	25
Blastogenesis (cpm)						
Phytohemagglutinin	44,000	34,200	51,250	27,500	22,500	10,400*
Concanavalin A	38,666	36,500	28,500	14,000*	24,500	15,000*
Pokeweed mitogen	12,000	22,300	25,500	24,000	13,250	14,200

*Significantly different from respective tumor-free stage value ($p < .05$).
Reproduced with permission of the American Academy of Otolaryngology from Ryan RE Jr et al.[28]

Other studies reported increased levels of suppressor T cells in the peripheral blood and regional lymph nodes of patients with head and neck cancer,[31–33] whereas another study reported increased ratios of helper T cells to suppressor T cells in patients with head and neck cancer.[34] These findings suggest that complex changes in cellular immune regulation may be more important than total T-cell number in explaining the failure of immunosurveillance and that ineffective cytotoxic responses may be caused by inappropriate stimulation of suppressor cell activities. In a detailed series of experiments, Wolf et al have shown that (1) T lymphocytes are the major lymphocyte component infiltrating head and neck carcinomas; (2) suppressor T cells are the dominant cell type in tumor parenchyma, but helper and suppressor T cells are equally represented in tumor stroma; and (3) extent of helper cell parenchymal infiltration correlates directly with clinical outcome in patients who have head and neck cancer.[35] In a subsequent flow cytometry study of T-cell subsets in patients with previously untreated head and neck cancers, this group of investigators provided firm evidence that decreased proportions of cytotoxic or suppressor cells are associated frequently with advanced head and neck cancers and portend a poor prognosis.[36]

The humoral immune system may interfere with the host response to tumor, contributing to a failure of immunosurveillance. Elevated titers of immunoglobulin A (IgA) have been demonstrated in the sera of patients with head and neck cancer,[37,38] and it has been proposed that the IgA functions as a blocking factor of lymphocyte-mediated cytotoxicity, at least in unique forms of nasopharyngeal carcinoma (NPC).[39,40] Furthermore, Seder et al reported successful palliative treatment of advanced head and neck cancers by plasmapheresis and hypothesized mediation of the beneficial response by a reduction of circulating immunosuppressive factors.[41]

CLINICAL APPLICATIONS OF TUMOR IMMUNOLOGY

The orchestrated activity of the immune system is capable of providing a natural defense against malignant disease, and a significant aspect of tumor immunology is concerned with manipulation of the immune system in the diagnosis and treatment of cancers that evade surveillance. Current emphasis is on tumor markers, monoclonal antibodies, and cell-based strategies for mediating the regression of established disease.

Tumor Markers

Oncogenesis is associated frequently with detectable changes in the serum content of various substances, which are referred to as tumor markers.[42] Many different types of tumor markers have been described. One scheme[43] groups them as (1) markers indicat-

ing a predisposition to malignant changes, (2) markers indicating inflammation or tumor necrosis, (3) markers relevant to tumor metabolism, (4) hormone receptors,[44] and (5) oncogenic virus antigens. The three basic uses for tumor markers include screening of high-risk populations, distinguishing among malignant tumors or between benign and malignant processes, and following the course of a tumor over time.[45]

Although there is growing evidence for the existence of tumor-specific antigens in squamous carcinomas of the head and neck,[18] two of the most widely used markers in the practice of head and neck oncology are calcitonin and Epstein-Barr virus (EBV)-associated antigens. Jackson et al demonstrated the value of serum calcitonin levels as a screening test for family members of patients with medullary carcinoma of the thyroid and for post-treatment surveillance for recurrent disease.[46]

A wealth of biologic, biochemical, and immunologic evidence supports an association between EBV and NPC. Immunologic studies have identified antibodies against certain EBV-associated antigens and support the concept that these markers are of value in the diagnosis and management of patients with NPC.[47–50] The most specific tests are those that measure the IgA antibody response to EBV-induced viral capsid antigen and the IgG antibody response to early antigen; the viral capsid antigen (IgA) is the more specific of the two tests.[49] The viral capsid antigen (IgA) and early antigen tests complement the process of diagnosis of NPC and are of special value in directing attention to the nasopharynx in patients with occult or small NPCs.[50–52]

The antibody-dependent cellular cytotoxicity assay titrates sera for antibody to the EBV-induced membrane antigen complex and predicts the clinical course and prognosis of patients with World Health Organization (WHO) types 2 and 3 NPC.[53] High titers at diagnosis signify a better prognosis. It is a significant risk factor and has been incorporated into a new staging (scoring) system for NPC.[54]

A striking relationship exists between the antibody titers and the histopathologic type of NPC. In a prospective collaborative study of NPC in North American patients,[49,51,55] antibody titers were consistently increased in cases of WHO types 2 and 3, averaging 85% positive. In cases of WHO type 1, however, the incidence of positive test results was similar to that for control groups.[49] Further refinement of these tests may lead to their use in screening programs in areas of the world with a high incidence of NPC.

The striking correlation between EBV and certain types of NPC led investigators to seek other viral associations in patients with head and neck malignancies. The recognized relationship between human papillomavirus (HPV) and cervical carcinoma has resulted in efforts to define the presence of HPV in malignant and premalignant lesions of the head and neck.[56] Recent studies of premalignant lesions have focused primarily on laryngeal papillomas and inverted papillomas of the nasal cavity. Garcia-Milian et al analyzed the presence of HPV in normal laryngeal mucosa, laryngeal carcinoma, and laryngeal papilloma by using polymerase chain reaction.[57] This study demonstrated an 82% and 48.5% association between the presence of HPV and laryngeal papillomas and carcinomas, respectively, significantly different from the 16% detection of HPV in normal controls.[57] In addition to the association of HPV with respiratory tract papillomas, there are several anecdotal reports of squamous cell carcinoma developing in the presence of respiratory papillomatosis without coexisting risk factors.[58,59]

Similar studies of the relationship between malignant degeneration of inverted papillomas and the presence of specific HPV subtypes are inconclusive. A study by Beck et al evaluating archival specimens of inverted papilloma demonstrated a significant correlation between the degree of dysplasia and specific HPV subtypes.[60] Other studies have failed to show an association between HPV and the presence of inverted papilloma or in the progression of this disease toward malignancy.[61,62] Resolution of this controversy will most likely require a prospective study using fresh tissue to avoid the limitations of DNA extraction from paraffin.

Interestingly, recent studies from our laboratory at the Mayo Clinic (A Savva, AE Brissett, SE Strome, B Gastout, JC Lewis, R McGovern, AM Weaver, D Persing, JL Kasperbauer, unpublished data, 2000; AE Brissett, A Savva, SE Strome, BS Gastout, JE Lewis, JL Weaver, RM McGovern, JL Kasperbauer, unpublished data, 2000) evaluating malignant and normal tonsillar tissue from age-matched controls demonstrated a significant increase of specific HPV subtypes, with a statistically significant improvement in disease-specific survival in patients with HPV. Several mechanisms have been postulated to characterize the potential pathophysi-

ologic relationship between the presence of HPV and the development of malignancy, including the ability of the HPV E6 protein to bind p53 and indirectly increase its rate of degradation.[63] A direct causal relationship between the presence of HPV and the development of squamous cell carcinoma has not been demonstrated conclusively.

Monoclonal Antibodies Antigen-specific monoclonal antibodies are products of hybridomas; they typically result from the fusion of immune-stimulated murine spleen cells and an appropriate B-cell tumor. Although this technology was first reported as recently as 1975,[64] it has generated an explosion of research on the diagnostic and therapeutic applications of monoclonal antibodies.[65]

In the diagnosis of malignant disease, monoclonal antibodies have been labeled with a radionuclide, injected into the patient's bloodstream, and then visualized with a radiation scanner, radioimmunodetection (RAID). The potential clinical applications for RAID in malignant disease include presurgical staging of the extent of the disease, postsurgical evaluation of residual disease, confirmation of viable tumor identified by other methods, identification of disease recurrence in patients with increasing serum tumor-marker titers, and confirmation of tumor targeting by antibody for immunotherapy. The best current indications for RAID are the detection of occult tumors and the confirmation of tumor sites revealed by other radiologic methods. Most studies of RAID have used ovarian and colorectal cancers, and with several thousand patients studied, between 60 and 90% of known lesions have been identified correctly. With the use of pancarcinoma antibodies and monoclonal agents with greater specificity, the tumor indications for RAID are expanding.

Several strategies have been followed in the therapeutic application of monoclonal antibodies. One strategy has involved administration of unconjugated, naked monoclonal antibody to produce complement-dependent or antibody-dependent cell-mediated cytotoxicity, to interfere directly with the growth or differentiation of tumor cells by binding to them, and to enhance the activities of other agents such as interferon and interleukin-2.

Another monoclonal antibody therapeutic strategy has involved the coupling of conventional cytotoxic agents to a tumor-seeking monoclonal antibody, with the expectations of decreased host toxicity and increased antitumor efficacy. Similarly, the conjugation of biologic toxins with carrier antibodies to produce potentially therapeutic immunotoxins has been explored. Despite their theoretic appeal, both of these approaches have been hampered by problems of conjugating the agent to a monoclonal antibody and the production of human antibodies against the murine monoclonal antibody and agent. Finally, with advances in the chemistry of antibody radiolabeling, many radionuclides are being investigated for radioimmunotherapy, with the prospect of sparing normal tissues from intensive radiation. Despite its theoretic appeal, radioimmunotherapy presents formidable obstacles, including bone marrow toxicity, a low dose of radiation delivered to the tumor, and limited use of multiple treatment courses because of mouse monoclonal antibody antigenicity. Monoclonal antibodies are not yet "magic bullets," but they do hold exciting prospects for enhanced oncologic selectivity and efficacy. In head and neck oncology, monoclonal antibodies have been used primarily in the laboratory for characterization of tumor-associated antigens and study of lymphocyte subpopulations.[18,35]

ACTIVE AND PASSIVE IMMUNOTHERAPY

Cellular Immunity to Malignant Tumors

Fundamental to the understanding of the role of cellular therapy in tumor immunology is an appreciation of the methods of antigen presentation and the role of the cellular immune response in mediating tumor growth in vivo. Both $CD8^+$ cytotoxic T lymphocytes and $CD4^+$ T cells are considered important effector arms of the cellular immune response to malignancy.[66] The initial signal for T-cell activation depends on antigen presentation in the context of appropriate MHC molecules. Subsequent nonspecific secondary signaling of the antigen-primed T cell by a group of glycoproteins, termed "costimulatory molecules," dictates antitumor reactivity.

Costimulatory molecules play diverse roles in regulating the antitumor T-cell response. Certain costimulatory molecules mediate T-cell proliferation and inhibit activation-induced cell death.[67] This is best evidenced by the interaction of B7-1(CD80)/B7-2(CD86) ligands on antigen-presenting cells with the CD28 receptor molecule constitutively expressed on

T cells.[68] Recently, a new group of costimulatory molecules has been characterized as belonging to the tumor necrosis factor family by virtue of primary, secondary, and tertiary structural homology. Certain members of this family, such as 4-1BB, are inducibly expressed on activated T cells and act in concert with B7/CD28 to enhance T-cell proliferation and antitumor reactivity.[69] Other tumor necrosis factor family members, such as LIGHT, act in a CD28-independent fashion to stimulate T-cell proliferation through the herpes virus entry mediator and lymphotoxin beta-receptor.[70]

In contrast to their stimulatory function, specific receptor–ligand interactions from both the tumor necrosis factor and B7 family are recognized to result in T-cell apoptosis. This is best evidenced by the Fas/Fas ligand interaction.[71] Interestingly, specific molecules have different costimulatory function, dependent on the status of T-cell activation. For example, B7-H1, a recently cloned member of the B7 family, is thought to provide stimulatory activity to naive T cells but to induce apoptosis in activated cells.[72] This suggests that costimulatory molecules play an essential role in maintaining the delicate balance between T-cell protective function and autoimmunity.

Recently, a specialized class of professional antigen-presenting cells (APCs), termed "dendritic cells" (DCs), has been recognized as essential for antigen processing and presentation in the presence of appropriate costimulation. Dendritic cells exist in two forms: an immature form in which antigen uptake is high and a mature form that expresses high levels of class I and II MHC molecules with a broad spectrum of constitutive and inducible costimulatory motifs. Dendritic cells have a migratory function and are detected in lymphoid and nonlymphoid organs. Perhaps most importantly, DCs with prior exposure to antigen can effectively educate naive APCs, resulting in priming of the immune response postulated to overcome effectively tumor-induced tolerance.[73,74]

Mechanisms of Tumor Escape

An understanding of the tumor's ability to evade the host immune response is premised on a working knowledge of two fundamental concepts: immune ignorance and immune tolerance.[75] A simple analogy for immune ignorance is lack of vision, whereas tolerance is characterized by the inability to respond to what is seen. Tumors have developed several mechanisms to prevent immune recognition, including down-regulation of major histocompatibility and transport-associated protein molecules necessary for antigen presentation. Additionally, in cancer patients, T cells are tolerized, or improperly activated, which might be caused by down-regulation of the zeta chain of the T-cell receptor and P56LCK signaling molecule.[76] Thus, at the most basic level, immunologic strategies for the treatment of malignancy can be conceptualized as breaking ignorance and overcoming tolerance. Because both ignorance and tolerance are important physiologically for the maintenance of immunologic homeostasis and the prevention of autoimmunity, any therapeutic strategy that alters this intricate balance has the potential to elicit an autoimmune response.

Cellular Strategies for the Treatment of Malignancy

Cellular strategies for cancer therapy can be broadly divided into two separate categories: active and passive. Active strategies are vaccines, which stimulate the host's cellular and humoral effector arms and ultimately depend on intact host immunity. Passive immunotherapy is defined as the transfer of tumor-specific effector cells to mediate the regression of established tumor. Because passive immunotherapy, also termed "adoptive immunotherapy," relies on effector cell transfer, these approaches may be therapeutic in the absence of functional host antitumor immune response.

Recent advances in the active immunotherapy of cancer have relied primarily on the use of DC vaccines. These DC approaches are premised on the knowledge that antigen presentation in the absence of appropriate costimulation results in activation-induced cell death.[77] When tumor cells serve as APCs, they may effectively induce tolerance by providing initial T-cell–specific antigenicity without costimulation. By presenting tumor antigens in the context of requisite costimulatory motifs, DCs loaded with tumor-associated antigens and tumor-specific antigens might overcome tolerance.[78]

Several methods of antigen loading of DCs are undergoing evaluation to determine therapeutic efficacy and clinical practicality. The primary approaches include the use of tumor lysates, tumor-specific peptides, tumor-specific ribonucleic acid (RNA), and

DC–tumor cell fusion.[78–84] Two recent studies documented the potential therapeutic efficacy of these strategies for the treatment of solid malignancies. Nestle et al, using DCs loaded with tumor lysates from melanoma patients, demonstrated a clinically effective antitumor immune response in 5 of 16 patients.[74] More recently, Kugler et al used allogeneic DCs fused with autologous tumor for the treatment of patients with metastatic renal cell carcinoma.[84] This strategy relied on autologous tumor cell antigen presentation, with allogeneic DCs providing appropriate costimulation and a nonspecific alloimmune response. In this trial, 17 patients had a 41% overall response rate at a mean follow-up of 13 months. Despite these initial encouraging results, several basic issues, including methods of DC generation and cytokine priming, utility of specific DC subsets (DC1 versus DC2) for initiating an immune response, and timing of antigen loading, require further characterization before broad clinical application.[85]

Unlike vaccine strategies, passive immunotherapy approaches may be successful in treating patients with large tumor burdens who have a poorly functional immune effector arm. In fact, adoptive immunotherapeutic approaches have demonstrated the greatest overall therapeutic efficacy of any immunologic treatment strategy.[86] Recent studies by Plautz et al documented the utility of tumor-specific activated effector T cells in mediating the regression of established tumor in human tumor models.[87,88] This improved therapeutic efficacy has resulted from subtle changes in lymph node priming and culture technique, including vaccine priming of lymph nodes with tumor in the presence of granuloma granulocyte-macrophage colony-stimulating factor, superantigen stimulation of pre-effector T cells in vitro, and shorter culture duration.[87,88]

Further improvement of these passive immunotherapy strategies will most likely rely on the identification and characterization of subpopulations of tumor-specific effector T cells within tumor-draining lymph nodes. One such example is the identification of a population of lymph node cells characterized by low levels of expression of the L-selectin molecule, which have been demonstrated to mediate all of the antitumor immune reactivity in mouse models.[89] Translation of these data in human squamous cell carcinoma models has demonstrated phenotypic and functional differences of L-selectin cells in vitro, which will improve our understanding of antitumor T-cell immunity and might further existing passive immunotherapy strategies for the treatment of solid malignancies.[66]

Monitoring of the Immune Response in Immunotherapy Trials

Further understanding and development of immunotherapy for the treatment of solid tumors will rely on the ability to create effective immunologic monitoring strategies. Specifically, functional gains will likely be made when we can identify and track the population of antigen-specific T cells responding to tumor and characterize the recovery of broad arrays of T-cell subsets as determined by specific Vα and Vβ chains of the T-cell receptor. Two recent developments, tetramers and T-cell spectrotyping, are critically important in this regard.

Tetramers are created by the fusion of four class I major histocompatibility receptors loaded with specific peptides. These molecules are used to identify T cells with T-cell receptors specific for the loaded peptide. When tetramers are conjugated with specific fluorochromes, they are effective in labeling peptide-specific T cells whose frequency can subsequently be determined by flow cytometry.[90] This technique allows accurate monitoring of the peptide-specific T-cell response in situations in which the antigen of interest is characterized and is unique (eg, peptide-loading strategies for DCs with a defined tumor-specific antigen). In certain circumstances, this technique may be used to replace the mixed lymphocyte response, which is labor intensive and experience dependent, allowing accurate monitoring of peptide-specific immunologic response.

T-cell spectrotyping relies on specific RNA primers to the variable regions of the T-cell receptor, Vα and Vβ, to define the presence of T-cell subsets within blood or tissue.[91] Although this technique does not demonstrate T-cell specificity for individual antigens, it defines the role of specific T-cell subsets in the antitumor immune response and characterizes changes that occur as a result of therapy. Tetramers and spectrotyping are becoming more widely available and serve as primary examples of the integral role of monitoring in the evolution of tumor immunotherapy.

REFERENCES

1. Gross L. Intradermal immunization of C3H mice against a sarcoma that originated in an animal of the same line. Cancer Res 1943;3:326–33.
2. Foley EJ. Antigenic properties of methylcholanthrene-induced tumors in mice of the strain of origin. Cancer Res 1953;13:835–7.
3. Prehn RT, Main JM. Immunity to methylcholanthrene-induced sarcomas. J Natl Cancer Inst 1957;18:769–78.
4. Klein G, Sjögren HO, Klein E, Hellstrom KE. Demonstration of resistance against methylcholanthrene-induced sarcomas in the primary autochthonous host. Cancer Res 1960;20:1561–72.
5. Ritts RE Jr, Neel HB III. An overview of cancer immunology. Mayo Clin Proc 1974;49:118–31.
6. Lewis MR, Aptekman PM. Atrophy of tumors caused by strangulation and accompanied by development of tumor immunity in rats. Cancer 1952;5:411–3.
7. Takeda K, Aizawa M, Kikuchi Y, et al. Tumor autoimmunity against methylcholanthrene-induced sarcomas of the rat. Gann 1966;57:221–40.
8. Neel HB III, Ketcham AS, Hammond WG. Requisites for successful cryogenic surgery of cancer. Arch Surg 1971;102:45–8.
9. Neel HB III, Ketcham AS, Hammond WG. Experimental evaluation of in situ oncocide for primary tumor therapy: comparison of tumor-specific immunity after complete excision, cryonecrosis and ligation. Laryngoscope 1973;83:376–87.
10. Neel HB III, Ritts RE Jr. Tumor-specific immunity and the otolaryngologist. Ann Otol Rhinol Laryngol 1973;82:323–9.
11. Neel HB III, Ritts RE Jr. Transfer of tumor-specific immunity with syngeneic spleen cells and serum from mice that have large tumors and metastases. Cancer 1977;40:1643–9.
12. Neel HB III, Ritts RE Jr. Immunotherapeutic effect of tumor necrosis after cryosurgery, electrocoagulation, and ligation. J Surg Oncol 1979;11:45–52.
13. Thomas L. Discussion. In: Lawrence HS, editor. Cellular and humoral aspects of the hypersensitive states New York: Paul B. Hoeber; 1959. p. 529–32.
14. Burnet FM. The concept of immunological surveillance. Prog Exp Tumor Res 1970;13:1–27.
15. Gatti RA, Good RA. Occurrence of malignancy in immunodeficiency diseases. A literature review. Cancer 1971;28:89–98.
16. Catalona WJ, Sample WF, Chretien PB. Lymphocyte reactivity in cancer patients: correlation with tumor histology and clinical stage. Cancer 1973;31:65–71.
17. Arons MS, Smith RR. Distant metastases and local recurrence in head and neck cancer. Ann Surg 1961;154:235–40.
18. Moller G, Moller E. Considerations of some current concepts in cancer research. J Natl Cancer Inst 1975; 55:755–9.
19. Outzen HC, Custer RP, Eaton GJ, Prehn RT. Spontaneous and induced tumor incidence in germfree "nude" mice. J Reticuloendothel Soc 1975;17:1–9.
20. Rygaard J, Povlsen CO. The nude mouse vs. the hypothesis of immunological surveillance. Transplant Rev 1976;28:43–61.
21. Schwartz RS. Another look at immunologic surveillance. N Engl J Med 1975;293:181–4.
22. Herberman RB, Holden HT. Natural cell-mediated immunity. Adv Cancer Res 1978;27:305–77.
23. Talmadge JE, Meyers KM, Prieur DJ, Starkey JR. Role of NK cells in tumour growth and metastasis in beige mice. Nature 1980;284:622–4.
24. Marx JL. Natural killer cells help defend the body. Science 1980;210:624–6.
25. Takasugi M, Ramseyer A, Takasugi J. Decline of natural nonselective cell-mediated cytotoxicity in patients with tumor progression. Cancer Res 1977;37:413–8.
26. Pross HF, Baines MG. Spontaneous human lymphocyte-mediated cytotoxicity against tumour target cells. I. The effect of malignant disease. Int J Cancer 1976;18:593–604.
27. Schantz SP, Goepfert H. Multimodality therapy and distant metastases. The impact of natural killer cell activity. Arch Otolaryngol Head Neck Surg 1987;113: 1207–13.
28. Ryan RE Jr, Neel HB III, Ritts RE Jr. Correlation of preoperative immunologic test results with recurrence in patients with head and neck cancer. Otolaryngol Head Neck Surg 1980;88:58–63.
29. Olkowski ZL, Wilkins SA Jr. T-lymphocyte levels in the peripheral blood of patients with cancer of the head and neck. Am J Surg 1975;130:440–4.
30. Wanebo HJ, Jun MY, Strong EW, Oettgen H. T-cell deficiency in patients with squamous cell cancer of the head and neck. Am J Surg 1975;130:445–51.
31. Balaram P, Vasudevan DM. Quantitation of Fc receptor-bearing T-lymphocytes (TG and TM) in oral cancer. Cancer 1983;52:1837–40.
32. Saxon A, Portis J. Lymphoid subpopulation changes in regional lymph nodes in squamous head and neck cancer. Cancer Res 1977;37:1154–8.

33. Yata J, Shimbo T, Sawaki S. Changes in T-cell subsets and their clinical significance in cancer patients. IARC Sci Publ 1978;20:511–21.
34. Wolf GT, Lovett EJ III, Peterson KA, et al. Lymphokine production and lymphocyte subpopulations in patients with head and neck squamous carcinoma. Arch Otolaryngol 1984;110:731–5.
35. Wolf GT, Hudson JL, Peterson KA, et al. Lymphocyte subpopulations infiltrating squamous carcinomas of the head and neck: correlations with extent of tumor and prognosis. Otolaryngol Head Neck Surg 1986; 95:142–52.
36. Wolf GT, Schmaltz S, Hudson J, et al. Alterations in T-lymphocyte subpopulations in patients with head and neck cancer. Correlations with prognosis. Arch Otolaryngol Head Neck Surg 1987;113:1200–6.
37. Katz AE, Yoo TJ, Harker LA. Serum immunoglobulin A (IgA) levels in carcinoma of the head and neck. Trans Am Acad Ophthalmol Otolaryngol 1976;82: 131–7.
38. Wara WM, Wara DW, Phillips TL, Ammann AJ. Elevated IgA in carcinoma of the nasopharynx. Cancer 1975;35:1313–5.
39. Mathew GD, Qualtiere LF, Neel HB III, Pearson GR. IgA antibody, antibody-dependent cellular cytotoxicity and prognosis in patients with nasopharyngeal carcinoma. Int J Cancer 1981;27:175–80.
40. Sundar SK, Ablashi DV, Kamaraju LS, et al. Sera from patients with undifferentiated nasopharyngeal carcinoma contain a factor which abrogates specific Epstein-Barr virus antigen-induced lymphocyte response. Int J Cancer 1982;29:407–12.
41. Seder RH, Vaughan CW, Oh SK, et al. Tumor regression and temporary restoration of immune response after plasmapheresis in a patient with recurrent oral cancer. Cancer 1987;60:318–25.
42. Veltri RW, Maxim PE. Tumor immunity and tumor markers in head and neck cancer. In: Wolf GT, editor. Head and neck oncology. Boston: Martinus Nijhoff; 1984. p. 411–30.
43. Pruet CW, Ghosh BC. Antigens and markers in head and neck cancer. In: Ghosh BC, Ghosh L, editors. Tumor markers and tumor-associated antigens. New York: McGraw-Hill; 1987. p. 313–7.
44. Berg NJ, Neel HB III, Weiland LH, Spelsberg TC. A new assay to assess steroid-hormone responsiveness in head and neck cancer. Laryngoscope 1987;97:286–90.
45. Longo DL. Tumor markers: current status of the quest—introductory overview. Semin Oncol 1987; 14:85–8.
46. Jackson CE, Tashjian AH Jr, Block MA. Detection of medullary thyroid cancer by calcitonin assay in families. Ann Intern Med 1973;78:845–52.
47. Henle G, Henle W. Epstein-Barr virus-specific IgA serum antibodies as an outstanding feature of nasopharyngeal carcinoma. Int J Cancer 1976; 17:1–7.
48. Klein G. The Epstein-Barr virus. In: Kaplan AS, editor. The herpesviruses. New York: Academic Press; 1973. p. 521–55.
49. Neel HB III, Pearson GR, Taylor WF. Antibodies to Epstein-Barr virus in patients with nasopharyngeal carcinoma and in comparison groups. Ann Otol Rhinol Laryngol 1984;93:477–82.
50. Neel HB III, Pearson GR, Weiland LH, et al. Application of Epstein-Barr virus serology to the diagnosis and staging of North American patients with nasopharyngeal carcinoma. Otolaryngol Head Neck Surg 1983;91:255–62.
51. Neel HB III. A prospective evaluation of patients with nasopharyngeal carcinoma: an overview. J Otolaryngol 1986;15:137–44.
52. Neel HB III, Pearson GR, Weiland LH, et al. Immunologic detection of occult primary cancer of the head and neck. Otolaryngol Head Neck Surg 1981;89:230–4.
53. Neel HB III, Pearson GR, Taylor WF. Antibody-dependent cellular cytotoxicity. Relation to stage and disease course in North American patients with nasopharyngeal carcinoma. Arch Otolaryngol 1984; 110:742–7.
54. Neel HB III, Taylor WF, Pearson GR. Prognostic determinants and a new view of staging for patients with nasopharyngeal carcinoma. Ann Otol Rhinol Laryngol 1985;94:529–37.
55. Neel HB III, Pearson GR, Weiland LH, et al. Anti-EBV serologic tests for nasopharyngeal carcinoma. Laryngoscope 1980;90:1981–90.
56. Gostout BS, Podratz KC, McGovern RM, Persing DH. Cervical cancer in older women: a molecular analysis of human papillomavirus types, HLA types, and p53 mutations. Am J Obstet Gynecol 1998;179:56–61.
57. Garcia-Milian R, Hernandez H, Panade L, et al. Detection and typing of human papillomavirus DNA in benign and malignant tumours of laryngeal epithelium. Acta Otolaryngol (Stockh) 1998;118:754–8.
58. Matsuba HM, Thawley SE, Spector GJ, et al. Laryngeal epidermoid carcinoma associated with juvenile laryngeal papillomatosis. Laryngoscope 1985;95: 1264–6.

59. Helmuth RA, Strate RW. Squamous carcinoma of the lung in a nonirradiated, nonsmoking patient with juvenile laryngotracheal papillomatosis. Am J Surg Pathol 1987;11:643–50.
60. Beck JC, McClatchey KD, Lesperance MM, et al. Human papillomavirus types important in progression of inverted papilloma. Otolaryngol Head Neck Surg 1995;113:558–63.
61. Weiner JS, Sherris D, Kasperbauer J, et al. Relationship of human papillomavirus to schneiderian papillomas. Laryngoscope 1999;109:21–6.
62. Orvidas LJ, Lewis JE, Olsen KD, Weiner JS. Intranasal verrucous carcinoma: relationship to inverting papilloma and human papillomavirus. Laryngoscope 1999;109:371–5.
63. Thomas M, Pim D, Banks L. The role of the E6-p53 interaction in the molecular pathogenesis of HPV. Oncogene 1999;18:7690–700.
64. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 1975;256:495–7.
65. Goldenberg DM. New developments in monoclonal antibodies for cancer detection and therapy. CA Cancer J Clin 1994;44:43–64.
66. Strome SE, Krauss JC, Chang AE, Shu S. Adoptive cellular immunotherapy of human melanoma. In: Leong SPL, editor. Immunotherapy of malignant melanoma. Medical Intelligence Unit. Austin (TX): RG Landes Company; 1996. p. 97–113.
67. Hurtado JC, Kim YJ, Kwon BS. Signals through 4-1BB are costimulatory to previously activated splenic T cells and inhibit activation-induced cell death. J Immunol 1997;158:2600–9.
68. Lenschow DJ, Walunas TL, Bluestone JA. CD28/B7 system of T cell costimulation. Annu Rev Immunol 1996;14:233–58.
69. Melero I, Bach N, Hellstrom KE, et al. Amplification of tumor immunity by gene transfer of the co-stimulatory 4-1BB ligand: synergy with the CD28 co-stimulatory pathway. Eur J Immunol 1998;28: 1116–21.
70. Tamada K, Shimozaki K, Chapoval AI, et al. LIGHT, a TNF-like molecule, costimulates T cell proliferation and is required for dendritic cell-mediated allogeneic T cell response. J Immunol 2000;164: 4105–10.
71. Nagata S. Fas ligand-induced apoptosis. Annu Rev Genet 1999;33:29–55.
72. Dong H, Zhu G, Tamada K, Chen L. B7-H1, a third member of the B7 family, co-stimulates T-cell proliferation and interleukin-10 secretion. Nat Med 1999;5:1365–9.
73. Steinman RM. The dendritic cell system and its role in immunogenicity. Annu Rev Immunol 1991;9:271–96.
74. Nestle FO, Alijagic S, Gilliet M, et al. Vaccination of melanoma patients with peptide- or tumor lysate-pulsed dendritic cells. Nat Med 1998;4:328–32.
75. Chen L, Linsley PS, Hellstrom KE. Costimulation of T cells for tumor immunity. Immunol Today 1993; 14:483–6.
76. Mizoguchi H, O'Shea JJ, Longo DL, et al. Alterations in signal transduction molecules in T lymphocytes from tumor-bearing mice. Science 1992; 258:1795–8.
77. Liebowitz DN, Lee KP, June CH. Costimulatory approaches to adoptive immunotherapy. Curr Opin Oncol 1998;10:533–41.
78. Mayordomo JI, Zorina T, Storkus WJ, et al. Bone marrow-derived dendritic cells pulsed with synthetic tumour peptides elicit protective and therapeutic antitumour immunity. Nat Med 1995;1:1297–302.
79. Boczkowski D, Nair SK, Snyder D, Gilboa E. Dendritic cells pulsed with RNA are potent antigen-presenting cells in vitro and in vivo. J Exp Med 1996; 184:465–72.
80. Gilboa E, Nair SK, Lyerly HK. Immunotherapy of cancer with dendritic-cell-based vaccines. Cancer Immunol Immunother 1998;46:82–7.
81. Ashley DM, Faiola B, Nair S, et al. Bone marrow-generated dendritic cells pulsed with tumor extracts or tumor RNA induce antitumor immunity against central nervous system tumors. J Exp Med 1997;186: 1177–82.
82. Porgador A, Snyder D, Gilboa E. Induction of antitumor immunity using bone marrow-generated dendritic cells. J Immunol 1996;156:2918–26.
83. Tsai V, Kawashima I, Keogh E, et al. In vitro immunization and expansion of antigen-specific cytotoxic T lymphocytes for adoptive immunotherapy using peptide-pulsed dendritic cells. Crit Rev Immunol 1998;18:65–75.
84. Kugler A, Stuhler G, Walden P, et al. Regression of human metastatic renal cell carcinoma after vaccination with tumor cell-dendritic cell hybrids. Nat Med 2000;6:332–6.
85. Arpinati M, Green CL, Heimfeld S, et al. Granulocyte-colony stimulating factor mobilizes T helper 2-inducing dendritic cells. Blood 2000;95:2484–90.
86. Rosenberg SA, Yang JC, Schwartzentruber DJ, et al. Immunologic and therapeutic evaluation of a syn-

thetic peptide vaccine for the treatment of patients with metastatic melanoma. Nat Med 1998;4:321–7.
87. Plautz GE, Miller DW, Barnett GH, et al. T cell adoptive immunotherapy of newly diagnosed gliomas. Clin Cancer Res 2000;6:2209–18.
88. Plautz GE, Barnett GH, Miller DW, et al. Systemic T cell adoptive immunotherapy of malignant gliomas. J Neurosurg 1998;89:42–51.
89. Kjaergaard J, Shu S. Tumor infiltration by adoptively transferred T cells is independent of immunologic specificity but requires down-regulation of L-selectin expression. J Immunol 1999;163:751–9.
90. Altman JD, Moss PAH, Goulder PJR, et al. Phenotypic analysis of antigen-specific T lymphocytes. Science 1996;274:94–6.
91. Gorski J, Yassai M, Zhu X, et al. Circulating T cell repertoire complexity in normal individuals and bone marrow recipients analyzed by CDR3 size spectratyping: correlation with immune status. J Immunol 1994;152:5109–19.

SUGGESTED READING

Shu S, Strome SE, Krauss JC, et al. Tumor-draining lymph nodes as a source for generating therapeutic immune T cells. In: Chang AE, Shu S, editors. Immunotherapy of cancer with sensitized T lymphocytes. Medical Intelligence Unit. Austin (TX): RG Landes Company; 1994. p. 1–14.

Strome SE, Krauss JC, Chang AE, Shu S. Adoptive cellular immunotherapy of human melanoma. In: Leong SPL, editor. Immunotherapy of malignant melanoma. Medical Intelligence Unit Austin (TX): RG Landes Company; 1996. p. 97–113.

CHAPTER 56

Molecular Biology of Squamous Cell Carcinoma

Carter Van Waes, MD, PhD

Head and neck squamous cell carcinoma (HNSCC) is the most common cancer that affects both human communication and survival. It accounts for 90% of the cancers arising in the upper aerodigestive tract, making it the most common type of cancer and cause of cancer deaths among patients with head and neck cancer. It accounts for over 40,000 new cancer cases and 11,000 deaths annually in the United States.[1] Despite progress in refinement of surgery, radiation, and chemotherapeutic approaches that have enhanced organ preservation, there has been no improvement in survival of patients with HNSCC over the past 25 years.[1] Recent advances in determining the identity and function of molecular events involved in the pathogenesis of HNSCC have provided a foundation for development of new methods for screening, prevention, and therapy in the new century.

Progress in understanding the molecular biology of HNSCC has been made possible through advances in technology that have permitted detection and mapping of deoxyribonucleic acid (DNA) sequences and genes throughout the human genome.[2] *Cytogenetic analysis* is a method that has enabled detection of chromosomal abnormalities that are large enough to be seen with special stains and light microscopy. Abnormalities such as deletions, amplifications, and translocations have been detected in certain chromosomal regions with increased frequency in HNSCC and other cancers. In some cases, these abnormalities have led to the identification of specific genes whose deletion, amplification, or truncation is involved in development of cancer. *DNA polymorphism analysis* is another important tool made possible by identification of differences in DNA sequence fragments between individuals at chromosomal locations throughout the genome. With this method, different-size DNA fragments cut by enzymes from the DNA contributed by each parent are identified with electrophoresis, allowing the detection of two copies (alleles) in normal cells from the patient. In cancer cells, loss of one marker, called a *loss of heterozygosity* (LOH), has been used to detect nearby alterations in DNA. DNA polymorphism and LOH analyses have permitted the detection and mapping of molecular alterations in DNA that were too small to be detected by cytogenetic analysis, thereby resulting in finer maps and detection of additional submicroscopic DNA abnormalities. Further mapping and sequencing of these regions have been made possible by development of methods for cloning large fragments of chromosomal DNA and methods for high through-put DNA sequencing, resulting in the positional cloning of many of the mapped genes. The recent development of a rough draft of the sequence of the human genome from high throughput DNA sequencing should rapidly result in the identification of genes remaining to be discovered.

The identification of genes and gene fragments has enabled studies of differences in expression in HNSCC of messenger ribonucleic acids (RNAs) using Northern blot and polymerase chain reaction methods and of proteins with immunoassays. Recently, technology has enabled reduction of these assays into *microarrays*, which have the capability to compare simultaneously and detect differences among thousands of different messenger RNAs or proteins in samples from normal and malignant tissue. Studies to determine the relative expression and function of genes in the cell are leading to a better understanding of the pathways involved in the pathogenesis of cancer and their potential usefulness for molecular diagnosis and therapy. As a result of these advances, it seems likely that in the next two decades, prevention and treatment of HNSCC will

increasingly involve use of molecular assays for prediction of prognosis and response to therapy and targeted prevention and therapy using molecular medicine and "gene surgery."

Carcinogenesis of Head and Neck Squamous Cell Carinoma

Development of HNSCC has been associated with repeated exposure to and injury by chemical carcinogens contained in tobacco and alcohol[2] or chronic infection by human papillomavirus (HPV)[3] or Epstein-Barr virus.[4] Several physical and viral causative agents associated with the development of HNSCC can cause defined molecular changes that result in malignancy. Worldwide, tobacco and alcohol products are the leading risk factors for development of HNSCC. Tobacco contains aromatic hydrocarbons that can cause DNA damage, most commonly resulting in G:C to T:A base-pair transversions and G:C to A:T base-pair transitions.[5] Alcohol appears to increase risk when used in combination with tobacco. Other physical agents that appear to contribute to the incidence of HNSCC at specific subsites can cause DNA damage. The increased incidence of lip and skin cancer in regions nearer the equator such as in Australia and the southwestern United States is associated with sun exposure. The ultraviolet spectrum in sunlight can cause thymidine dimer formation, resulting in C to T base transitions.[5]

Certain viruses have been associated with HNSCC, and these viruses can commandeer vital cellular control pathways and cause DNA disruption on integration. Human papillomavirus type 16, which is the subtype that causes cervical carcinoma in women, has recently been shown to be prevalent in oropharyngeal carcinomas, including nonsmokers.[3] Cells infected with HPVs express proteins encoded by viral genes called E6 and E7 that can inactivate the p53 and retinoblastoma proteins that help regulate cell proliferation and death. Epstein-Barr virus has been associated with nasopharyngeal carcinoma, particularly in Asia.[4] It encodes viral proteins that can promote immortalization and proliferation of epithelia and lymphocytes. Epstein-Barr virus also carries a gene that encodes a homologue of human interleukin-10 (IL-10), an immune hormone that can suppress development of cytotoxic T-lymphocyte immunity, which is needed for immune destruction of cancer cells.

Critical Events in the Development of Cancer

Tumor development following exposure to carcinogenic agents involves molecular changes that affect key cellular functions necessary for malignant behavior. These changes include overriding the normal program for cell differentiation and death and favoring an increase in cell proliferation and life span.[6] Malignant tumor progression involves additional changes that result in establishment of a blood supply, migration, invasion, and metastasis within the patient host.

Weinberg has shown that molecular alteration in expression or function of a minimum of three genes is required to alter important cellular functions and cause transformation and tumor development from normal human cells under experimental conditions.[7] One important requirement for neoplastic transformation and tumor formation is an increase in *cell proliferation*. He showed that increased cell-cycle progression and proliferation can result from activation of a signaling kinase called Ras. Ras may be activated by mutation or by activation of growth factor receptors. However, Ras activation and increased proliferation alone are not sufficient as such cells eventually reach the end of their life span, undergoing "crisis" and death. Thus, a second requirement is an increase in *cell life span*, or *immortalization*. Cells normally stop proliferating and reach the end of their life in a finite number of cell divisions. The life-span limit of normal cells has been shown to be associated with a gradual shortening of the telomeres, which are the ends of chromosomes. In contrast, cancer cells do not exhibit shortening of the telomeres, and this has been found to be attributable to increased expression of an enzyme called telomerase. Experimental expression of telomerase in normal cells transformed with Ras provides a second event needed to extend cell life span. Third, cancer cells exhibit a decrease in the rate of programmed cell death, called *apoptosis*. Weinberg showed that inactivation of mechanisms that prevent apoptosis, such as by a viral gene, is a requirement for transformation and tumor development. Inactivation of certain genes in the cell, by viruses, mutation, or loss, can affect the cell-death program. Inactivation or loss of one such gene called p53, which encodes a 53,000 kDa

protein, can reduce the ability of the cell to undergo cell-cycle arrest, repair damaged DNA, or undergo apoptosis when DNA damage cannot be repaired. Because inactivation of p53 may result in replication of cells with DNA mutations and damage, changes in expression of additional genes may occur, leading to further increases in malignant behavior. Genes that encode proteins necessary for normal behavior that stop uncontrolled growth or are necessary for death of aging or damaged cells, such as p53, are termed *tumor suppressor genes*.[7] When abnormal expression of genes "turns on" transformation and growth, these genes have been called *oncogenes*.[7] Genes expressed by viruses, such as E6 and E7, are called *viral oncogenes*. Genes such as the signal kinase Ras are termed *cellular oncogenes*. Altered expression of different combinations of oncogenes and tumor suppressor genes may result in increased proliferation and life span and decreased cell death, leading to tumor formation.

Cellular Pathologic Changes during Development of Head and Neck Squamous Cell Carcinoma

The initial development of HNSCC involves a series of progressive pathologic changes in behavior of cells in squamous cell mucosa and underlying stroma of the upper aerodigestive tract, as illustrated in Figure 56–1, A. Cells injured by exposure to physical or viral agents may undergo irreversible damage and cell death or undergo repair and survive. The surviving cells often acquire increased cellular resistance and proliferate and migrate, healing the area of injury. With repeated injury, increased numbers of cells may accumulate to produce a thickened epithelium, termed *hyperplasia*. With repeated exposure to mutagenic agents, irreversible genetic mutations may occur, with changes in cell behavior that include a partial- or full-thickness increase in proliferating and morphologically atypical cells called *dysplasia* or *carcinoma in situ* (Cis), respectively. Dysplasia and Cis lesions are premalignant. Invasion and migration into the stroma are the hallmark of *invasive carcinoma*, which is the stage of malignant tumor development. Hyperplasia, dysplasia, or Cis may be accompanied by hyperkeratinization, which may be visible clinically as *leukoplakia*. Dysplasia, Cis, or microinvasive carcinoma is accompanied by a progressive increase in inflammation, vascularity, and proliferation of fibrous stroma, which may be visible clinically as induration and vascular erythema, called *erythroplasia*.

Molecular Pathologic Changes during Development of Head and Neck Squamous Cell Carcinoma

Early Expression of Epidermal Growth Factor Receptor and Transforming Growth Factor α in Head and Neck Squamous Cell Carcinoma The histopathologic cellular changes that occur during the stages of tumor development of HNSCC have been associated with an accumulation in molecular changes, as highlighted in Figure 56–1, B. One of the earliest events identified in development of HNSCC is overexpression of a growth factor receptor called epidermal growth factor receptor (EGFR) and one of its stimulatory factors, transforming growth factor-α (TGF-α)[8] (see Figure 56–1, B). Grandis and colleagues have demonstrated that over 90% of HNSCCs overexpress EGFR and TGF-α.[8,9] Increased expression of TGF-α and EGFR has been detected in tumor cells and mucosa from patients with HNSCC, when compared with normal mucosa from controls, indicating that TGF-α and EGFR expression may be an early event in carcinogenesis of HNSCC.[8,9] Increased expression of both TGF-α and EGFR occurs with progression to carcinoma (see Figure 56–1, B). Epidermal growth factor receptor and TGF-α appear to be overexpressed owing to transcriptional activation of the genes for this receptor and ligand in most HNSCC.[8] Patients with carcinomas expressing higher levels of these factors have been shown to have a shortened disease-free survival, independent of cervical lymph node stage.[9]

Production of TGF-α and expression of EGFR establish a potential autocrine signal pathway for continuous stimulation of proliferation of squamous cells, and the importance of TGF-α and EGFR expression in growth of HNSCC has been established. Inhibition of either TGF-α or EGFR expression or function using antisense oligonucleotides[10,11] or pharmacologic inhibitors in combination with radiation[12,13] was found to decrease proliferation and growth of HNSCC cells in vitro and of xenografts in mice in vivo. The effects of EGFR activation may be mediated through several signal pathways, which include Ras, mitogen-activated protein kinase, phosphatidylinositol-3-kinase, and signal transduc-

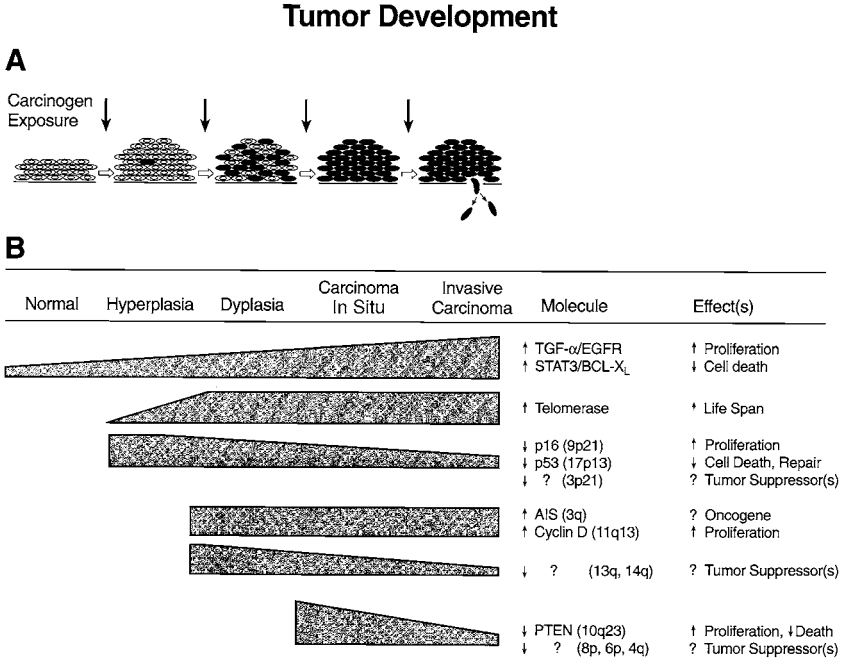


FIGURE 56–1. Model of histopathologic and molecular changes with tumor development of head and neck squamous cell carcinoma (HNSCC). *A*, Histopathologic changes during tumor development of HNSCC occur following repeated carcinogen exposure, as described in the text. *B*, Molecular changes with tumor development. TGF = transforming growth factor; EGFR = epidermal growth factor receptor; STAT = signal transducers and activators of transcription; BCL-X_L = BCL-XLong; p = protein; AIS = activated in squamous carcinoma; PTEN = phosphatase and tensin homolog deleted on chromosome 10.

ers and activators of transcription 3 (STAT3). Grandis et al have shown that EGFR activation of STAT3 prevents HNSCC cells from undergoing cell death (apoptosis) and stimulates proliferation.[14] Thus, overexpression of EGFR appears to contribute to the increase in proliferation as well as decrease in cell death involved in tumorigenesis.

Early Activation of Telomerase in Head and Neck Squamous Cell Carcinoma Early activation of telomerase has been detected in HNSCC. Mao et al reported that although telomerase is not detected in normal mucosa, increased telomerase activity is detected with development of squamous hyperplasia, dysplasia, and invasive carcinoma.[15] Califano et al have also shown that increased telomerase activity may be detected in dysplasia and invasive carcinoma and, with lower sensitivity, in shed cells in oral rinses from patients.[16] The increased expression of telomerase observed in HNSCC is consistent with the requirement for immortalization of cancer cells in a study of carcinogenesis by Weinberg.[7]

Tumor Suppressor Genes and Oncogenes in Development of Head and Neck Squamous Cell Carcinoma There is evidence for inactivation of tumor suppressor genes and activation of oncogenes at several chromosomal loci during development of HNSCC. Califano et al have shown that the LOH at several chromosomal loci occurs with increasing frequency with different stages of development of HNSCC[17] (see Figure 56–1, B).

A locus on the short (p) arm of chromosome 9 located at 9p21 has been found to exhibit LOH at an early stage during development of hyperplasia.[17] Chromosome 9p21 has been found to encode a protein called p16, which is a cyclin-dependent kinase called CDKN2/MTS-1/INK4A that normally inhibits cell-cycle progression. p16 was found to be inactivated in the majority of HNSCC by homozygous deletion, by methylation of the regulatory promoter region of the gene, or by mutation,[18] thereby releasing squamous cells to proliferate. Re-expression of p16 in HNSCC cells by gene transfer suppressed cell growth in vitro.[19] Loss of p16 in HNSCC may be detected by immunostaining as well as molecular methods.

A locus on the long (q) arm of chromosome 11 located at 11q13 demonstrates LOH in increasing frequency during the Cis stage[17] and is associated with amplification of a cell-cycle regulatory protein called cyclin D1.[20] Increased expression of cyclin D1 is prevalent in HNSCC. Cyclin D1 is required for progression of cells through the cell cycle, thereby stimulating squamous cells to proliferate. Inhibition of cyclin D1 results in inhibition of growth of HNSCC in vitro.

Chromosome 3p is altered with increasing frequency with development of dysplasia in HNSCC

and lung cancers and contains at least three putative loci, the identity of which remain to be determined.[17,21] Recently, a homologue of p53 has been identified on 3q, and this protein, designated p40AIS (amplified in squamous cell carcinoma), was found to occur with high frequency in HNSCC and lung squamous cell carcinoma.[22] Amplified in squamous cell carcinoma is a homologue of the p53 tumor suppressor gene but lacks the tumor suppressor function of p53. Expression of AIS was found to promote growth of rat cells in soft agar and in mice, indicating that it may play an early role in transformation as an oncogene. Increased p40 expression appears to be correlated with loss of p53, and p40 may interact and inhibit normal p53 function.

Chromosome 17p13 encodes the p53 gene, which is deleted or mutated with relatively low frequency in dysplasia and Cis and with increased frequency in primary carcinomas.[17] A variety of functions have been attributed to the p53 gene,[23] which are thought to result primarily in tumor suppression. p53 is expressed in response to DNA damage, resulting in cell-cycle arrest for the repair of DNA damage or induction of cell death when damage is irreversible.[23] Loss of p53 presumably eliminates mechanisms by which cancer cells undergo repair or cell death.

Chromosome 13q21 has been shown to exhibit increasing LOH with Cis and invasive carcinoma.[17] It includes the region encoding the retinoblastoma gene, another tumor suppressor gene. However, mutation of the retinoblastoma gene occurs relatively infrequently in the HNSCC studied, indicating the possible presence of another gene at this location.

Loss of heterozygosity of loci on 4q occurs with increased frequency in invasive carcinomas, suggesting that the gene(s) located therein are associated with later stages of primary tumor development. To date, the identity of genes and predictive value of alterations of these loci in prognosis have not been demonstrated.[24]

Loss of heterozygosity involving chromosomes 8p and 6p is observed in invasive HNSCC, and a locus on chromosome 8p23 is associated with poor prognosis.[25] The putative tumor suppressor genes at the chromosome 8p loci remain to be identified. A p53 regulated growth inhibitory gene that is a member of the tumor necrosis factor–related apoptosis-inducing ligand (TRAIL) family has been mapped to the chromosome 8p21 region in HNSCC and infrequently undergoes mutations that result in loss of function.[26]

Loss of heterozygosity at chromosome 10q23 is detected with intermediate frequency in primary HNSCC and has been found to reflect deletion or inactivation of a gene called phosphatase and tensin homolog deleted on chromosome 10 (*PTEN*).[27] *PTEN* functions in the inactivation of Akt, a protein kinase. Akt is activated by EGFR and other cytokine and growth factor receptors. Early activation of Akt by overexpression of EGFR, and later by loss of *PTEN*, could lead to increased survival and growth of HNSCC and other cancers.

In summary, HNSCCs have been found to accumulate a series of genetic changes during the histopathologic stages of tumor development. The functions of the affected genes identified to date are consistent with the changes required for development of cancer, namely genes involved in regulation of cell-cycle progression and proliferation, cell life span, and cell death. The early occurrence of multiple events and detection of events at different stages in individual cancers from different patients suggest that different combinations of multiple genes affecting cell-cycle progression and proliferation, cell life span, and cell death can contribute to the common histopathologic stages of tumor development observed in HNSCC.

Tumor Progression and Metastasis of Squamous Cell Carcinoma The late stages of cancer involve progressive tumor invasion and metastasis, which are the stages that ultimately affect vital functions and cause death in patients. Figure 56–2 highlights some of the important histopathologic and molecular events associated with tumor progression and metastasis.

Development of *invasive carcinoma* is associated with focal dissolution of the basement membrane and extracellular matrix (ECM), detachment, and migration of cells into the submucosal tissue (see Figure 56–2, A). Head and neck squamous cell carcinomas that exhibit a streaming pattern of small clusters of cells through the ECM are associated with more aggressive behavior and poor prognosis.[28] *Tumor progression* to a size that becomes visible and has an effect on adjacent structures requires an increase in supply of oxygen and nutrients and removal of waste. Folkman is largely responsible for establishing the concept that new blood vessel for-

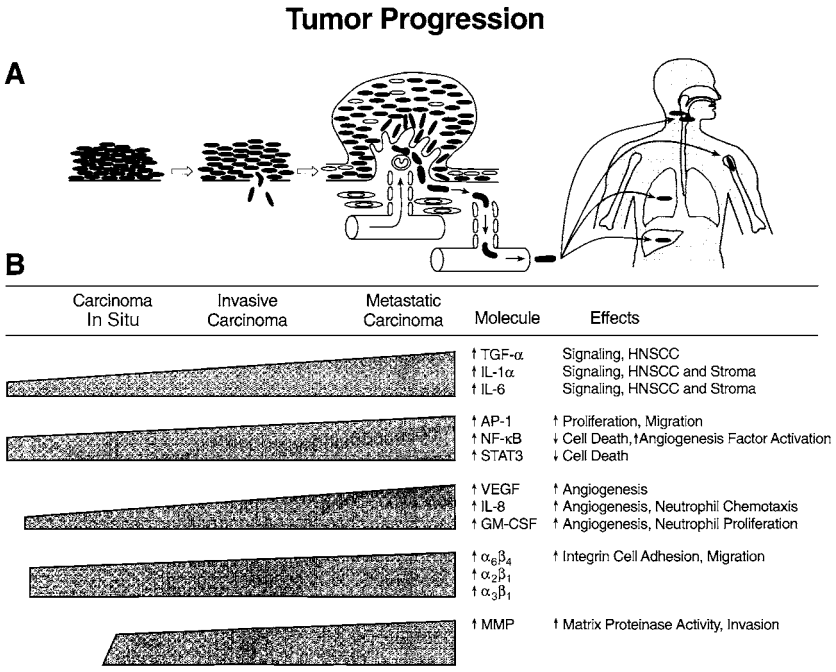


FIGURE 56–2. Model of histopathologic and molecular changes with tumor progression and metastasis of head and neck squamous cell carcinoma (HNSCC). *A*, Histopathologic changes during tumor development of HNSCC as described in the text. *B*, Molecular changes with tumor progression. TGF = transforming growth factor; IL = interleukin; AP = activator protein; NF-κB = nuclear factor-kappa B; STAT = signal transducers and activators of transcription; VEGF = vascular endothelial growth factor; GM-CSF = granulocyte-macrophage colony-stimulating factor; MMP = matrix metalloproteinase.

mation is critical in cancer.[29] Enlargement of tumors to a size beyond 0.5 cm exceeds the range for diffusion of oxygen from existing vessels and necessitates new blood vessel formation, called *neoangiogenesis*. Such new vessel formation has been demonstrated in all cancers and is commonly associated with an increase in inflammatory cells. Increased vessel density and inflammation within tumors have been associated with more rapid growth, metastasis, and a decrease in prognosis, suggesting that the increase in vessels may somehow relate to increased access for metastasizing cancer cells. Invasion of the lymphatics and blood vessels and *circulation* of cells are necessary for regional and distant spread of HNSCC.

Growth of the tumor epithelia and angiogenesis is also accompanied by increased infiltration of inflammatory cells and proliferation of fibrous stroma. Several studies have suggested that tumor cells that induce host inflammatory and stromal cell responses grow, invade, and metastasize more rapidly. Chen et al have shown that during progression, squamous cell carcinomas undergo additional changes needed for growth and metastasis that depend on the host.[30] Young and colleagues have shown that inflammatory cells infiltrating human and murine squamous cell carcinomas are one of the host components that promote growth and metastasis.[31] These inflammatory cells bear a stem cell marker called CD34 and appear to differentiate into granulocytes and the endothelial cells that form new blood vessels. Young[32] and Pekarek and colleagues[33] have shown that the granulocytes promote increased growth and metastasis. Granulocytes from the host can release growth factors[33] and proteases[34] that stimulate growth and invasion of tumor cells. Squamous cell carcinomas also induce proliferation of stromal fibroblasts. Fibroblasts also secrete factors and ECM substances that can promote growth.[35] The establishment of *metastases* requires cell arrest and vessel formation in a new location. Head and neck squamous cell carcinoma shows a predilection for metastases to the lymphatics, lungs, liver, and bone marrow, suggesting that the cells and substrate of the reticuloendothelial system provide a favorable environment for arrest and formation of squamous cell carcinoma metastases.

Molecules Involved in Cell Adhesion, Migration, and Invasion by Head and Neck Squamous Cell Carcinoma Head and neck squamous cell carcinomas exhibit alterations in expression of a repertoire of cell adhesion molecules and ECM substances that function in attachment and migration. Increased expression of cell adhesion molecules called *integrins* has been detected in HNSCC by Van Waes et al.[36] The integrins are heterodimers comprised of α and β subunits that form a superfamily of cell surface receptors involved in cell-cell and cell-

ECM adhesion and recognition. The integrin α6β4, α2β1, and α3β1 heterodimers are normally expressed among proliferating layers of squamous cell epithelium but are expressed in suprabasilar layers of many squamous cell carcinomas in association with increased proliferation and immortalization that occur during early tumor development.[36] Increased suprabasilar expression of α6β4, as detected by monoclonal antibody A9, was found to be correlated with a poorer prognosis in a prospective study of 80 patients with HNSCC.[37] Expression of integrin α6β4 has been shown to promote aggressive tumor behavior.[38] The α6β4, α2β1, and α3β1 integrins have been found to be receptors for ECM laminin, and α2β1 and $\alpha_3\beta_1$ integrins also bind to collagen. Head and neck squamous cell carcinomas secrete the basement membrane component laminin in vitro and in situ, and blockade of α6β4, α2β1, and α3β1 completely inhibits attachment of HNSCC to laminin.[39] Monoclonal antibody to the α6 integrin also has been shown to reduce binding to activated endothelial cells. Thus, HNSCCs exhibit constitutive alterations in expression of a repertoire of integrin cell adhesion molecules, and expression of integrin α6β4 in particular is linked with aggressive tumor behavior.

Increased expression and activation of enzymes involved in remodeling of the ECM have been detected in HNSCC and are associated with increased invasiveness and pathogenicity. The matrix metalloproteinases (MMPs) comprise a family of proteases that digest ECM, which are up-regulated in HNSCC (see Figure 56–2, B). Head and neck squamous cell carcinoma exhibits increased expression of urokinase-type plasminogen activator (uPA) and uPA receptor as well as the MMPs membrane-type MMP-1, collagenase 1, stromelysin 1, and gelatinase B.[40] Expression of these factors and invasiveness appears to be inducible by epidermal growth factor (EGF) or scatter factor (SF).[40–42] Invasion of EGF- or SF-stimulated cells is completely suppressed by recombinant and synthetic MMP inhibitors.[40] These studies suggest that MMPs may be more important than the plasminogen activator-plasmin system in mediating EGF- or SF-induced tumor cell invasion of interstitial matrix barriers. Inflammatory cells infiltrating tumors also can express MMPs. Matrix metalloproteinase-9 may be expressed by host-derived neutrophils, macrophages, and mast cells attracted to the tumor site. Matrix metalloproteinase-9 expressed by bone marrow–derived cells appears to promote malignant tumor progression of squamous cell carcinoma in vivo.[43]

Cytokine and Growth Factors Involved in Inflammatory and Angiogenesis Responses to Head and Neck Squamous Cell Carcinoma The demonstration of the role of host inflammatory and angiogenesis responses in progression and metastasis of squamous cell carcinoma has led to efforts to identify the molecules involved. Head and neck squamous cell carcinomas have been found to express a number of cytokines and growth factors that mediate inflammatory and angiogenesis responses (see Figure 56–2, B). Squamous cell carcinomas produce a repertoire of factors in vitro and in situ that include IL-1α, IL-6, IL-8, granulocyte-macrophage colony-stimulating factor (GM-CSF), growth-regulated oncogene-α (GRO-α), and vascular endothelial growth factor (VEGF).[44,45] Tumor fibroblasts have been shown to produce SF/hepatocyte growth factor (HGF).[35] Interleukin-6, IL-8, GRO-α, VEGF, and HGF are of sufficient stability and are produced at high enough concentrations that they have been detected in the serum of patients.[44]

Interleukin-1 is a cytokine that can regulate expression of several of the other factors detected in squamous cell carcinoma. Interleukin-1 serves as an autocrine factor to stimulate HNSCC to produce IL-8, GM-CSF, and VEGF [46] and as a paracrine factor to stimulate production of HGF by fibroblasts. In addition, IL-1 has important systemic regulatory effects as an important mediator of acute-phase reactions, increased catabolic state and cachexia often observed in patients with aggressive HNSCC. Interkeukin-1 and IL-6 both have direct effects as autocrine factors that can stimulate proliferation of HNSCC cells.[46,47]

Interleukin-8 and GRO-α are both members of a related family of chemoattractant and proliferative factors that contain a cysteine-X-cysteine (C-X-C) amino acid motif. Interleukin-8 and GRO-α have been shown to serve as chemoattractants for granulocytes and endothelial cells, which are major constituents of the inflammatory response in HNSCC. Loukinova et al have shown that expression of the murine homologue of GRO-α and IL-8 promotes aggressive growth and metastases, angiogenesis, and inflammatory cell infiltration in squamous cell carcinoma.[45] The aggressive pattern of growth is reversed in knockout mice deficient in CXC receptor

2, the receptor for the chemokine. These results provide direct evidence that tumor factors and the host response induced by them are critical in tumor progression and metastasis of squamous cell carcinoma. Kitadai et al have shown that IL-8 has similar effects on growth of other histologic types of human tumors as xenografts in mice.[48] Young et al have shown that GM-CSF produced by squamous cell carcinoma may also play a role in activity of CD34 progenitors of granulocytes and endothelium, which they reported are associated with metastasis.[31] Vascular endothelial growth factor also promotes angiogenesis and stimulates expression of αvβ3 and αvβ5 integrin heterodimers that are involved in migration of endothelial cells during angiogenesis. Inhibition of αv integrins inhibits tumorigenesis of squamous cell carcinomas in mice.[49]

Activation of Oncogenic Signal Pathways and Transcription Factors in Head and Neck Squamous Cell Carcinoma Identification of growth factors, cytokines, and other molecules that contribute to pathogenesis of squamous cell carcinoma has facilitated characterization of mechanisms of activation and potential targets for molecular therapy of HNSCC. Cytokines as well as integrins and MMPs expressed by HNSCC are also frequently expressed during the early response to injury, and these so-called "early response" genes are regulated by a group of transcription factors, which are proteins that usually bind DNA in the upstream (5′) noncoding portion of each gene. Transcription factors regulate expression of genes through transcription of messenger RNA. Transcription factor binding sequences in the regulatory region of cytokine genes expressed by HNSCC include nuclear factor-kappa B (NF-κB), activator protein-1 (AP-1), and nuclear factor IL-6 (NF IL-6), and these transcription factors were found to be activated in HNSCC.[50] Grandis et al have identified another transcription factor, called STAT3, which is activated and appears to regulate expression of BCL-XL, a molecule that prevents apoptosis of HNSCC cells.[14]

Studies performed to determine the effects of activation of transcription factors NF-κB and STAT3 in HNSCC have shown that they are important in the pathogenesis of HNSCC. Activation of both NF-κB and STAT3 appears to be important for preventing HNSCC cells from undergoing apoptosis.[14,51] Head and neck squamous cell carcinoma cells in which activation of NF-κB and STAT3 has been blocked by genetic methods exhibit increased cell death and inhibition of tumorigenesis in immune-deficient mice.[14,51] Inhibition of NF-κB also sensitizes HNSCC cells to cell death induced by radiation treatment and immune killing by tumor necrosis factor.[52,53] Molecular inhibitors of signal pathways that activate NF-κB, AP-1, and STAT3 are in development as potential agents for pharmacologic therapy of HNSCC.

Several of these transcription factors identified in HNSCC can be activated by signal receptors expressed on the surface of HNSCC, as illustrated in the model in Figure 56–3. Autocrine activation of TGF-α/EGFR, IL-1/IL-1 receptor, and IL-6/IL-6 receptor can potentially induce activation of NF-κB, AP-1, and STAT3.[47] In addition, squamous cell carcinomas have also been shown to express c-Met, which is a tyrosine kinase receptor for SF,[40,41] a factor produced as a result of paracrine activation of stromal fibroblasts by tumor cell–derived IL-1[35] (see Figure 56–3). c-Met may activate genes in HNSCC through the Ets oncogene family transcription factor or E1AF.[41] Together, these various signal pathways and transcription factors activate expression of many of the genes that have been shown to contribute to pathogenesis in HNSCC (see Figure 56–3). New drugs that target these receptors or intermediate kinase pathways are the subject of intensive preclinical and clinical development.

Molecular Diagnosis of Head and Neck Squamous Cell Carcinoma

Progress in the determination of molecular mechanisms involved in the pathogenesis of development and progression of HNSCC, together with sensitive molecular assay methods, has expanded opportunities for use of molecular diagnosis in patients with HNSCC. Using polymerase chain reaction–based methods for amplification of DNA, LOH of specific markers associated with HNSCC has been shown to occur in biopsy specimens of lesions prior to development of invasive malignancy.[17] Mutant p53 detected in surgical margins from patients has been shown to predict positive margins and sites of future recurrence.[54] Overexpression of telomerase and mutant p53 has been detected in tissue biopsies as well as in saliva,[16,23] opening possibilities for screening. Recently, detection of microsatellite and methylated DNA, which are markers for inactivated tumor

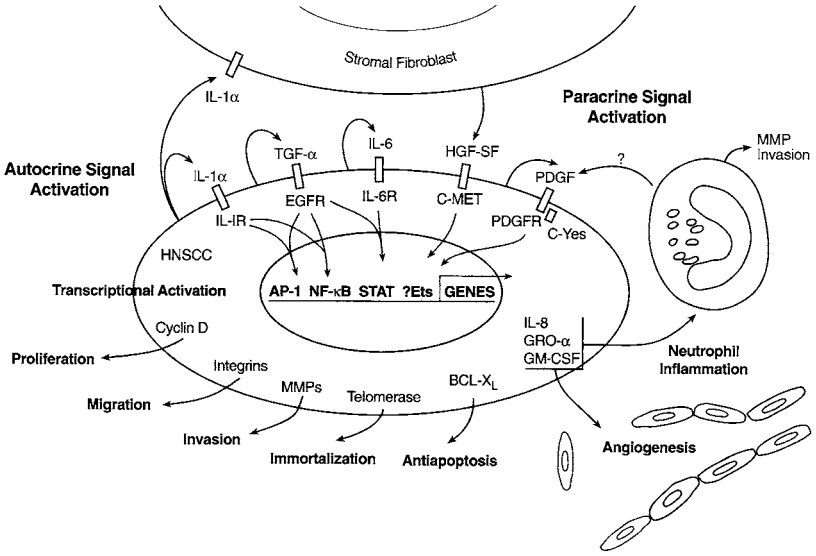


FIGURE 56–3. Summary of molecular changes affecting tumor and host cells that promote tumor development and progression of head and neck squamous cell carcinoma (HNSCC). Factors produced by tumor and host cells result in autocrine and paracrine signals that activate HNSCC transcription factors and genes. Genes activated by transcription factors cause cellular changes in HNSCC that lead to tumor development and host inflammatory and angiogenesis responses that promote malignant tumor progression and metastasis. TGF = transforming growth factor; EGFR = epidermal growth factor receptor; IL = interleukin; HGF-SF = hepatocyte growth factor-scatter factor; PDGF = platelet-derived growth factor; AP = activator protein; NF-κB = nuclear factor-kappa B; STAT = signal transducers and activators of transcription; GRO = growth-regulated oncogene; VEGF = vascular endothelial growth factor; GM-CSF = granulocyte-macrophage-colony-stimulating factor; MMP = matrix metalloproteinase.

suppressor genes, has been assayed in blood[55,56] and may eventually be useful for diagnosis and detection of recurrence. Detection of overexpressed proteins in blood, such as cytokines and angiogenesis factors, may also eventually prove useful in predicting behavior and recurrence.[44] These studies have established the feasibility of using molecular methods to screen patients for diagnosis, prognosis, and detection of recurrence. Further understanding of the "anatomy" of activation of pathways and target genes, as in Figure 56–3, will eventually offer the possibility of determining susceptibility or resistance of HNSCC to different therapies.

Molecular Therapy of Head and Neck Squamous Cell Carcinoma

The determination of molecular mechanisms involved in pathogenesis of development and progression of HNSCC has enabled the development and testing of specific molecular therapies. Replacement of mutant p53 by gene transfer inhibits growth of experimental tumors, and p53 gene therapy in patients with advanced HNSCC has established the feasibility of gene therapy in HNSCC.[57] Limited expression of genes administered within the tumor and the tendency of cancer cells lacking p53 to survive and become dominant provide obstacles to successful treatment of cancer by replacement of a normal gene. Identification of the role of EGFR in HNSCC and other cancers has led to testing of therapy with DNA antisense sequence to the target gene. Grandis et al and He et al have shown that EGFR and STAT3 antisenses inhibit growth of HNSCC,[10,11] and a clinical trial is under way to determine safety in patients and effects of expression of antisense DNA on tumor gene expression. Antibodies that inhibit activation of EGFR have been shown to inhibit growth of squamous cell carcinomas when given concurrently with radiation.[12,13,58] Small molecule inhibitors of EGFR and of components of the NF-κB and AP-1 pathways have been developed and are under study for use in patients with HNSCC.

REFERENCES

1. Greenlee RT, Murray T, Bolden S, Wingo PA. Cancer statistics, 2000. CA Cancer J Clin 2000;50:7–33.
2. Sidransky D. Molecular genetics of head and neck cancer. Curr Opin Oncol 1995;7:229–233.
3. Gillison ML, Koch WM, Capone RB, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst 2000;92:709–20.
4. Niedobitek G. Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma. Mol Pathol 2000;53:248–54.
5. Ostwald C, Gogacz P, Hillmann T, et al. p53 mutational spectra are different between squamous-cell carcinoma of the lip and the oral cavity. Int J Cancer 2000;88:82–6.
6. Hahn WC, Counter CM, Lundberg AS, et al. Creation of human tumor cells with defined genetic elements. Nature 1999;400:464–8.
7. Weinberg RA. Oncogenes and tumor suppressor genes. CA Cancer J Clin 1994;44:160–70.
8. Song JI, Grandis JR. STAT signaling in head and neck cancer. Oncogene 2000;19:2489–95.
9. Grandis JR, Melhem MF, Gooding WE, et al. Levels of TGF-alpha and EGFR protein in head and neck squamous cell carcinoma and patient survival. J Natl Cancer Inst 1998;90:824–32.
10. Grandis JR, Chakraborty A, Zeng Q, et al. Downmodulation of TGF-alpha protein expression with antisense oligonucleotides inhibits proliferation of head and neck squamous carcinoma but not normal mucosal epithelial cells. J Cell Biochem 1998;69:55–62.
11. He Y, Zeng Q, Drenning SD, et al. Inhibition of human squamous cell carcinoma growth in vivo by epidermal growth factor receptor antisense RNA transcribed from the U6 promoter. J Natl Cancer Inst 1998;90:1080–7.
12. Bonner JA, Raisch KP, Trummell HQ, et al. Enhanced apoptosis with combination C225/radiation treatment serves as the impetus for clinical investigation in head and neck cancers. J Clin Oncol 2000;18:47S–53S.
13. Milas L, Mason K, Hunter N, et al. In vivo enhancement of tumor radioresponse by C225 antiepidermal growth factor receptor antibody. Clin Cancer Res 2000;6:701–8.
14. Grandis JR, Drenninig SD, Zeng Q, et al. Constitutive activation of Stat3 signaling abrogates apoptosis in squamous cell carcinogenesis in vivo. Proc Natl Acad Sci U S A 2000;97:4227–32.
15. Mao L, El-Naggar AK, Fan YH, et al. Telomerase activity in head and neck squamous cell carcinoma and adjacent tissues. Cancer Res 1996;56:5600–4.
16. Califano J, Ahrendt SA, Meininger G, et al. Detection of telomerase activity in oral rinses from head and neck squamous cell carcinoma patients. Cancer Res 1996;56:5720–2.
17. Califano J, van der Riet P, Westra W, et al. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res 1996; 56:2488–92.
18. Reed AL, Califano J, Cairns P, et al. High frequency of p16 (CDKN2/MTS-1/INK4A) inactivation in head and neck squamous cell carcinoma. Cancer Res 1996;56:3630–3.
19. Liggett WH Jr, Sewell DA, Rocco J, et al. p16 and p16 beta are potent growth suppressors of head and neck squamous carcinoma cells in vitro. Cancer Res 1996;56:4119–23.
20. Okami K, Reed AL, Cairns P, et al. Cyclin D1 amplification is independent of p16 inactivation in head and neck squamous cell carcinoma. Oncogene 1999; 18:3541–5.
21. Lerman MI, Minna JD. The 630-kb lung cancer homozygous deletion region on human chromosome 3p21.3: identification and evaluation of the resident candidate tumor suppressor genes. The International Lung Cancer Chromosome 3p21.3 Tumor Suppressor Gene Consortium. Cancer Res 2000;60:6116–33.
22. Hibi K, Trink B, Patturajan M, et al. AIS is an oncogene amplified in squamous cell carcinoma. Proc Natl Acad Sci U S A 2000;97:5462–7.
23. Sidransky D, Hollstein M. Clinical implications of the p53 gene. Annu Rev Med 1996;47:285–301.
24. Shah SI, Yip L, Greenberg B, et al. Two distinct regions of loss on chromosome 4q in primary head and neck squamous cell carcinoma. Arch Otolaryngol Head Neck Surg 2000;126:1073–6.
25. Scholnick SB, Haughey BH, Sunwoo JB, et al. Chromosome 8 allelic loss and the outcome of patients with squamous cell carcinoma of the supraglottic larynx. J Natl Cancer Inst 1996;88:1676–82
26. Pai SI, Wu GS, Ozoren N, et al. Rare loss-of-function mutation of a death receptor gene in head and neck cancer. Cancer Res 1998;58:3513–8.
27. Okami K, Wu L, Riggins G, et al. Analysis of PTEN/MMAC1 alterations in aerodigestive tract turmors. Cancer Res 1998;58:509–11.
28. Truelson JM, Fisher SG, Beals TE, et al. DNA content and histologic growth pattern correlate with prognosis in patients with advanced squamous cell carci-

noma of the larynx. The Department of Veterans Affairs Cooperative Laryngeal Cancer Study Group. Cancer 1992;70:56–62.
29. Folkman J. Fighting cancer by attacking its blood supply. Sci Am 1996;275:150–4.
30. Chen Z, Smith CW, Kiel D, Van Waes C. Metastatic variants derived following in vivo tumor progression of an in vitro transformed squamous cell carcinoma line acquire a differential growth advantage requiring tumor-host interaction. Clin Exp Metastasis 1997;15:527–37.
31. Young MR, Wright MA, Lozano Y, et al. Increased recurrence and metastasis in patients whose primary head and neck squamous cell carcinomas secreted granulocyte-macrophage colony-stimulating factor and contained CD34+ natural suppressor cells. Int J Cancer 1997;74:69–74.
32. Young MR. Chemokines and cancer. Arch Pathol Lab Med 2000;124:642.
33. Pekarek LA, Starr BA, Toledano AY, Schreiber H. Inhibition of tumor growth by elimination of granulocytes. J Exp Med 1995;181:435–40.
34. Itoh T, Tanioka M, Matsuda H. Experimental metastasis is suppressed in MMP-9-deficient mice. Clin Exp Metastasis 1999;17:177–81.
35. Tamura M, Arakaki N, Tsubouchi H, et al. Enhancement of human hepatocyte growth factor production by interleukin-1 alpha and -1 beta and tumor necrosis factor-alpha by fibroblasts in culture. J Biol Chem 1993;268:8140–5.
36. Van Waes C, Surh DM, Chen Z, et al. Increase in suprabasilar integrin adhesion molecule expression in human epidermal neoplasms accompanies increased proliferation occurring with immortalization and tumor progression. Cancer Res 1995;55: 5434–44.
37. Wolf GT, Carey TE, Schmaltz SP, et al. Altered antigen expression predicts outcome in squamous cell carcinoma of the head and neck. J Natl Cancer Inst 1990;82:1566–72.
38. Rabinovitz I, Mercurio AM. The integrin alpha 6 beta 4 and the biology of carcinoma. Biochem Cell Biol 1996;74:811–21.
39. Van Waes C, Surh DM, Chen Z, Carey TE. Inhibition of integrin mediated cell adhesion of human head and neck squamous cell carcinoma to extracellular matrix laminin by monoclonal antibodies. Int J Oncol 1997;11:457–64.
40. Rosenthal EL, Johnson TM, Allen ED, et al. Role of the plasminogen activator and matrix metalloproteinase systems in epidermal growth factor- and scatter factor-stimulated invasion of carcinoma cells. Cancer Res 1998;58:5221–30.
41. Hanzawa M, Shindoh M, Higashino F, et al. Hepatocyte growth factor upregulates E1AF that induces oral squamous cell carcinoma cell invasion by activating matrix metalloproteinase genes. Carcinogenesis 2000;21:1079–85.
42. O-charoenrat P, Modjtahedi H, Rhys-Evans P, et al. Epidermal growth factor-like ligands differentially up-regulate matrix metalloproteinase 9 in head and neck squamous carcinoma cells. Cancer Res 2000;60: 1121–8.
43. Coussens LM, Tinkle CL, Hanahan D, Werb Z. MMP-9 supplied by bone marrow-derived cells contributes to skin carcinogenesis. Cell 2000;103:481–90.
44. Chen Z, Malhotra PS, Thomas GR, et al. Expression of proinflammatory and proangiogenic cytokines in patients with head and neck cancer. Clin Cancer Res 1999;5:1369–79.
45. Loukinova E, Dong G, Enamorado-Ayalya I, et al. Growth regulated oncogene-alpha expression by murine squamous cell carcinoma promotes tumor growth, metastasis, leukocyte infiltration and angiogenesis by a host CXC receptor-2 dependent mechanism. Oncogene 2000;19:3477–86.
46. Chen Z, Colon I, Ortiz N, et al. Effects of interleukin-1alpha, interleukin-1 receptor antagonist, and neutralizing antibody on proinflammatory cytokine expression by human squamous cell carcinoma lines. Cancer Res 1998;58:3668–76.
47. Hong SH, Ondrey FG, Avis IM, et al. Cyclooxygenase regulates human oropharyngeal carcinomas via the proinflammatory cytokine IL-6: a general role for inflammation? FASEB J 2000;14:1499–507.
48. Kitadai Y, Takahashi Y, Haruma K, et al. Transfection of interleukin-8 increases angiogenesis and tumorigenesis of human gastric carcinoma cells in nude mice. Br J Cancer 1999;81:647–53.
49. Van Waes C, Enamorado-Ayala I, Hecht D, et al. Effects of the novel alphav integrin antagonist SM256 and cis-platinum on growth of murine squamous cell carcinoma PAM LY8. Int J Oncol 2000;16:1189–95.
50. Ondrey FG, Dong G, Sunwoo J, et al. Constitutive activation of transcription factors NF-(kappa) B, AP-1, and NF-IL6 in human head and neck squamous cell carcinoma cell lines that express proinflammatory and pro-angiogenic cytokines. Mol Carcinog 1999;26:119–29.
51. Duffey DC, Chen Z, Dong G, et al. Expression of a dominant-negative mutant inhibitor-κBα of nuclear factor-κB in human head and neck squamous cell

carcinoma inhibits survival, proinflammatory cytokine expression, and tumor growth in vivo. Cancer Res 1999;59:3468–74.
52. Duffey DC, Crowl-Bancroft CV, Chen Z, et al. Inhibition of transcription factor nuclear factor-κB by a mutant inhibitor-κBα attenuates resistance of human head and neck squamous cell carcinoma to TNF-α caspase-mediated cell death. Br J Cancer 2000;83:1367–74.
53. Kato T, Duffey DC, Ondrey FG, et al. Cisplatin and radiation sensitivity in human head and neck squamous carcinomas are independently modulated by glutathione and transcription factor NF-kappaB. Head Neck 2000;22:748–59.
54. Koch WM, Brennan JA, Zahurak M, et al. p53 mutation and locoregional treatment failure in head and neck squamous cell carcinoma. J Natl Cancer Inst 1996;88:1580–6.
55. Nawroz H, Koch W, Anker P, et al. Microsatellite alterations in serum DNA of head and neck cancer patients. Nat Med 1996;2:1035–7.
56. Sanchez-Cespedes M, Esteller M, Wu L, et al. Gene promoter hypermethylation in tumors and serum of head and neck cancer patients. Cancer Res 2000; 60:892–5.
57. Clayman GL. The current status of gene therapy. Semin Oncol 2000;27:39–43.
58. Saleh MN, Raisch KP, Stackhouse MA, et al. Combined modality therapy of A431 human epidermoid cancer using anti-EGFR antibody C225 and radiation. Cancer Biother Radiopharmacol 1999;14: 451–63.

CHAPTER 57

Chemotherapy for Head and Neck Cancer

Bruce Brockstein, MD, Janardan D. Khandekar, MD

Head and neck cancer (HNC) accounts for only 5% (or 40,000 patients) of malignancies in the United States, but the associated pronounced cosmetic and functional morbidities resulting from both the disease and its treatment heighten its relative importance. As a result, these tumors require the multidisciplinary care and expertise of head and neck surgeons, radiation and medical oncologists, prosthodontists, speech and swallowing therapists, pathologists, radiologists, dietitians, physiatrists, nurses, and social workers.

Multimodality care has led to clear improvements in the outcome for HNC patients. Chemotherapy has assumed an increasing role in HNC in both the curative and palliative settings. In just the last decade, it has been demonstrated that larynx preservation is possible in up to two-thirds of stage III and IV larynx and hypopharynx patients who receive chemotherapy prior to radiation therapy (RT).[1,2] Concomitant chemotherapy and RT (CRT) have recently been definitively shown to improve survival compared with RT alone for patients with unresectable HNC.[3–5]

Chemotherapy plays a diverse role in the many diseases for which it is used. The ultimate objective of chemotherapy is to cure the patient of a malignant disease. In practice, cure has been realized for only a few types of metastatic malignancies, and most patients with metastatic solid tumors cannot be treated with curative intent. Chemotherapy as a single modality has the potential to cure patients with testicular cancer, small cell lung cancer, ovarian cancer, lymphoma, leukemia, and sarcomas of children or young adults. In most disseminated solid tumors, cure is not yet possible, although significant palliation and lengthening of survival can be achieved.

The use of chemotherapy in treating malignancies aims at the eradication of systemic cancer or the increase of locoregional control when used with surgery or RT. Patients are treated with chemotherapy for either "macroscopic" or "microscopic" metastases or as an adjunct in the management of primary, localized tumors. In the setting of macroscopic metastases ("metastatic disease"), patients have clinical or radiologic evidence of tumor spread. Microscopic metastasis refers to clinically unrecognizable small metastatic tumor deposits, which, if untreated, will become macroscopic. It is in this setting that *adjuvant* or *neoadjuvant* chemotherapy is used.

Traditionally, chemotherapy had been used only after the failure of irradiation and/or surgery to control head and neck tumors. These local therapeutic modalities, however, fail to cure a sizeable percentage of patients. Although patients with limited disease (T1–2, N0, M0) can be treated effectively with local modalities such as surgery and radiation, with 60 to 90% cure rates, over 70% of patients with extensive disease (T3–4, N1–3, M0) have recurrence at 2 years after these treatments. This formed the impetus for clinical trials of chemotherapy in the primary treatment of HNC in the 1970s to 1990s that have demonstrated that chemotherapy is an important adjunct to surgery and/or radiation.

This chapter reviews the principles of chemotherapy and describes its uses in HNC. The understanding of molecular biology and genetics is exploding and is beginning to have an impact on therapy of HNC. These topics are also briefly touched on.

CELL CYCLE

Malignant tissues are largely made up of dividing cells, which synthesize deoxyribonucleic acid (DNA) at some point in their life cycle, and nondividing

cells. The dividing cells pass through four different phases (Figure 57–1). Mitosis occupies a discrete phase (M) of the cell cycle. Following the division, the cell enters the G_1 phase. This phase is then followed by DNA synthesis, which is termed the S phase. After completion of the DNA synthesis, the cell then enters a phase of apparent rest (G_2) before the initiation of the next mitosis.

In most dividing cells, the periods for S, G_2, and M phases are of relatively constant duration. Variation of the length of the cell cycle generally occurs in the G_1 phase. When the proliferating cells stop dividing, they do so in the G_1 phase. This phase seems to hold the key to the proliferative activity of the tissue. When the cell-cycle time (time taken for a proliferative cell to move from one mitosis to another) is short, and thus the proliferative activity is high, the G_1 phase is of brief duration. When proliferative activity is slow, the G_1 phase is long.

It is not clear at present how a mammalian eukaryotic cell decides to leave the G_1 phase and to start DNA synthesis. Based on the differences in ribonucleic acid (RNA) content, two distinct subcompartments, G_{1A} and G_{1B}, have been recognized. It is postulated that, after the mitosis, cells reside in the low RNA G_{1A} compartment and that an increase in RNA above this critical level is required for G_1 cells to be able to initiate DNA replication. The G_1 cells with high RNA values are in G_{1B}. During the S phase, biosynthesis of enzymes essential for pyrimidine and nucleic acids is accentuated, and the DNA content of the cell doubles. The phase of DNA synthesis is then complete, and the cell enters G_2 phase. None of the phases is quiescent, and RNA and protein synthesis occurs throughout each of them.

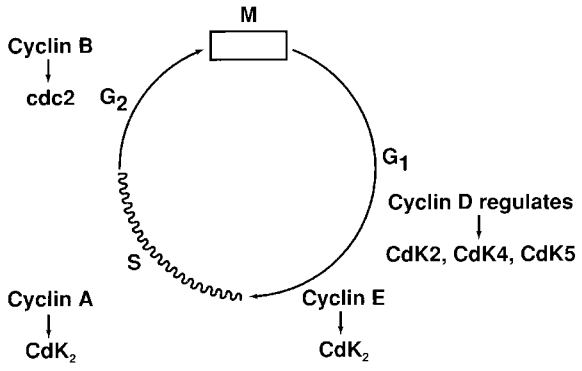


FIGURE 57–1. The cell-cycle phases: M = mitosis; G_1 = preceding DNA synthesis; S = DNA synthesis; G_2 = premitotic resting.

The molecular mechanisms controlling cell cycle are being rapidly elucidated. Cell-cycle control is based on two key families of proteins. The first is the family of cyclin-dependent protein kinases (CdK for short), which phosphorylates selected proteins. The second is a family of specialized activating proteins called cyclins that bind to the CdK molecules and control their ability to phosphorylate appropriate target proteins. Thus, there are two checkpoints in the cell cycle, namely at G_1/S and G_2/M. These checkpoints are controlled by various genes such as the retinoblastoma and p53 genes.

In addition to differences in the biochemical mechanism of action, discussed later, antineoplastic agents also differ in the point in the cell cycle at which they act. The drugs can be grouped as follows on the basis of their effects on phases of the cell cycle[6–8]:

1. *S phase specific.* These are the agents that inhibit the DNA synthesis that occurs in the S phase of the cell cycle. Drugs such as methotrexate (MTX) inhibit not only DNA but also RNA and protein synthesis and are called S phase specific.
2. *Non S specific.* Alkylating agents and antitumor antibiotics exert a direct effect on DNA; thus, their activities are not dependent on the phase of the cell cycle.
3. *Antimitotics.* Plant alkaloids such as vincristine arrest mitosis during metaphase.

Studies of the cell kinetics indicate that the often-held belief that all malignant tumors are rapidly proliferating is not necessarily true. The mean doubling time of human tumors (time taken by a tumor to double its volume) can be as short as 1 week in Burkitt's tumor or as long as several years in some adenocarcinomas.

Malignant tumor formation occurs when any of the following three events occur together or separately: (1) the growth fraction, defined as a proportion of tumor cells in the mitotic cycle at a given time, may become large; (2) the cell-cycle time may become shorter (on average, the cell-cycle time of tumors is 2 days, which is not faster than normal cells); (3) normally, renewal of normal cells is accompanied by cell loss, an active process known as "programmed" cell death or apoptosis. When this cell loss is reduced, tumor formation can occur. This

seems to be the major factor in the production of most tumors.

Epidermoid neoplasms have about 35% growth fraction (dividing cells) in contrast to 3 to 5% in breast cancer and virtually 100% in embryonal tumors. Because most of the chemotherapeutic agents act on these dividing cells, epidermoid cancers would be intermediate in their response to drugs. Recent work suggests that, at the metastatic site, a tumor may still be proliferating rapidly, whereas the parent neoplasm has a low growth fraction. This suggests that chemotherapy given to these patients may still be useful. Methods for measurement of cell kinetic parameters in patients are cumbersome, however. The results of therapy based on these considerations are also at present not much different from those with conventional therapy. Nevertheless, newer advances in this field may considerably alter the present practice of chemotherapy.

MODE OF ACTION OF ANTINEOPLASTIC AGENTS

Traditional Chemotherapy Drugs

The DNA synthesized during the S phase of the cell cycle acts as a selective template for the production of specific forms of messenger, ribosomal, or transfer RNA. The specific sequences of the messenger RNA determine which enzymes will be synthesized in the cell. The enzymes, in turn, are responsible for the structure, function, metabolism, and proliferative rate of the cells.

Most of the clinically useful antineoplastic drugs appear to work by inhibition of synthesis of either enzymes or substrates essential for the nucleic acid synthesis and/or function. Based on the mechanism of drug action, agents that are useful in HNC can be classified as follows:

1. *Antimetabolites*. These drugs inhibit purine or pyrimidine biosynthesis. Methotrexate, for example, inhibits formation of reduced folates, which are essential for thymidine synthesis.
2. *Drugs that interfere with the structure or function of the DNA molecule*. Alkylating agents such as cyclophosphamide change the structure of DNA, thereby arresting cell replication. On the other hand, antibiotics such as dactinomycin and doxorubicin bind and intercalate with nucleotide sequences of the DNA molecule and thereby block messenger RNA (mRNA) production.
3. *Mitotic inhibitors*. Vinca alkaloids, such as vincristine and vinblastine, and the taxanes paclitaxel and docetaxel arrest cell division by disrupting the microfilaments of the mitotic spindle.

New or Investigational Classes of Chemotherapy

A major focus of drug development in academia, industry, and government is drugs that work via different, nonclassic mechanisms. There is hope, and already proof of principle, that some of the drugs from these new classes will improve outcomes or cure rates in HNC and other cancers:

1. *Monoclonal antibodies and vaccines*. These agents target specific cell antigens. Monoclonal antibodies targeted against a single antigen are already approved by the US Food and Drug Administration (FDA) and are in clinical use. These monoclonal antibodies may be used alone or in combination with chemotherapy or RT. They may be linked to radioactive substrates ("radioimmunoconjugates") or may be "bispecific" to an antigen and another target such as an immune effector cell. Trastuzumab (Herceptin) combined with chemotherapy improves outcome for patients with advanced breast cancer who overexpress Her2, a transmembrane receptor protein. Rituximab (Rituxan), directed against the B-cell antigen CD-20, is approved for use in low-grade lymphomas and is being investigated in aggressive B-cell lymphomas. The epidermal growth factor receptor (EGFR) is overexpressed in most HNCs and is one of the most common molecular abnormalities. Upon ligand binding, the receptor undergoes conformational change and dimerizes with another EGFR molecule or another member of the EGFR family of receptors. This then initiates a cascade of signal transduction that is mitogenic and may be antiapoptotic. Epidermal growth factor receptor has become a target for therapy in HNC and other cancers. C225 is a chimeric human-mouse monoclonal antibody directed against EGFR, which is currently in phase III trials. The antibody binds to EGFR, preventing binding of the natural ligands

so that activation does not occur.[9] The antibody has antitumor effects as well as radio-sensitizing effects.[10] The antibody is currently in phase III trials.

2. *Signal transduction modulators*. A variety of these drugs are in clinical trials, and one has recently gained FDA approval for the treatment of chronic myelogenous leukemia. STI-571 (Gleevec) is a tyrosine kinase inhibitor that has shown unprecedented activity in refractory chronic myelogenous leukemia and gastrointestinal stromal tumors. This drug is likely to lead the way in the approval of many other drugs in this class. Head and neck cancer will likely also be a target of such therapies, with targets such as EGFR and downsteam molecules.
3. *Angiogenesis inhibitors*. This broad term refers to drugs that target the synthesis or integrity of tumor neovasculature. It is an attractive target for cancer therapy since vascular biosynthesis is much more active in tumors than normal tissue and may thus provide a more rational target for therapy. No drugs from this class are yet approved, but many are in clinical trial.
4. *"Gene therapy/gene transfer."* Head and neck cancer is a prime target for gene therapy. Head and neck cancer trials of the drugs ONYX-015 and RPR/INGN-201 are among the first gene therapy trials in oncology. ONYX-015 uses an adenovirus with the *E1B* gene deleted. This modification renders it unable to replicate in cells with normal p53 but allows for selective replication and cell lysis in p53 mutant cells.[11] This drug is being extensively studied as a single agent and in combination therapies, and via several routes of delivery.[12,13] One interesting study was recently reported with preliminary results. Thirty-seven patients with recurrent HNC received intratumoral injections of p53 in combination with cisplatin and 5-fluorouracil. Almost all patients had prior surgery and RT. The complete response (CR) rate was 27%, and the overall response rate (RR) was 63%. More impressively, none of the 19 injected tumors that responded to therapy had progressed at the time of an early report (5-month median follow-up).[14]

   RPR/INGN 201 (Ad5CMVp53) uses a replication-deficient adenovirus to deliver a wild-type *P53* gene driven by a cytomegalovirus promoter. Early studies in over 200 patients have shown intratumoral injections to be tolerable, with some evidence of activity of this drug. Phase III studies are ongoing.
5. *Photosensitizers*. Photosensitizers such as porfimer sodium (Photofrin) are used in photodynamic therapy. Given systemically, these drugs are preferentially taken up in tumor cells, rendering them more sensitive to light-mediated cell killing than their surrounding tissues. Although not approved specifically for HNC, these drugs have been used for palliation in advanced HNC or the treatment of early-stage HNC.
6. *Inhibitors of invasion and metastasis.* An important step in the invasion and metastasis process is the disruption of the epithelial basement membrane. This step is integrally related to the angiogenesis process. A number of such drugs have been tested or are currently in clinical trial, although none are yet approved.
7. *Antisense oligonucleotides.* Once the gene sequence of a cancer-related gene or oncogene is identified, an exogenous, complementary oligodeoxynucleotide (oligo) can be synthesized against the gene or oncogene mRNA. Oligos of large enough base-pair size can be synthesized to be uniquely complementary to the target mRNA. The antisense oligo, by formation of an mRNA-DNA duplex, could specifically prevent translation of that mRNA into protein. These drugs continue in the developmental and clinical trial phase.
8. *Biologics*. Drugs such as interleukin-2 and interferon-α have an established role in the therapy of malignant melanoma and renal cell carcinoma. The role of interferon in HNC has diminished, but research in this field is continuing.

CHEMOTHERAPY-RELATED ISSUES

General Toxicity of Antineoplastic Agents

Most of the approved chemotherapy agents act on the proliferative fraction of cells. Therefore, an undesirable consequence of these drugs is the damage inflicted on the normal cells undergoing rapid cell division. This correlation between proliferative activity and toxicity is most notable in the hair follicles, skin, gastrointestinal mucosa, and normal bone marrow elements. Clinically, then, damage to these tis-

sues is manifested as alopecia, stomatitis, nausea, vomiting, anemia, leukopenia, and thrombocytopenia. In addition to these effects, specific problems can occur with each drug, and these are discussed below.

Chemotherapy Drug Resistance

Resistance to chemotherapy drugs, either inherent or acquired, remains a major problem in cancer chemotherapy. Resistance to natural product drugs, known as *multidrug resistance*, has been shown to be acquired by the tumor cells to drugs that are structurally and functionally unrelated. Initially, this type of drug resistance was shown to be owing to the presence of a P-glycoprotein encoded by the human multidrug resistance (*MDR*) gene. Subsequently, it was shown that the gene product, P-glycoprotein, is inserted in the plasma membrane and acts as an adenosine triphosphate–driven drug efflux pump. Other proteins, such as multidrug resistance–associated protein, have also been implicated in this drug resistance. Many other mechanisms, such as altered blood delivery of drugs and activation of target enzymes as well as modulation of DNA, may also confer drug resistance.

Organ Preservation

A major goal of combined chemotherapy and radiation is that of decreasing morbidity. The organs in the head and neck area are of paramount importance for normal daily function; therefore, organ preservation has assumed an important role. An important emerging concept not only in the chemotherapy of cancer in general but also in all of cancer treatment is the impact that the disease and/or treatment may have on the quality of life. Full consideration of organ preservation strategies must consider the maintenance of normal organ anatomy and function, as well as global quality of life. Assessment of function outcome is a major goal in organ preservation trials. Specific issues of organ preservation are discussed below.

SPECIFIC CHEMOTHERAPEUTIC AGENTS

US Food and Drug Administration–approved chemotherapy drugs used, but generally not specifically approved for HNC, are described below. Mechanisms, RRs to the drug as a single agent, and toxicities are described. The definition of RR or clinical benefit in HNC requires special consideration.

Definition of Clinical Benefit

Clinical benefit from systemic chemotherapy in patients with HNC or other cancers can be defined by the *objective RR* in measurable disease, the *time to progression*, the impact on *overall survival*, or the success of *organ preservation strategies*. Objective responses are defined as partial or complete. Partial response is generally defined as a $\geq 50\%$ decrease in the sum of the product of the perpendicular diameters of all tumors. Complete response is defined as 100% shrinkage. Most commonly, objective RRs are used, although they may not be the best assessment of clinical benefit. When comparing objective RRs between various regimens in HNC, certain caveats apply. Objective RRs may vary greatly between studies, and a number of factors can account for these discrepancies. Prior chemotherapy can greatly alter (decrease) the RR to a subsequent regimen. Neoadjuvant chemotherapy for patients with locoregional disease generally results in two-fold higher RRs than those achieved in patients with recurrent and metastatic disease who have received no prior chemotherapy. Response rates in multi-institutional phase III trials are generally lower than those reported in single-institution or phase II trials. The definition of response is variable among studies; some use computed tomographic (CT) scans for evaluation, whereas others rely on clinical examination for tumor measurement. Defining objective response, particularly for patients who have had prior locoregional therapy, is difficult in HNC. Head and neck cancers tend to form less well-defined, discrete lesions than are seen in other sites (eg, lung, liver, or brain), making precise measurement difficult and introducing a variable in response reporting. Despite these difficulties, most clinicians use the objective RR as the primary end point in metastatic disease.

The reported objective response of various drugs in head and neck cancer is shown in Table 57–1.

Multiple other drugs (eg, mitomycin, cyclophosphamide, doxorubicin, vinblastine, hydroxyurea) have shown RRs in smaller studies of 10 to 40% but are rarely used in HNC. Newer, approved

TABLE 57–1. Systemic Single-Agent Chemotherapy*

Drugs	*Response Rate (%)*
Cisplatin	19
Carboplatin	25
Paclitaxel	37
Docetaxel	30
Methotrexate	22
5-Fluorouracil	14
Ifosfamide	24
Bleomycin	28

*Combined results of clinical trials (phases II and III) of single-agent chemotherapy for head and neck cancer. Trials took place in different settings and in different eras, using different evaluation techniques. Therefore, no direct comparisons regarding efficacy can be made from this table.

drugs such as irinotecan-11, gemcitabine, and vinorelbine have activity in HNC, but their role is undefined.

Specific Drugs

Methotrexate *Methotrexate* was the first active chemotherapeutic agent for patients with HNC. The biochemical effects of MTX have been extensively studied. Methotrexate acts primarily by inhibiting the activity of the enzyme dihydrofolate reductase, which converts dihydrofolate to tetrahydrofolate. The reduced folates are required for the formation of thymidylate, which then becomes incorporated into DNA. Other evidence suggests that the drug also inhibits purine biosynthesis.

Methotrexate has been administered by and studied in multiple doses, routes (oral, intravenous, intra-arterial), and schedules. In multi-institutional phase III studies using standard-dose MTX (usually 40 mg/m^2 per week) as the control arm, objective RRs ranged from 12 to 35% (average 22%).[15–20] The disparity in reported RRs of this one drug is an example of the difficulty in comparing RRs across trials in HNC. The superiority of high-dose MTX with leucovorin rescue versus standard dose has not been demonstrated. Response durations to MTX are comparatively short-lived because of development of resistance to the drug. There are several mechanisms of resistance to MTX. These include decreased transport of the drug across cell membranes and amplification of the gene for dihydrofolate reductase (gene amplification results in more copies of the gene and thus the enzyme, and increased enzyme level results in decreased inhibition by MTX). Additionally, decreased formulation of polyglutamates may result in decreased inhibition of DNA synthesis. The main side effects of MTX, when given weekly or biweekly, include bone marrow suppression, stomatitis, mucositis, and sometimes nausea, vomiting, and diarrhea.

Randomized trials have shown that, compared with single-agent cisplatin, single-agent MTX has a lower RR (12 versus 28% and 24 versus 29%), a shorter overall survival, but a better toxicity profile.[15–20] Because MTX is easy to administer and has manageable toxicities, some experts still consider MTX the standard single agent for comparison in HNC trials.

Cisplatin Cisplatin is an alkylating agent that exerts its effect by the formation of intrastrand and interstrand DNA cross-links. Cisplatin has been extensively studied in a variety of doses, routes, and schedules for patients with advanced HNC. In large phase II and multicenter phase III trials, the single-agent RR is about 20%.[16,21,22] In randomized trials, single-agent cisplatin appears to produce a superior RR when compared to MTX. Survival was also significantly better in one of these studies. As a single agent, it may be superior to single-agent bleomycin and appears to be equivalent to 5-fluorouracil (5-FU) alone. Cisplatin has not been directly compared to paclitaxel (Taxol) or docetaxel (Taxotere).

Higher than standard doses of cisplatin have been studied in advanced HNC. However, the incremental benefit may not be sufficient to warrant the added toxicity. If given without forced hydration and diuresis, nephrotoxicity is extremely common. Severe nausea and vomiting are at least partly ameliorated by the serotonin receptor antagonist antiemetics such as ondansetron or granisetron. After multiple doses, peripheral neuropathy and ototoxicity become increasingly common.

Carboplatin Carboplatin is an analogue of cisplatin that causes less nephrotoxicity, ototoxicity, and neuropathy, although it is more myelosuppressive. The single-agent RR of carboplatin is 25%, similar to that of cisplatin.[23] Although single-agent randomized comparative studies are lacking, extrap-

olation from combination studies combining each agent with 5-FU suggests a marginal superiority of cisplatin over carboplatin.[24,25]

Taxanes In advanced squamous cell HNC, the taxanes, paclitaxel and docetaxel, have emerged as perhaps the most active of all single agents. Both drugs act by stabilization of microtubules against depolymerization, blocking mitosis. Paclitaxel was originally isolated and supplied from the bark of the Pacific yew *Taxus brevifolia*; however, both drugs are now manufactured using a semisynthetic process from natural precursors.

Paclitaxel In two phase II studies of patients with advanced HNC, single-agent paclitaxel was associated with an RR between 35 and 40%.[26,27] The major toxicity is hematologic. An ongoing EORTC (European Organization for Research and Treatment of Cancer) Head and Neck Cancer Cooperative Group trial is comparing two schedules of paclitaxel to bolus MTX. Preliminary analysis shows a lower RR for paclitaxel compared with that seen in the earlier studies.

Docetaxel Docetaxel has been evaluated in four phase II trials of 144 patients, about one half of whom had received prior chemotherapy. The RR ranges from 21 to 42%, with a median response duration of 11 weeks in one study.[28–31] There are no randomized trials directly comparing docetaxel with paclitaxel for patients with advanced HNC.

Side effects of the taxanes include myelosuppression, peripheral neuropathy, nail damage, skin rash, and fluid retention (docetaxel).

5-Fluorouracil 5-Fluorouracil and its deoxynucleoside, floxuridine, are both prodrugs that require intracellular metabolism for cytotoxicity. Three different targets are known to be affected by these fluorpyridimines. The metabolite of 5-FU, known as 5-fluorodeoxyuridine monophosphate, forms a stable complex with thymidylate synthetase, inhibiting de novo thymidylate synthesis and thus causing cell death by thymidine starvation. Further, 5-FU may induce DNA damage by two distinct mechanisms: fraudulently being incorporated into DNA, which causes DNA fragmentation and cell death, and inhibiting DNA repair. The degree of inhibition of thymidylate synthetase can be augmented by exogenous folates. This forms the basis for using leucovorin with 5-FU.

5-Fluorouracil, despite its extensive use as a radiosensitizer and widespread experience in combination with cisplatin (see below), has not been well studied as a single agent for HNC. An older study reported a 15% objective RR in patients with HNC. In a more contemporary randomized phase II study, 5-FU alone (1,000 mg/m^2 per day $\times$ 4 days) had a RR of 13%, compared with 17% for cisplatin and 32% for cisplatin plus 5-FU.[22] The side effects of 5-FU are schedule dependent but include nausea, vomiting, myelosuppression, diarrhea, and stomatitis.

Ftorafur and Uracil UFT, a 1:4 molar mixture of ftorafur (which is converted to 5-FU in the liver) and uracil (which inhibits the catabolism of 5-FU), is an oral drug that has been predominantly developed in Japan and Spain. In phase II trials of patients with advanced HNC, the overall RR exceeded 30%.[32]

In the United States, several other oral fluorinated pyrimidines, including Orzel, which combines UFT and oral leucovorin, and capecitabine, a 5-FU prodrug, are undergoing investigation in HNC. In a preliminary report of 29 patients with advanced HNC, the overall RR with single-agent Orzel was 28%, including three CRs.

Ifosfamide Ifosfamide is a prodrug that requires hepatic activation to an active alkylating agent. There are at least eight trials evaluating ifosfamide as a single agent in advanced HNC. In these reports, the objective RR has ranged from 4 to 53%.[33,34] Some but not all of this difference may be accounted for by a dose-response relationship. One of the highest RRs, 43%, was from a study using 17.5 g/m^2 ifosfamide; in contrast, the lowest RR (4.3%) was noted with a lower dose (1.5 g/m^2 per day for 5 days). However, this same regimen was used in a third trial that reported the highest RR of 53%. The side effects of ifosfamide include nausea, vomiting, myelosuppression, hemorrhagic cystitis, a generally reversible central nervous system toxicity, and proximal renal tubule damage.

Bleomycin Bleomycin has modest activity when used as a single agent for HNC. In multicenter phase II trials, objective RRs range from 6 to 56% (average 28%). [35,36] Despite minimal myelosuppression with this agent, nonhematologic toxicity (eg, pneumonitis, stomatitis, and skin changes) has precluded its widespread use in patients with HNC.

Cyclophosphamide, Doxorubicin, and Hydroxyurea Cyclophosphamide, doxorubicin, and hydroxyurea have been studied in a wide variety of doses and schedules. Many of these reports predated CT scanning, complicating the interpretation of RRs. In five trials, the objective RR with single-agent cyclophosphamide was 36%. The single-agent RR has ranged from 13 to 23% with doxorubicin and from 32 to 39% with hydroxyurea. The main utility of hydroxyurea in HNC has been as a radiosensitizer. Despite these "high" RRs from older studies, these drugs are probably less active than those above and are not commonly used for HNC.[37,38] Newer drugs such as vinorelbine, gemcitabine, and irinotecan have shown activity against HNC in small studies, and their roles continue to be defined.

COMBINATION CHEMOTHERAPY

For most cancers, drug combinations give better results than single agents. These results may be explained on the basis of a number of theoretical and practical considerations. Drugs with multiple mechanisms of action may kill cells resistant to only one drug or may prevent the emergence of resistant tumor cells. Similarly, by acting on different phases of the cell cycle, the drugs may have additive or synergistic effects. Finally, with the combination of drugs with different spectrums of clinical toxicity, administration of full or nearly full doses of each of the active agents is possible.

In HNC, most combinations have been based on cisplatin or methotrexate. At least six large randomized studies comparing cisplatin-based com-

TABLE 57–2. Randomized Trials of Cisplatin-Based Combination Chemotherapy Compared with Single-Agent Chemotherapy in Recurrent or Metastatic Head and Neck Cancer

Study/Reference	*Regimen*	*No. of Patients*	*RR (%)*	*Median Survival (mo)*	*Comments*
SWOG[18]	Cis/5-FU	87	32	7	$p < .001$ for Cis/5-FU vs MTX
	Carbo/5-FU	86	21	5	$p = .05$ for Carbo/5-FU vs MTX
	MTX	88	10	6	
Stanford University[22]	Cis/5-FU	79	32	6	$p = .035, .005$ for CF vs Cis, 5-FU
	Cis	83	17	5	
	5-FU	83	13	6	
Liverpool*[16]	Cis/5-FU	50	24	5	MTX given only every other week
	Cis/MTX	50	22	5	No difference in RR
	Cis	50	28	7	
	MTX	50	12	4	
EORTC[21]	Cis/5-FU	116	34	6	$p = .001, .003$ for CABO, PF vs Cis
	Cis/MTX/Bleo/VCR (CABO)	127	37	8	Survival difference not significant
	Cis	122	16	5	
SECSG	Cis/Vinbl/Bleo	92	24	6	p = NS
	MTX	98	16	6	
ECOG	Cis/MTX/Bleo	80	48	6	p = NS
	MTX	83	35	6	

SWOG = Southwest Oncology Group; Cis = cisplatin; 5-FU = 5-fluorouracil; Carbo = carboplatin; MTX = methotrexate; CF = cisplatin and 5-fluorouracil; RR = response rate; EORTC = European Organization for Research and Treatment of Cancer; Bleo = bleomycin; VCR = vincristine; PF = cisplatin and 5-fluorouracil; SECSG = Southeastern Cancer Study Group; Vinbl = vinblastine; ECOG = Eastern Cooperative Oncology Group; NS = not significant.
*Minority of patients on this trial had no prior therapy and no metastases.

bination regimens to single-agent chemotherapy have demonstrated superior RRs for combination therapy compared with monotherapy in patients with metastatic or recurrent disease (Table 57–2).[16–18,20–22,39] However, survival duration, at least within the power of these studies, has not been improved by the use of multidrug regimens in the palliative setting. These studies attempted to increase RRs by combining, to achieve additive or synergistic activity, drugs with proven single-agent activity, such as cisplatin, 5-FU, MTX, and bleomycin. Cisplatin combined with 5-FU is one combination that is synergistic in vitro, and, clinically, cisplatin followed by a 4- to 5-day continuous intravenous infusion of 5-FU is a very active combination. For patients with recurrent disease, it has reproducibly yielded RRs ranging from 30 to 40%.[15] In the neoadjuvant (prior to chemotherapy or radiation) setting with locally advanced, nonmetastatic disease, RRs are about 80%, with 10 to 40% CR rates.[1,2,40,41]

The combination of cisplatin and 5-FU has been compared with each of these drugs delivered separately as single agents in a three-arm randomized trial. Although the RR of the combination (32%) was significantly higher than that of cisplatin (17%) or 5-FU alone (13%), there was no significant difference in the median survival, 5 to 6 months, for all groups.[22] A randomized trial by the Southwest Oncology Group (SWOG) compared cisplatin and 5-FU with a combination of carboplatin and 5-FU (postulated to be equally active but less toxic) and single-agent MTX.[27] Both cisplatin and carboplatin combined with infusional 5-FU resulted in an improved RR as compared with methotrexate alone. Both combinations were more toxic, and survival was not affected.

The taxanes are now commonly used in combination regimens.[42–47] A recently completed trial compared cisplatin plus 5-FU versus cisplatin plus paclitaxel for recurrent and metastatic HNC. Both regimens were essentially equivalent in RR and survival, although cisplatin with paclitaxel may be slightly less toxic. Several other three- or four-drug combination regimens containing taxanes have shown high RRs and continue under development.

The following conclusions can be made about combination chemotherapy:

1. Combinations produce statistically significantly higher RRs than the single agents with which they have been directly compared.
2. Cisplatin and infusional 5-FU and taxane combinations generally produce higher RRs than other combinations.
3. In no direct comparison has survival been meaningfully increased when comparing combinations with single agents. No well-conducted studies have been performed to ascertain a survival benefit of chemotherapy versus best supportive care.
4. The toxicities of cisplatin and infusional 5-FU (especially nausea and vomiting) are significantly higher than with single agents.

Clinical research now focuses on the identification of new agents and combinations with activity in HNC. In particular, the new agents paclitaxel, docetaxel, irinotecan, topotecan, and gemcitabine have only begun to be tested in combinations for HNC. Drugs with novel mechanisms, such as monoclonal antibodies, gene transfer agents, angiogenesis inhibitors, and other inhibitors of invasion and metastasis, are currently under investigation in combination. The role of retinoids, selenium, and other molecules in the reversal of premalignant lesions and prevention of primary or second malignancies is also under active investigation.[48]

In summary, chemotherapy for recurrent or metastatic HNC may be palliative in some patients, although the impact on survival duration is small. Methotrexate in weekly low doses, cisplatin or carboplatin, infusional 5-FU, paclitaxel, and docetaxel remain the most active single agents, producing RRs of approximately 20 to 30% that last for 2 to 6 months. Combination chemotherapy, particularly the combination of cisplatin and 5-FU and cisplatin (or carboplatin) and paclitaxel, produce higher RRs, although long-term survival is rarely achieved. Because cure is unlikely to occur as a result of any of these drugs or combinations in the previously treated patient population and in view of the poor outcome with standard single agents and combinations, whenever possible, patients should be treated on a clinical trial.

COMBINED MODALITY THERAPY

Neoadjuvant Chemotherapy

In the 1980s and 1990s, enormous enthusiasm was generated for neoadjuvant chemotherapy for HNC. Through early clinical trials, it quickly became clear

that tumors treated with chemotherapy prior to surgery or radiotherapy usually shrank substantially.[49] This concept was, therefore, extensively studied on the following theoretical grounds: untreated tumors with intact vascular supply have better drug delivery, as compared with tumors previously treated with radiation or surgery in which normal vasculature and drug delivery are disrupted; regression of primary and lymph nodal tumors may permit subsequent local therapies such as surgery and radiation to have enhanced efficacy; the extent of surgical resections can be reduced to more conservative procedures; reducing the bulk of initially unresectable lesions may allow eventual surgical removal; eradication of subclinical micrometastasis at the earliest possible time can reduce the risk of systemic relapse; tolerance of intensive combination chemotherapy is better in patients with better performance and nutritional status (which generally worsen after chemotherapy and RT); a subpopulation of patients can be identified who might further benefit from additional adjuvant chemotherapy following local treatments.

The outcome of two decades of clinical research in neoadjuvant chemotherapy in HNC was disappointing and can be summarized as follows:

1. Overall RRs (at least 50% tumor shrinkage) exceeding 80% are frequently achieved.
2. Complete RRs usually range from 20 to 40%, with most trials reporting approximately 30%. Many of the clinical CRs are confirmed histologically at surgery.
3. Toxicity is usually moderate to severe, but the administration of subsequent standard local therapy is not compromised (although a small percentage of patients with CRs will refuse subsequent planned local therapy).
4. The rate of development of distant metastases is decreased.
5. Patients achieving CRs have a better prognosis, particularly if the responses are histologically confirmed. This observation, though, does not establish that chemotherapy is responsible for better results as chemotherapy-responsive patients may have an intrinsically favorable course.
6. *Organ preservation is made possible, at least in laryngeal and hypopharyngeal carcinoma.*
7. *Survival, assessed in randomized studies, is generally not improved (with the possible exception of patients receiving neoadjuvant cisplatin plus 5-FU); thus, the use of* neoadjuvant *chemotherapy today is very limited.*

This last observation was made by analyzing the randomized trials comparing standard local therapy given with or without neoadjuvant chemotherapy. A large number of such randomized trials have been conducted, allowing for conclusions to be drawn, despite the patient heterogeneity, by analyzing aggregate trials. Most individual trials have been too small or poorly designed to be conclusive. About 10 studies, however, have entered large numbers of patients. The most important of these are summarized in Table 57–3.[1,2,40,41,50–53] In none of these 10 studies was overall survival prolonged. In all 8 studies reporting rates of distant metastases as a site of first failure, the rate of distant metastases was decreased among the patients receiving chemotherapy.[54] Unfortunately, most patients still died of locoregional disease complications; therefore, the decreased rate of distant metastases did not translate into a survival benefit. These negative results have been confirmed by three recent meta-analyses.[3–5] In the individual patient meta-analysis performed by the MACH-NC (Meta-analysis of Chemotherapy on Head and Neck Cancer) Cooperative Group, however, it was noted that in the subset of patients receiving neoadjuvant cisplatin plus 5-FU, there was a significant survival benefit (hazard ratio 0.88, 95% confidence interval [CI] = 0.79 to 0.97).[3] Additionally, since metastases are decreased using neoadjuvant chemotherapy, it is possible with improved locoregional treatment modalities that neoadjuvant chemotherapy may further improve survival rates.

As mentioned, organ preservation may be possible with the use of neoadjuvant chemotherapy. This has been demonstrated in two well-conducted randomized trials. One is the randomized trial conducted by the Veterans Administration Cooperative Study Program. In this study, patients with advanced laryngeal cancer were randomized to standard therapy with surgery and postoperative RT or to three cycles of neoadjuvant cisplatin and 5-FU followed by RT.[1] Response was assessed after two cycles of chemotherapy. Partial and complete responders continued with a third cycle of chemotherapy. Only patients who failed to respond to the first two cycles of chemotherapy or had residual disease after RT proceeded with surgery on the experimental study arm. Two goals were pursued in this study: improved survival and larynx preservation. The 2-year actuar-

TABLE 57–3. Randomized Studies of Neoadjuvant Cisplatin and/or 5-Fluorouracil versus Locoregional Therapy Only

Author Years of Patient Accrual	*Evaluated Patients (N)*	*Chemotherapy*	*Radiation/ Surgery*	*Overall Survival (%) Chemotherapy/ Control*	*Comments*
Schuller et al[50] (SWOG) 1980–1985	158	CDDP, MTX, Bleo, VCR Q 3 wk × 3 cycles	Surgery + RT in all	(5 y) 28/28	Operable stages III/IV CR to chemotherapy 19% Decreased metastases
Laramore et al[40] (Intergroup) 1983–1990	442	CDDP, 5-FU Q 3 wk × 3 cycles	Surgery, then RT	(4 y) 48/44 NS	Stages II–IV operable; all with negative margins Decreased metastases
Veterans Affairs Larynx[1] 1985–1989	325	CDDP, 5-FU Q 3 wk × 3 cycles	Chemotherapy + RT vs surgery + RT	(2 y) 68/68	Stages III, IV surgery in chemotherapy nonresponders Decreased metastases Histologic CR 45% to chemotherapy
Martin[53] 1986–1989	156	CDDP, 5-FU Q 3 wk × 3 cycles	Surgery + RT in all	(5 y) 34/38 $p = .2$	CR to chemotherapy 47%; (abstract only)
Paccagnella et al[41] 1986–1990	237	CDDP, 5-FU Q 3 wk × 4 cycles	Surgery (if resectable = 66 patients) + RT (all)	(2 y) 37/29 (all patients) $p = .2$ (inoperable patients 30/19, $p = .04$])	Stages III/IV 28% resectable; CR to chemotherapy 20%; 8 patients with angina/MI Decreased metastases
Domenge et al[52] 1986–1992	318	CDDP, 5-FU Q 3 wk × 3 cycles	RT	(Median) 5.1/3.3 ($p = .03$)	Oropharynx only
Lefebvre et al[2] 1987–1993	194	CDDP, 5-FU Q 3 wk × 3 cycles	Surgery + RT vs chemotherapy + RT	(3 y) 57/43 (p = NS)	Hypopharynx Decreased metastases
Depondt et al[51] 1988–1991	300	Carboplatin, 5-FU Q 3 wk × 3 cycles	Varied	(3 yr) 52/46 NS	Primary T2–4; CR to chemotherapy 31%; less mutilating surgery with neoadjuvant

SWOG = Southwest Oncology Group; CDDP = *cis*-diamminodichloroplatinum (cisplatin); MTX = methotrexate; Bleo = bleomycin; VCR = vincristine; RT = radiation therapy; CR = complete response; 5-FU = 5-fluorouracil; NS = not significant.

ial rates of overall survival were identical in the two groups at 68%. The most important finding was the high rate of larynx preservation. Sixty-four percent of patients on the chemotherapy arm had their larynges preserved, with a median follow-up of 33 months. Thirty-nine percent of patients remained disease free with an intact larynx. Only two salvage laryngectomies were performed after the first year. A similar

rate of disease-free survival with larynx preservation (28%) was achieved in another study with similar design in hypopharyngeal cancer.[2]

Generally, based on these two studies, all patients with laryngeal or hypopharyngeal carcinoma who would otherwise be required to have total laryngectomy or pharyngolaryngectomy should be offered organ preservation therapy. The protocol used in the Department of Veterans Affairs larynx trial is acceptable at a minimum, although the addition of concomitant chemotherapy to RT is likely to further improve survival and organ preservation. It should be noted that in the MACH-NC meta-analysis, the inclusion of a third small larynx preservation study has created a nonsignificant negative effect of chemotherapy, although a similar outcome continues to exist in the two major studies. An ongoing Intergroup study is currently aimed at determining optimal larynx preservation strategy and compares RT alone versus RT with neoadjuvant or concomitant chemotherapy. It is notable that several new aggressive multiagent neoadjuvant chemotherapy regimens, given prior to RT, are showing promise at improving survival further.

Neoadjuvant chemotherapy has not conclusively improved survival, and its use remains investigational for sites other than the larynx and hypopharynx. Ongoing roles include organ preservation in larynx and hypopharynx cancer and a possible role in the treatment of nasopharyngeal cancer. Finally, the neoadjuvant setting may serve as an appropriate time to test promising new chemotherapy agents or drug combinations.

Adjuvant Chemotherapy

Adjuvant chemotherapy generally refers to the postoperative (or post-RT) use of chemotherapy to reduce the risk of recurrence. This strategy is clearly effective in breast and colon cancer, along with several other malignancies. In HNC, this strategy has been difficult to implement and not very successful and is therefore rarely used today.

Combined Radiation and Chemotherapy

Concomitant, or simultaneous, chemotherapy and RT (CRT), as opposed to neoadjuvant chemotherapy, does have a large and growing role in the treatment of locoregionally advanced HNC. Because HNC manifests clinically predominantly as a *locoregional* disease, concomitant chemoradiotherapy is valuable because it focuses early treatment on the site that determines prognosis. Clinical trials have proven the utility of CRT, which was hypothesized as useful for some of the following reasons: chemotherapy drugs and RT may be active against different tumor cell populations, based on factors such as cell-cycle specificity, pH, or oxygen supply. Cells that are resistant to one modality of treatment may respond to the other. Some chemotherapy drugs may "recruit" cells from the G_0 phase of the cell cycle to the more RT-responsive cell-cycle phases. Chemotherapy may shrink tumors, allowing for improved efficacy of RT by decreasing tumor interstitial pressure and increasing blood flow and oxygen delivery. Early eradication of tumor cells in the radiation field may prevent the emergence of drug or RT resistance; early treatment of microscopic tumor spread outside the RT field may allow for the eradication of microscopic metastases. Chemotherapy may prevent inherent causes of resistance to RT such as tumor cell repopulation between RT fractions and sublethal radiation damage and may inhibit recovery from potentially lethal radiation damage.[54,55]

Although a number of drugs can enhance the efficacy of RT, most also increase the damage to normal tissue. Thus, the balance of efficacy and toxicity with respect to the patient and tumor characteristics is a determining factor for the use of a particular CRT regimen. A large number of chemotherapeutic drugs and combinations have been used in conjunction with radiation.[54–69] A brief overview of the results will be considered.

There are several clinical settings in which CRT may be used, and these will be addressed individually. Unresectable locoregionally advanced HNC is clearly better treated with CRT than RT alone in most settings. In the postoperative setting, several studies suggest that CRT, for high-risk patients, confers an advantage over RT alone.[57,62] For patients with resectable disease who, for medical reasons or refusal of surgery, are to receive nonsurgical treatment, CRT is generally superior to RT alone. Chemoradiotherapy has not been directly compared to surgery and RT for resectable and operable patients; however, comparison with phase II and phase III CRT arms of other studies suggests that appropriately administered intensive CRT may be at least equivalent to surgery plus RT.[56,58,64,69]

TABLE 57–4. Randomized Studies of Concomitant Chemoradiotherapy versus Radiotherapy Alone

Authors/Year	*Patients (N)*	*Chemotherapy (+ RT in All)*	*Overall Survival (%) Chemotherapy/ Control*	*Comments*
Single-agent cisplatin or 5-FU				
Browman et al[60]	175	5-FU infusion	(2 y) 63/50 ($p = .08$)	
Lo et al[65]	151	5-FU bolus	(5 y) 32/14 ($p < .05$)	
Jeremic et al[63]	130	Cisplatin, daily	(5 y) 46/25 ($p = .008$)	Hyperfractionated RT
Bachaud et al[57]	83	Cisplatin, weekly	(5 y) 38/13 ($p < .01$)	
Combination chemotherapy				
Taylor et al[67]	214	Cisplatin, 5-FU	(2.5 y) 41/36	CRT was "split course," control group was neoadjuvant cisplatin/5-FU→RT
Merlano et al[66]	157	Cisplatin, 5-FU	(3 y) 41/23 ($p < .05$)	Chemotherapy alternated with RT
Brizel et al[58]	116	Cisplatin, 5-FU weeks 1, 6	(3 y) 55/34 ($p = .07$)	Hyperfractionated RT
Wendt et al[69]	270	Cisplatin, 5-FU weeks 1, 4, 7	(3 y) 48/24 ($p < .0003$)	Hyperfractionated RT
Adelstein et al[56]	100	Cisplatin, 5-FU weeks 1, 4	(5 y) 50/48	Far less salvage surgery needed in chemotherapy arm
Calais et al[61]	226	Carboplatin, 5-FU, weeks 1, 4, 7	(3 y) 51/31 ($p = .02$)	
Al-Sarraf et al[74]	295	Arm A, RT; Arm B, cisplatin; weeks 1, 4, 7, or Arm C, cisplatin, 5-FU weeks 1, 5, 9	(3 y) 20/37/29 (for RT/cisplatin/ cisplatin, 5-FU) ($p = .016/.13$)	Arm C had 4-week RT break (weeks 5 to 8)

RT = radiotherapy; 5-FU = 5-fluorouracil; CRT = chemotherapy and radiation therapy.
All studies used cisplatin, 5-FU, or the combination of cisplatin (or carboplatin) and 5-FU. Several principles are demonstrated, including the superiority of CRT over RT for unresectable patients and the superiority of CRT (hyperfractionated) over hyperfractionated RT alone.

Representative randomized studies in which cisplatin (or carboplatin) and/or 5-FU were used concomitantly with RT are detailed in Table 57–4.

Unresectable Head and Neck Cancer

There is great variability on the definition of resectability, which is highly dependent on the surgeon's and patient's opinion and preferences. This is, of course, an important factor in interpreting the outcome in trials of patients with "unresectable" HNC. Despite these problems, historically, long-term survival with RT alone for unresectable HNC has been 10 to 30%. Numerous trials have been performed over the last two to three decades in which patients with unresectable

HNC have been randomized to receive RT with or without concomitant chemotherapy. Multiple different drugs have been used, alone or in combination, on various schedules with respect to timing of chemotherapy, dose of radiation, etc. Some general conclusions can be drawn from these studies, particularly from three recent, large-scale meta-analyses.[3–5] In general, CRT using most single-agent chemotherapy drugs studied have produced locoregional control and survival superior to RT alone. Comparisons have not been directly made between the various radiosensitizing chemotherapy drugs. The most frequently used and perhaps best single-agent chemotherapy drugs are 5-FU and cisplatin. Although they have not been directly compared to RT alone, paclitaxel and docetaxel are also commonly used. Several other drugs are effective but less commonly used because of concerns about side effects or inferior efficacy. Many combination regimens have been used, the most effective of which appears to be the combination of cisplatin and 5-FU. The combination of carboplatin and paclitaxel is undergoing rigorous study based on its potential efficacy and relative tolerability.[70] The combination of 5-FU and hydroxyurea has been extensively studied at the University of Chicago and appears to be a synergistic regimen, which may be as effective or more effective than cisplatin and 5-FU.

Results from the three mentioned meta-analyses have mostly involved unresectable patients. In the summary meta-analysis by El-Sayed and Nelson of 11 studies of concomitant CRT for which there were adequate survival data, a relative reduction in the hazard of death of 22% was found (95% CI = 8 to 33%, $p < .005$).[5] Another meta-analysis by Munro of 16 trials found an absolute benefit of 12% (95% CI = 5 to 19%).[4] More recently, the MACH-NC Cooperative Group performed an individual patient meta-analysis involving over 10,000 patients.[3] The hazard ratio of death in the concomitant group was 0.81 (95% CI = 0.76 to 0.88), representing an improvement in survival from 32 to 40% at 5 years. The subset of trials using multiagent chemotherapy demonstrated a hazard ratio of 0.69! The benefit of CRT must be weighed against the increased toxicity that occurs with its use. Acute mucositis, dermatitis, and chemotherapy-specific side effects are greater. Short-term gastric feeding devices are needed more frequently, and although few data exist, long-term pharyngeal dysfunction may be more common with CRT.

The latest generation of studies continues to show improved survival with aggressive CRT. This advantage is maintained even when chemotherapy is added to hyperfractionated RT,[58,63,69] so the benefit of chemotherapy is not solely owing to more intensive treatment but rather to a true radioenhancing effect.

Postoperative Adjuvant Chemoradiotherapy Only a few studies have prospectively assessed postoperative use of CRT versus RT alone in a randomized setting. In one study of 83 patients with stage III or IV HNC who had surgery but had extracapsular lymph node spread, patients were treated with RT alone or with RT plus weekly cisplatin. The combined-modality group experienced improved locoregional control as well as improved overall survival (36 versus 13% 5-year overall survival, $p < .01$).[57] In a second trial from Yale University, over 200 patients were treated with RT or RT plus mitomycin.[62] Sixty percent of patients were treated postoperatively. The combined-modality group experienced improved locoregional control and cause-specific survival. Other studies are currently completing accrual and will be published in the next few years.

Chemoradiotherapy in Resectable Head and Neck Cancer Until quite recently, CRT has had a lesser role in resectable HNC. It is now becoming more widely accepted that at least some intensive CRT programs, when used to treat patients who are appropriately medically fit, can yield results that appear to be comparable and in some cases better than definitive surgery plus RT. As of the time of the writing of this chapter, however, this concept had not yet gained universal acceptance. It should be noted that even when primary site surgery is not initially used, the head and neck surgeon still plays a major role in diagnostic procedures, follow-up, neck dissection in some patients, and surgical salvage for CRT failures.

No randomized data have compared concomitant CRT to surgery ± RT. Several indirect comparisons suggest CRT as an alternative to surgery. Adelstein et al randomized 100 patients, mostly stage IV, to RT alone or RT plus concomitant cisplatin and 5-FU on weeks 1 and 4 of RT.[56] In the CRT arm, 5-year survival was 50%. Eleven patients required

surgical salvage, which was successful in 8. Overall survival with primary site preservation was 42%. Although survival was equivalent in the RT-only arm, 27 patients required surgical salvage, which was successful in 17. Brizel et al randomized 116 stage III and IV patients to hyperfractionated RT alone or chemotherapy (cisplatin and 5-FU) plus hyperfractionated RT.[58] The 3-year survival was 55% in the CRT arm, which is comparable to or better than most surgical series. Although the survival in the subset of patients with resectable disease (47% of patients) was not reported, it was likely as good as or better than the group as a whole. In a similar study, 130 mostly stage IV patients received hyperfractionated RT ± daily cisplatin.[63] Again, 47% of CRT patients were resectable and likely had an overall survival better than the 46% 5-year survival experienced by the group as a whole. Several phase II trials have also examined this issue. Kies et al administered three cycles of induction chemotherapy using cisplatin, 5-FU, leucovorin, and interferon-α followed by CRT with 5-FU, hydroxyurea, and radiotherapy to 93 stage IV HNC patients.[64] Although not specifically reported, a large number had potentially resectable disease. The 5-year survival of the group as a whole was 62%. Since, in general, resectable patients have less biologically aggressive disease, it is likely that the resectable patients, not specifically reported on, had an even better survival. Several other studies have yielded similar results.

Based on these and other data, it is our belief that stage III or IV patients with resectable disease, who are likely to have significant functional or cosmetic sequelae from their surgery and who will require postoperative RT anyway, can be offered an aggressive definitive CRT regimen if they are medically fit.

Three published trials randomized patients into either concomitant radiation and combination chemotherapy or induction chemotherapy (using the same drugs) followed by radiation. The groups receiving concomitant therapy generally had a better disease-free and overall survival.

In summary, randomized clinical trials have shown survival benefits by combining single-agent or combination chemotherapy with radiation. However, treatment is intensive, with potential acute and chronic side effects that need to be considered in ongoing and future studies.

TREATMENT OF OTHER RARE TUMORS OF THE HEAD AND NECK

Salivary Gland Tumors

These tumors are usually treated with surgery and radiation. Chemotherapy has a palliative role only. The role of chemotherapy as a radiosensitizer for salivary gland tumors is not well studied. Because of the rarity of these tumors, as well as their heterogeneity, the entire spectrum of chemotherapy drugs has not yet been tested. The combination of cyclophosphamide, doxorubicin, and cisplatin (the "CAP" regimen) is the most active known combination.

Nasopharyngeal Carcinoma

Chemotherapy plays a crucial role in nasopharyngeal cancer. Because of the location of these tumors and the existence of other treatment options, surgery is generally not used. Early-stage nasopharyngeal tumors are generally treated for cure with RT alone. Stage III or IV tumors can be cured with RT alone, but the combination of chemotherapy and RT nearly doubles the cure rate. The most compelling study results are from an Intergroup trial in the United States that randomized 147 patients to RT alone or concomitant cisplatin and post-RT cisplatin and 5-FU. The study was stopped early when a significant difference in 2-year survival was noted in favor of the chemotherapy arm. The 3-year survival rates in the chemotherapy and nonchemotherapy groups, respectively, were 78 and 47% ($p = .005$).[71] Early follow-up of another study of 321 patients in an endemic area of Asia confirmed similar results.[72]

Metastatic undifferentiated carcinoma, or lymphoepithelioma, of the nasopharynx is highly sensitive to chemotherapy. In four consecutive studies of a total of 165 patients treated with cisplatin-containing regimens for metastatic lymphoepithelioma, 19% achieved a CR, and 64% responded with at least a partial response.[73] Twelve percent of patients were free of disease at 3 years after chemotherapy, and 14 of 165 were still disease free at least 82 months. Chemotherapy plays an important role in nasopharyngeal cancer and should now be considered a standard part of treatment for all but the few early-stage cases. The optimal timing and role of chemotherapy remain to be determined.

PREVENTION

Head and neck cancers are primarily caused by smoking and chewing tobacco. Primary prevention, therefore, continues to be of paramount importance. It has also been seen that patients with HNC who continue to smoke during RT have a lower RR and survival than those patients who quit, even at that stage. Patients with HNC treated successfully with surgery or radiation develop second primary cancer (mostly in the head and neck or lung) at an annual rate of 3 to 7%. Strategies have been devised to prevent the second primary cancers by administration of chemicals that suppress carcinogenesis and thus prevent invasive cancer. This strategy, known as chemoprevention, has assumed great importance in recent years. Lippman and Matrisian recently reviewed the dynamics of progression from normal cells to invasive oral cancers and pointed out the molecular changes that occur during progression and identified targets for chemoprevention.[48]

The premalignant lesion of HNC is usually leukoplakia. Several studies have examined the ability of synthetic (such as 13-*cis*-retinoic acid) and natural retinoids in reversing leukoplakia. Initial studies indicated a high RR but also a rapid relapse when therapy is stopped. Likewise, randomized studies to date have not demonstrated the success of retinoids in preventing the progression of premalignancy to malignancy.

High-dose 13-*cis*-retinoic acid has also been shown in some studies to reduce the development of second cancer after treatment of primary head and neck tumors. A large randomized study is currently under way to evaluate the role of retinoids in secondary prevention in this disease.

SUMMARY AND CONCLUDING REMARKS

In the past several years, chemotherapy has assumed an expanded and important role in the management of HNC. Once used purely, and relatively ineffectively, for palliation, chemotherapy now has diverse roles. Given concomitantly with RT, chemotherapy enhances cure rates in comparison with RT alone. Chemotherapy can be used prior to definitive RT for larynx or hypopharyngeal cancer without sacrificing survival opportunity and allowing for possible larynx preservation. For metastatic or incurable HNC, chemotherapy continues to be palliative. Several new chemotherapy drugs with novel mechanisms of action are in various stages of clinical trials and may provide an additional chance to increase cure rates for HNC.

REFERENCES

1. Induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. The Department of Veterans Affairs Cancer Study Group. N Engl J Med 1991;324:1685–90.
2. Lefebvre JL, Chevalier D, Luboinski B, et al. Larynx preservation in pyriform sinus cancer: preliminary results of a European Organization for Research and Treatment of Cancer phase III trial. EORTC Head and Neck Cancer Cooperative Group. J Natl Cancer Inst 1996;88:890–9.
3. Pignon JP, Bourhis J, Domenge C, Designe L. Chemotherapy added to locoregional treatment for head and neck squamous cell carcinoma: three meta-analyses of updated individual patient data. Lancet 2000;355:949–55.
4. Munro A. An overview of randomized controlled trials of adjuvant chemotherapy in head and neck cancer. Br J Cancer 1995;71:83–91.
5. El-Sayed S, Nelson N. Adjuvant and adjunctive chemotherapy in the management of squamous cell carcinoma of the head and neck region: a meta-analysis of prospective and randomized trials. J Clin Oncol 1996;14:838–47.
6. Chabner B, Collins JM. Cancer chemotherapy, principles and practice. Philadelphia: JB Lippincott; 1990.
7. DeVita VT. Principles of cancer management: chemotherapy. In: DeVita VT, Hellman S, Rosenberg SA, editors. Cancer: principles and practice of oncology. 5th ed. Philadelphia: Lippincott-Raven; 1997. p. 333–47.
8. Khandekar JD, Lawrence GA. Fundamentals in cancer management. Niles (IL): Mel Publishers; 1982. p. 75–99.
9. Ryan PD, Chabner BA. On receptors and chemotherapy. Clin Cancer Res 2000;6:4607–9
10. Huang SM, Bock JM, Harari PM. Epidermal growth factor receptor blockade with C225 modulates proliferation, apoptosis, and radiosensitivity in squamous cell carcinomas of the head and neck. Cancer Res 1999;59:1935–40.

11. Bischoff JR, Kirn DH, Williams A, et al. An adenovirus mutant that replicates selectively in p53-deficient human tumor cells. Science 1996;274:373–6.
12. Heise C, Sampson-Johannes A, Williams A, et al. ONYX-015, an E1B gene-attenuated adenovirus, causes tumor-specific cytolysis and antitumoral efficacy that can be augmented by standard chemotherapeutic agents. Nat Med 1997;3:639–45.
13. Ganly I, Eckhardt SG, Rodriguez GI, et al. A phase I study of Onyx-015, an EiB attenuated adenovirus, administered intratumorally to patients with recurrent head and neck cancer. Clin Cancer Res 2000; 6:798–803.
14. Khuri FR, Nemunaitis J, Ganly I, et al. A controlled trial of intratumoral ONYX-015, a selectively-replicating adenovirus, in combination with cisplatin and 5-fluorouracil in patients with recurrent head and neck cancer. Nat Med 2000;6:879–85.
15. Browman GP, Cronin L. Standard chemotherapy in squamous cell head and neck cancer: what we have learned from randomized trials. Semin Oncol 1994;21:311–9.
16. A phase III randomized trial of cisplatinum, methotrexate, cisplatinum + methotrexate and cisplatinum + 5-FU in end stage squamous carcinoma of the head and neck. The Liverpool Head and Neck Oncology Group. Br J Cancer 1990;61:311–5.
17. Drelichman A, Cummings G, Al-Sarraf M. A randomized trial of the combination of cis-platinum, oncovin and bleomycin (COB) versus methotrexate in patients with advanced squamous cell carcinoma of the head and neck. Cancer 1983;52:399–403.
18. Forastiere AA, Metch B, Schuller DE, et al. Randomized comparison of cisplatin plus fluorouracil and carboplatin plus fluorouracil versus methotrexate in advanced squamous-cell carcinoma of the head and neck: a Southwest Oncology Group study. J Clin Oncol 1992;10:1245–51.
19. Hong WK, Schaefer S, Issell B, et al. A prospective randomized trial of methotrexate versus cisplatin in the treatment of recurrent squamous cell carcinoma of the head and neck. Cancer 1983;52:206–10.
20. Vogl SE, Schoenfeld DA, Kaplan BH, et al. A randomized prospective comparison of methotrexate with a combination of methotrexate, bleomycin, and cisplatin in head and neck cancer. Cancer 1985;56:432–42.
21. Clavel M, Vermorken JB, Cognetti F, et al. Randomized comparison of cisplatin, methotrexate, bleomycin and vincristine (CABO) versus cisplatin and 5-fluorouracil (CF) versus cisplatin (C), in recurrent or metastatic squamous cell carcinoma of the head and neck. A phase III study of the EORTC Head and Neck Cancer Cooperative Group. Ann Oncol 1994;5:521–6.
22. Jacobs C, Lyman G, Velez-Garcia E, et al. A phase III randomized study comparing cisplatin and fluorouracil as single agents and in combination for advanced squamous cell carcinoma of the head and neck. J Clin Oncol 1992;10:257–63.
23. Eisenberger M, Hornedo J, Silva H, et al. Carboplatin (NSC-241-240): an active platinum analog for the treatment of squamous-cell carcinoma of the head and neck. J Clin Oncol 1986;4:1506–9.
24. De Andres L, Brunet J, Lopez-Pousa A, et al. Randomized trial of neoadjuvant cisplatin and fluorouracil versus carboplatin and fluorouracil in patients with stage IV-M0 head and neck cancer. J Clin Oncol 1995;13:1493–500.
25. Jaulerry C, Rodriguez J, Brunin F, et al. Induction chemotherapy in advanced head and neck tumors: results of two randomized trials. Int J Radiat Oncol Biol Phys 1992;23:483–9.
26. Brockstein B. Integration of taxanes into primary chemotherapy for squamous cell carcinoma of the head and neck: promise fulfilled? Curr Opin Oncol 2000;12:221–8.
27. Forastiere AA, Shank D, Neuberg D, et al. Final report of a phase II evaluation of paclitaxel in patients with advanced squamous cell carcinoma of the head and neck: an Eastern Cooperative Oncology Group Trial (PA390). Cancer 1998;82:2270–4.
28. Colevas AD, Posner MR. Docetaxel in head and neck cancer: a review. Am J Clin Oncol 1998;21:482–6.
29. Catimel G, Verweij J, Mattijssen V, et al. Docetaxel (Taxotere): an active drug for the treatment of patients with advanced squamous cell carcinoma of the head and neck. EORTC Early Clinical Trials Group. Ann Oncol 1994;5:533–7.
30. Couteau C, Chouaki N, Leyvraz S, et al. A phase II study of docetaxel in patients with metastatic squamous cell carcinoma of the head and neck. Br J Cancer 1999;81:457–62.
31. Dreyfuss AI, Clark JR, Norris CM, et al. Docetaxel: an active drug for squamous cell carcinoma of the head and neck. J Clin Oncol 1996;14:1672–8.
32. Colevas AD. Future directions in the treatment of squamous cell carcinoma of the head and neck: the role of UFT. Oncology (Huntingt) 1997;11:86–9.
33. Huber MH, Lippman SM, Benner SE, et al. A phase II study of ifosfamide in recurrent squamous cell

carcinoma of the head and neck. Am J Clin Oncol 1996;19:379–82.
34. Sandler A, Saxman S, Bandealy M, et al. Ifosfamide in the treatment of advanced or recurrent squamous cell carcinoma of the head and neck: a phase II Hoosier Oncology Group trial. Am J Clin Oncol 1998;21:195–7.
35. Bleomycin in advanced squamous cell carcinoma: a random controlled trial. Report of Medical Research Council Working Party on Bleomycin. BMJ 1976;1: 188–92.
36. Haas CD, Coltman CA, Gottlieb JA, et al. Phase II evaluation of bleomycin. A Southwest Oncology Group study. Cancer 1976;38:8–12.
37. Patel D. Hydroxyurea: its value in 50 cases of advanced cancers. Indian J Cancer 1973;10:423–7.
38. Livingston RB, Carter SK. Single agents in cancer chemotherapy. New York: IFI/Plenum; 1970.
39. Williams SD, Velez-Garcia E, Essessee I, et al. Chemotherapy for head and neck cancer: comparison of cisplatin + vinblastine + bleomycin versus methotrexate. Cancer 1986;57:18–23.
40. Laramore GB, Scott CB, Al-Sarraf M, et al. Adjuvant chemotherapy for resectable squamous cell carcinomas of the head and neck: report on Intergroup study 0034. Int J Radiat Oncol Biol Phys 1992;23:705–13.
41. Paccagnella A, Orlando A, Marchiori C, et al. Phase III trial of initial chemotherapy in stage III or IV head and neck cancers: a study by the Gruppo di Studio sui Tumori della Testa e del Collo. J Natl Cancer Inst 1994;86:265–72.
42. Fountzilas G, Athanassiadis A, Samantas E, et al. Paclitaxel and carboplatin in recurrent or metastatic head and neck cancer: a phase II study. Semin Oncol 1997;24 Suppl 2:S2–65.
43. Hitt R, Paz-Ares L, Hidalgo M, et al. Phase I/II study of paclitaxel/cisplatin as first-line therapy for locally advanced head and neck cancer. Semin Oncol 1997;24(6 Suppl 19):S19–S20.
44. Hussain M, Gadgeel S, Kucuk O, et al. Paclitaxel, cisplatin, and 5-fluorouracil for patients with advanced or recurrent squamous cell carcinoma of the head and neck. Cancer 1999;86:2364–9.
45. Janinis J, Papadakou M, Xidakis E, et al. Combination chemotherapy with docetaxel, cisplatin, and 5-fluorouracil in previously treated patients with advanced/recurrent head and neck cancer: a phase II feasibility study. Am J Clin Oncol 2000;23:128–31.
46. Schoffski P, Catimel G, Planting AS, et al. Docetaxel and cisplatin: an active regimen in patients with locally advanced, recurrent or metastatic squamous cell carcinoma of the head and neck. Results of a phase II study of the EORTC Early Clinical Studies Group. Ann Oncol 1999;10:119–22.
47. Specht L, Larsen SK, Hansen HS. Phase II study of docetaxel and cisplatin in patients with recurrent or disseminated squamous-cell carcinoma of the head and neck. Ann Oncol 2000;11:845–9.
48. Lippman SM, Matrisian LM. Protease inhibitors in oral carcinogenesis and chemoprevention. Clin Cancer Res 2000;6:4599–603.
49. Vokes EE, Weichselbaum RR, Lippman SM, Hong WK. Head and neck cancer. N Engl J Med 1993;328:184–94.
50. Schuller DE, Metch B, Stein DW, et al. Preoperative chemotherapy in advanced resectable head and neck cancer: final report of the Southwest Oncology Group. Laryngoscope 1988;98:1205–11.
51. Depondt J, Gehanno P, Martin M, et al. Neoadjuvant chemotherapy with carboplatin/5-fluorouracil in head and neck cancer. Oncology 1993;50 Suppl 2:23–7.
52. Domenge C, Hill C, Lefebvre JL, et al. Randomized trial of neoadjuvant chemotherapy in oropharyngeal carcinoma. French Groupe d'Etude des Tumeurs de la Tete et du Cou (GETTEC). Br J Cancer 2000; 83:1594–8.
53. Martin M, Malaurie E, Langlet PM, et al. A randomized prospective study of CDDP and 5-FU as neoadjuvant chemotherapy in head and neck cancer: a final report [abstract]. Proc ASCO 1995;14:843
54. Brockstein BE, Vokes EE. Chemoradiotherapy for head and neck cancer. Principles and Practice of Oncology Updates 1996;10:1–19.
55. Brockstein B, Weichselbaum RW, Vokes EE. Clinical update: head and neck cancer; chemoradiotherapy. Adv Oncol 1998;14:8–15.
56. Adelstein DJ, Lavertu P, Saxton JP, et al. Mature results of a phase III randomized trial comparing concurrent chemoradiotherapy with radiation therapy alone in patients with stage III and IV squamous cell carcinoma of the head and neck. Cancer 2000;88:876–83.
57. Bachaud JM, Cohen-Jonathan E, Alzieu C, et al. Combined postoperative radiotherapy and weekly cisplatin infusion for locally advanced head and neck carcinoma: final report of a randomized trial. Int J Radiat Oncol Biol Phys 1996;36:999–1004.
58. Brizel DM, Albers ME, Fisher R, et al. Hyperfractionated irradiation with or without concurrent

chemotherapy for locally advanced head and neck cancer. N Engl J Med 1998;338:1798–804.
59. Brockstein B, Haraf DJ, Stenson K, et al. Phase I study of concomitant chemoradiotherapy with paclitaxel, fluorouracil, and hydroxyurea with granulocyte colony stimulating factor support for patients with poor prognosis cancer of the head and neck. J Clin Oncol 1998;16:735–44.
60. Browman GP, Cripps C, Hodson I, et al. Placebo controlled randomized trial of infusional fluorouracil during standard radiotherapy in locally advanced head and neck cancer. J Clin Oncol 1994; 12:2648–53.
61. Calais G, Alfonsi M, Bardet E, et al. Randomized trial of radiation versus concomitant chemotherapy and radiation for advanced-stage oropharynx carcinoma, J Natl Cancer Inst 1999;15:2081–6.
62. Haffty BG, Son YH, Papac R, et al. Chemotherapy as an adjunct to radiation in the treatment of squamous cell carcinoma of the head and neck: results of the Yale mitomycin randomized trials. J Clin Oncol 1997;1:268–76.
63. Jeremic B, Shibamoto Y, Milicic B, et al. Hyperfractionated radiation therapy with or without concurrent low dose cisplatin in locally advanced squamous cell carcinoma of the head and neck: a prospective randomized trial. J Clin Oncol 2000;18:1458–64.
64. Kies MS, Haraf DJ, Athanasiadis I, et al. Induction chemotherapy followed by concurrent chemoradiation for advanced head and neck cancer: improved disease control and survival. J Clin Oncol 1998;16: 2715–21.
65. Lo TC, Wiley AL, Ansfield FJ, et al. Combined radiation therapy and 5-fluorouracil for advanced squamous cell carcinoma of the oral cavity and oropharynx: a randomized study. AJR Am J Roentgenol 1976;126:229–35.
66. Merlano M, Benasso M, Corvo R, et al. Five-year update of a randomized trial of alternating radiotherapy and chemotherapy compared with radiotherapy alone in treatment of unresectable squamous cell carcinoma of the head and neck. J Natl Cancer Inst 1996;88:583–9.
67. Taylor S, Murphy AK, Vannetzelj M. Randomized comparison of neoadjuvant cisplatin and fluorouracil infusion followed by radiation versus concomitant treatment in advanced head and neck cancer. J Clin Oncol 1994;12:385–95.
68. Vokes EE, Kies M, Haraf DJ, et al. Concomitant chemoradiotherapy as primary therapy for locoregionally advanced head and neck cancer. J Clin Oncol 2000;18:1652–61.
69. Wendt TG, Grabenbauer GG, Rodel CM. Simultaneous radiotherapy versus radiotherapy alone in advanced head and neck cancer: a randomized multicenter study. J Clin Oncol 1998;16:1318–24.
70. Suntharalingam M, Haas ML, Conley BA, et al. The use of carboplatin and paclitaxel with daily radiotherapy in patients with locally advanced squamous cell carcinomas of the head and neck. Int J Radiat Oncol Biol Phys 2000;47:49–56.
71. Al-Sarraf M, LeBlanc M, Giri PG, et al. Chemoradiotherapy versus radiotherapy in patients with advanced nasopharyngeal cancer: phase III randomized Intergroup study 0099. J Clin Oncol 1996;16: 1310–7.
72. Chan AT, Teo PM, Ngan RK, et al. A phase III randomized trial comparing concurrent chemotherapy-radiotherapy with radiotherapy alone in locoregionally advanced nasopharyngeal carcinoma [abstract]. Proc ASCO 2000;19:415a.
73. Fandi A, Bachouchi M, Azli N, et al. Long-tern disease free survivors in metastatic undifferentiated carcinoma of the nasopharynx type. J Clin Oncol 2000;18:1324–30.
74. Adelstein DJ, Adams GL, Li Y, et al. A phase III comparison of standard radiation therapy (RT) vs RT plus concurrent cisplatin (DDP) versus split course. RT plus concurrent DDP and 5-fluorouracil in patients with unrejectable squamous cell head and neck cancer: an intergroup study. Proc ASCO 2000; 19(1624):411a.

CHAPTER 58

Nutrition of the Head and Neck Patient

Laurence J. DiNardo, MD, Elizabeth G. Miller, RD

Malnutrition has been reported in 30 to 50% of head and neck cancer patients.[1] Four factors are generally responsible for calorie and protein malnutrition in these individuals. First, the tumor itself may present a barrier to deglutition through obstruction or pain. Second, tumor metabolism remains constant despite changes in nutritional status. Consequently, starvation responses such as lipolysis are not used by tumor cells, and glucose is continuously derived from protein catabolism. Third, recent evidence suggests that cancer cachexia may be induced by catabolic factors secreted by tumor cells.[2] Finally, tumor treatment may induce anorexia through discomfort, obstruction, and loss of taste. Accelerated radiotherapy has been associated with grade 3 to 4 reactions necessitating nasogastric (NG) or percutaneous gastrostomy feedings in 26% of patients with a median weight loss of 4.1 kg.[3] Difficulties are not limited to high-dose, large-field radiation treatments. A significant decrease in body mass index (BMI) has been observed in 80% of patients undergoing radiation therapy for early laryngeal carcinoma despite dietetic consultation and protein/calorie supplementation.[4] Tables 58–1 to 58–3 list the possible consequences of cancer treatment—surgery, radiation, and chemotherapy—that may impact a patient's nutritional status.[5]

Conversely, complications resulting from modern, aggressive treatment regimens for head and neck cancer are exacerbated by poor nutrition. Patients receiving accelerated fraction radiation therapy, chemoradiation, or extended resection with free flap reconstruction are particularly at risk. Preoperative weight loss of greater than 10% has been identified as the most prominent nutrition-related factor in the development of postoperative complications.[1] In addition, preoperative weight loss of greater than 5% has been associated with higher mortality rates for men with head and neck cancer.[6]

Recent work has identified relationships among immune function, nutritional status, and postoperative wound infection. Malnourished head and neck cancer patients demonstrate a significantly lower human leukocyte antigen (HLA)-DR expression on monocytes than well-nourished, matched controls.[7] Theoretically, this places these patients at a higher risk for postoperative infection. Intervention with immune-enhancing formulas has resulted in a significant decrease in the incidence of postoperative infectious complications.[8]

ASSESSMENT OF NUTRITIONAL STATUS

Recognition and assessment of nutritional status are important to the management of the head and neck cancer patient. The initial nutritional assessment includes anthropometric, biochemical, clinical, and dietary history data.

Anthropometric Measurements

The percentage of weight loss (% wt loss) that a patient has experienced is the most common estimate of malnutrition. The patient's usual body weight (UBW) is defined as his or her weight 6 months prior to the assessment. Actual, or current,

TABLE 58–1. Surgery Related Nutritional Consequences

Negative nitrogen balance
Interference with mastication
Aglutition
Dysphagia
Aspiration

TABLE 58–2. Radiation-Related Nutritional Consequences

Mucositis
Xerostomia
Odynophagia
Dysgeusia, ageusia
Dysosmia
Dental caries
Osteoradionecrosis

TABLE 58–3. Chemotherapy-Related Nutritional Consequences

Nausea
Vomiting
Diarrhea
Cheilosis, glossitis
Pharyngitis
Esophagitis
Anorexia

body weight is then measured. The % wt loss is calculated as follows:

$$\% \text{ wt loss} = (\text{UBW} - \text{current wt})/\text{UBW} \times 100$$

Severe malnutrition is defined as greater than 10% weight loss in 6 months.[9]

A patient's height is needed to allow determination of ideal body weight (IBW), percent IBW (%IBW), and additional anthropometric measurements. Ideal body weight may be determined using reference values for a given height and frame size[10] or by the Hamwi method[11]:

- For men: 106 pounds for the first 5 feet, with an additional 6 pounds for each inch over 5 feet
- For women: 100 pounds for the first 5 feet, with an additional 5 pounds for each inch over 5 feet

Percent IBW is then calculated as follows:

$$\%\text{IBW} = \text{actual (current) wt}/\text{IBW} \times 100$$

The extent of malnutrition is then estimated as follows[9]:

- Mild malnutrition = 80 to 90% IBW
- Moderate malnutrition = 70 to 79% IBW
- Severe malnutrition = < 69% IBW

Another method for assessing weight status is by determination of BMI:

$$\text{BMI} = \text{wt (kg)}/\text{ht}^2\,\text{(m)}$$

Once calculated, the patient's BMI is compared to reference values:

- Underweight: BMI < 18.5
- Adequate weight: BMI = 18.5 to 24.9
- Overweight: BMI = 25 to 29.9
- Obese: BMI > 30

Further assessment of a patient's nutritional status using anthropometric data can be performed through measurement of triceps skinfold thickness and mid-arm muscle circumference using skinfold calipers and a measuring tape. These determine the adequacy of a patient's fat and somatic protein, or skeletal muscle status. Given inconsistency in measurements among clinicians and the influence of hydration on results, these are now uncommonly performed tests. Body fat, lean body mass, and total body water can be calculated from bioelectrical impedance analysis and may be used as a bedside measurement of a patient's fluid status.[12]

Biochemical Assessment

Assessment of laboratory data comprises another facet of the nutritional status evaluation. The serum proteins that are often used to assess nutritional status include albumin, transferrin, and prealbumin or transthyretin. Each have different synthesis rates and half-lives; therefore, these parameters vary in their reflection of alterations in nutritional status.

The general availability of albumin measurement renders it a useful test. It reflects visceral protein status. Since it has a long half-life of approximately 20 days, however, it does not detect early protein deficiency. Various levels of depletion have been defined[9]:

- Adequate = > 3.5 g/dL
- Mild depletion = 2.7 to 3.5 g/dL
- Moderate depletion = 2.1 to 2.6 g/dL
- Severe depletion = < 2.1 g/dL

Albumin is influenced by many non-nutritional variables, including hydration status, liver or renal

disease, sepsis, trauma, surgery, and congestive heart failure, as well as infusion of albumin, fresh frozen plasma, or whole blood. Unfortunately, this necessitates cautious interpretation of albumin values.

In protein-calorie malnutrition from cancer-related cachexia, there appears to be an insufficient supply of amino acids for liver protein synthesis, resulting in a decline in serum albumin. A decline in serum albumin is correlated with cachexia and is a predictor of mortality. In cancer patients, postoperative complications are more frequent with serum albumin < 3.0 g/dL. Decreased tube feeding tolerance has also been observed with a serum albumin < 2.5 g/dL.

Transferrin is more useful than albumin as a measure of visceral protein status. Owing to its relatively short half-life (8 days), it is more sensitive to short-term changes in nutrient intake. The degree of malnutrition relative to serum transferrin levels is as follows:

- Adequate = 200 to 400 mg/dL
- Mild malnutrition = 180 to 200 mg/dL
- Moderate malnutrition = 160 to 180 mg/dL
- Severe malnutrition = < 160 mg/dL

As with albumin, transferrin can be impacted by non-nutritional variables. Levels may be increased with iron deficiency anemia, some malignancies, acute hepatitis, oral contraceptive use, and pregnancy. Levels may be decreased with chronic infection, cancer, iron overload, liver disease, and protein-losing enteropathy. Despite its shorter half-life, transferrin is not as commonly used as albumin or prealbumin in evaluating malnutrition and visceral protein status.

Prealbumin, also named transthyretin or thyroxine-binding globulin, has a half-life of 2 to 3 days and therefore has replaced the transferrin assay in many institutions. Its shorter half-life makes it a sensitive indicator of visceral protein status. It is most useful when assessing a patient for acute changes in nutritional status and short-term response to nutritional support through weekly measurements. Assessment of visceral protein status based on the serum transthyretin value is estimated as follows:

- Adequate = 18 to 42 mg/dL
- Mild depletion = 13 to 17 mg/dL
- Moderate depletion = 8 to 12 mg/dL
- Severe depletion = ≤ 7 mg/dL

Prealbumin levels are also impacted by non-nutrition factors. Levels can be increased with renal dysfunction, dehydration, and blood transfusions. Levels may be decreased by stress, surgery, inflammation, cirrhosis, hepatitis, and hyperthyroidism.

Retinol-binding protein, with a half-life of 10 hours, measures acute protein deprivation or minor stress. Normal values range from 2.6 to 7.6 mg/dL.

Evaluation of immune function can also provide information relevant to the patient's nutritional status and serves as a predictor of postoperative complications. In the past, intradermal antigen challenge for the detection of delayed hypersensitivity was used as a measure of immune competence. It has fallen out of favor owing to its lack of specificity. Total lymphocyte count (TLC) is currently the most frequently employed test of immune status in the head and neck cancer patient. It is calculated as follows[1]:

$$\text{TLC} = \frac{[\%\ \text{lymphocytes} \times \text{white blood count (WBC)}]}{100}$$

Immune competence is then evaluated as follows[9]:

- Adequate immunocompetence = > 1,800 mm^3
- Mild depletion = 1,500 to 1,800 mm^3
- Moderate depletion = 900 to 1,500 mm^3
- Severe depletion = < 900 mm^3

The use of TLC as an indicator of nutritional status is limited since marked changes can occur in the WBC and differential counts on a day-to-day basis. Also, numerous non-nutrition-related factors can alter the value, such as severe stress, infection, corticosteroid therapy, cancer, renal failure, surgery, and cancer treatments such as radiation therapy and chemotherapy.

Several parameters have been combined to create the prognostic nutritional index (PNI), which has been defined by various formulas.[1] An example:

$$\text{PNI} = [0.14 \times \text{alb (g/dL)}] + [0.03 \times \%\text{IBW}] + [0.73 \times \text{TLC } (10^9/\text{mm}^3)] - 8.90$$

A correlation between the PNI and surgical risk in patients with head and neck cancers has been observed. A value less than 1.31 is considered below normal.

Clinical Data and Dietary History

The physical examination is an important component of the nutritional status assessment. Physical

findings that imply vitamin, mineral, and protein-calorie deficiencies and excesses are listed in Table 58–4. However, most of the physical findings are not specific for individual nutrient deficits and must be considered together, along with diet history, anthropometric data, and laboratory results.

A diet history is used to identify underlying risks for nutrition depletion or excess. Causes include inadequate intake, compromised metabolism (altered absorption, decreased use, or increased losses), and heightened requirements for nutrients. Patients with the characteristics included in Table 58–5 have increased susceptibility for nutritional problems.[13]

The diet history of the head and neck cancer patient should include questions about alcohol consumption. When alcohol is substituted for good nutrition, malnutrition owing to inadequate nutrient intake occurs. Alcoholic persons are prone to develop malabsorption of nutrients owing to inflammation of the gastrointestinal tract. Altered absorption or metabolism of thiamin, folic acid, pyridoxine, vitamin A, vitamin B_{12}, sodium, potassium, magnesium, calcium, phosphorus, zinc, and selenium can result. Long-term alcoholism can also induce cirrhosis, with the potential development of ascites and/or hepatic encephalopathy. Glucose intolerance may be observed in alcoholics secondary to pancreatic inflammation or injury.

Commonly used methods for obtaining a diet history include the 24-hour recall, a 1- to 7-day food intake record, and a food frequency questionnaire. A number of computer programs allow rapid estimation of nutrient intake based on these records or questionnaires. It is important to remember that in the course of the complete medical history, information elicited from the patient in all of the areas noted in Table 58–6[13] will aid in identifying possible nutrient deficits not gathered from the dietary recall.

ASSESSMENT OF NUTRITIONAL NEEDS

The assessment of nutritional needs must include calorie, protein, lipid, vitamin, mineral, and water requirements. Calorie and protein needs vary depending on several factors: (1) a patient's current nutritional status, (2) whether the need is for maintenance or repletion, (3) whether a patient is at risk for refeeding syndrome, and (4) whether a patient is in a catabolic state. Adult daily calorie requirements are estimated at 30 to 35 kilocalories per kilogram of body weight (kcal/kg).[14] As many as 10 kcal/kg body weight may be added to regain lost weight and compensate for surgical stress. The Harris-Benedict equation was derived to estimate more accurately a patient's basal energy expenditure (BEE), the energy expended at rest[15]:

$$\text{For men: BEE} = 66.47 + [13.75 \times \text{wt (kg)}] + [5.0 \times \text{ht (cm)}] - [6.76 \times \text{age (y)}]$$

$$\text{For women: BEE} = 665.10 + [9.56 \times \text{wt (kg)}] + [1.85 \times \text{ht (cm)}] - [4.68 \times \text{age (y)}]$$

In general, to calculate the total calorie need of a head and neck cancer patient, the BEE is multiplied by a stress and activity factor totaling 1.3 to 1.5, with the higher number used for sepsis or other hypermetabolic states. If weight gain or nutritional repletion is desired, approximately 500 to 1,000 calories/day is added to the total calorie need. Since these calculations are estimates and not based on actual measurement of caloric expenditure, frequent monitoring of a patient's response to the nutrition regimen is required.

Acceleration in protein turnover and derangements in protein metabolism have been observed in cancer patients. In contrast to simple starvation, protein use increases when a cancer patient is under metabolic stress. Protein needs are usually calculated based on actual or adjusted body weight. The approximate protein need can then be adjusted based on the degree of protein depletion and metabolic stress factors. Under normal conditions with no stress, protein requirements are estimated at 1.0 g/kg/day. For the well-nourished, mildly stressed cancer patient, the protein need may only be 1.2 g/kg/day. However, with mild to moderate depletion along with metabolic stress, up to 1.5 g/kg/day may be necessary. If a head and neck cancer patient undergoes surgery with reconstruction, protein needs for healing may be up to 2.0 g/kg/day. The amount of protein provided may require further adjustments if renal or liver disease is present.

The best method to determine if protein needs are met in the malnourished patient is by monitoring for gradual weight gain and positive nitrogen balance. In the well-nourished patient, weight maintenance and nitrogen equilibrium are the goals.

TABLE 58–4. Physical Findings and Associated Nutrient Deficiencies/Excesses

Finding	*Deficiency*	*Excess*
Hair and nails		
Transverse hair depigmentation	Protein	
Sparse hair	Protein, biotin, zinc	Vitamin A
Corkscrew hair	Vitamin C	
Transverse nail ridging	Protein	
Skin		
Scaling	Vitamin A, zinc, fatty acids	Vitamin A
Cellophane appearance	Protein	
Cracking dermatitis	Protein	
Follicular hyperkeratosis	Vitamins A, C	
Petechiae	Vitamin C	
Purpura	Vitamins C, K	
Pigmentation, desquamation of sun-exposed areas	Niacin	
Yellow pigmentation-sparing sclerae		Carotene
Eyes		
Papilledema		Vitamin A
Night blindness	Vitamin A	
Perioral		
Angular stomatitis	Riboflavin, pyridoxine, niacin	
Cheilosis	Riboflavin, pyridoxine, niacin	
Oral		
Atrophic lingual papillae	Riboflavin, niacin, folate, vitamin B_{12}, protein, iron	
Glossitis	Riboflavin, niacin, pyridoxine, folate, vitamin B_{12}	
Hypogeusia, hyposmia	Zinc	
Swollen, retracted, bleeding gingiva	Vitamin C	
Neurologic		
Headache		Vitamin A
Drowsiness, lethargy, vomiting		Vitamins A, D
Dementia	Niacin, vitamin B_{12}, folate	
Confabulation, disorientation	Thiamin	
Ophthalmoplegia	Thiamin, phosphorus	
Peripheral neuropathy	Thiamin, pyridoxine, vitamin B_{12}	Pyridoxine
Tetany	Calcium, magnesium	
Other		
Parotid enlargement	Protein (consider bulimia)	
Heart failure	Thiamin, phosphorus	
Hepatomegaly	Protein	Vitamin A
Edema	Protein, thiamin	
Poor wound healing, decubitus ulcers	Protein, vitamin C, zinc	

TABLE 58–5. Patients at High Risk for Nutritional Deficits

Underweight: less than 80% standard and/or recent loss greater than 10% usual body weight
Poor intake: anorexia, nothing by mouth more than 5 days
Protracted nutrient loss: malabsorption, enteric fistulae, draining abscesses or wounds, renal dialysis
Hypermetabolic state: sepsis, fever, extensive burn, or trauma
Chronic alcohol/drug use: corticosteroids, antimetabolites, immunosuppressants
Impoverishment, isolation, advanced age

Weekly serial monitoring of a patient's prealbumin or transthyretin level is another method for tracking protein supplementation.

Ideally, ≤ 30% of nonprotein calories should be derived from fat. The appropriate administration of essential fatty acids helps to blunt the insulin response by decreasing the need for infusions rich in glucose. Deriving calories from fat also has a protein-sparing effect.[16] Using greater amounts of fat in combination with less glucose is especially important in patients with diabetes mellitus in whom fat can provide as much as 60% of nonprotein calories.

Although recommended daily doses of vitamins, minerals, and trace elements exist,[17] requirements for nutritionally depleted cancer patients have not been established. It is recommended that patients requiring prolonged parenteral nutrition receive regular doses of multivitamins and trace elements. More definitive guidelines will likely evolve as knowledge is gained regarding the role of vitamins, minerals, and trace elements in critically ill patients. Experimental evidence exists in mice, for example, which demonstrates increased survival in endotoxin shock with the administration of vitamin D analogues.[18]

NUTRITIONAL SUPPORT

Whenever possible, nutritional support should be provided through the enteral route. The use of an intact gastrointestinal tract is not only more cost effective, safer, and more practical than parenteral nutrition, it also maintains the integrity of the gut mucosal barrier and immune function.[19] In addition, no survival advantage or reduction in treatment toxicity has been demonstrated with the routine administration of total parenteral nutrition in cancer patients.[20]

Oral alimentation should always be considered first. Simple interventions such as the eradication of oral candidiasis or the treatment of mucositis may facilitate oral intake. Nutrition counseling and provision of nutritional supplements can further support oral intake in head and neck cancer patients. Nutritious liquid formulas can be taken by mouth when other foods cannot be tolerated. The administration of appetite stimulants such as megestrol acetate has been advocated in patients with cancer-related anorexia or cachexia. The risk of side effects such as adrenal suppression, hyperglycemia, and thromboembolic events must be considered with this medication.[21] However, if surgery, radiation, or chemotherapy side effects prevent a patient from taking oral liquids, nutritional support can be provided via a feeding tube. Despite the ability to take some oral nutrition, the seriously malnourished patient should be considered for early feeding tube placement owing to the impending side effects of cancer therapy. A review of head and neck cancer patients who were treated with radiation therapy found that early initiation of enteral nutrition lowered the need for treatment breaks and possibly led to enhanced treatment efficacy.[22]

Types and Selection of Feeding Tubes

Enteral feeding involves the administration of nutrient solutions through a tube into the upper gastrointestinal tract. There are two major routes for tube placements: those entering the gastrointestinal tract via the nose (NG and nasointestinal) and those entering the gastrointestinal tract through the abdominal wall (gastrostomy, duodenostomy, jejunostomy). Occasionally, a feeding may be given through a tube inserted in the mouth (orogastric) or esophagus (cervical esophagostomy).

The timing and type of enteral tube feeding are largely dependent on the anticipated duration of nutritional supplementation. Placement of a pliable red rubber or Dobbhoff tube for feeding is preferred if short-term nutritional support is expected. Nasogastric tube feeding should be restricted to short-term use owing to nasopharyngeal and laryngeal

TABLE 58–6. Nutrition History Survey

Mechanism of Deficiency	*History*	*Deficiency to Suspect*
Inadequate intake	Alcoholism	Calories, protein, thiamin, niacin, folate, pyridoxine, riboflavin
	Avoidance of fruit, vegetables, grains	Vitamin C, thiamin, niacin, folate
	Avoidance of meat, dairy products, eggs	Protein, vitamin B_{12}
	Constipation, hemorrhoids, diverticulosis	Dietary fiber
	Isolation, poverty, dental disease, food idiosyncrasies	Various nutrients
	Weight loss	Calories, other nutrients
Inadequate absorption	Drugs (especially antacids, anticonvulsants, cholestyramine, laxatives, neomycin, alcohol)	Varies, dependent on drug
	Malabsorption (diarrhea, weight loss, steatorrhea)	Vitamins A, D, and K; calories; protein; calcium; magnesium; zinc
	Parasites	Iron, vitamin B_{12} (fish tapeworm)
	Pernicious anemia	Vitamin B_{12}
	Surgery Gastrectomy Intestinal resection	Vitamin B_{12} (if distal ileum), iron, others as in malabsorption
Decreased utilization	Drugs (especially anticonvulsants, antimetabolites, oral contraceptives, isoniazid, alcohol)	Varies, dependent on drug
	Inborn errors of metabolism (by family history)	Various nutrients
Increased losses	Alcohol abuse	Magnesium, zinc
	Blood loss	Iron
	Paracentesis (ascitic, pleural)	Protein
	Diabetes, uncontrolled	Calories
	Diarrhea	Protein, zinc, electrolytes
	Draining abscesses, wounds	Protein, zinc
	Nephrotic syndrome	Protein, zinc
	Peritoneal dialysis or hemodialysis	Protein, water-soluble vitamins, zinc
Increased requirements	Fever	Calories
	Hyperthyroidism	Calories
	Physiologic demands (infancy, adolescence, pregnancy, lactation)	Various nutrients
	Surgery, trauma, burns, infection	Calories, protein, vitamin C, zinc
	Tissue hypoxia	Calories (inefficient utilization)
	Cigarette smoking	Vitamin C, folate

irritation as well as potential for necrosis of the nasal alae. With the placement of an NG tube, the patient can be weaned quickly from intravenous hydration after surgery, receive adequate nutrition postoperatively, and have access for medications. If there is difficulty transitioning to complete oral intake, then the patient is considered for placement of a percutaneous endoscopic gastrostomy (PEG), which is more appropriate for long-term feeding. If it is anticipated that the patient will require long-term tube feeding postoperatively, a PEG tube can be placed intraoperatively. Intraoperative PEG tube placement has been associated with fewer complications than preoperative insertion.[23] Open or laparoscopic gastrostomy tube placement is considered in cases for which PEG placement is not feasible.

Intestinal feedings are indicated when gastric reflux results in aspiration, with depressed gastric motility and emptying, with significant pulmonary disease, or when ulcer disease, cancer, or surgery involves the stomach. Tube feeding into the intestine may be provided via a nasointestinal tube, PEG-jejunostomy, or a surgical jejunostomy tube. If a patient is dependent on a feeding jejunostomy tube for medication administration, an 8 French tube or larger should be used to prevent clogging.

Tube Feeding Formulas

Numerous oral and tube feeding solutions are available. Some are intended for general nutrition; others are designed to meet specific metabolic or clinical needs. Occasionally, blenderized foods can be used for the provision of adequate nutrition by oral or feeding tube routes.

The majority of head and neck cancer patients are administered commercially available polymeric formulas. Polymeric formulas consist of macronutrients in the form of isolates of intact protein, fats, and carbohydrate polymers. In most of these solutions, protein makes up 12 to 20% of total calories; carbohydrates, 40 to 60%; and fats, 30 to 40%. In the standard formulas, the ratio of nonprotein calories to nitrogen is approximately 150 to 1. In the high-nitrogen formulas, this ratio is much lower, approximately 75 to 1. Table 58–7 provides an overview of these formulas.[24]

Polymeric formulas contain whole proteins isolated from casein, whey, lactalbumin, and egg white. The source of carbohydrate is usually glucose polymers from starch and its hydrolysates. The fats are from vegetable sources. Essential vitamins and minerals are present in sufficient amounts so that 100% of the United States Recommended Daily Allowance for each nutrient is met through a daily intake of 1,500 to 2,000 calories of the formula. The amounts of minerals such as sodium and potassium vary greatly, allowing for adjustment based on requirements.

Polymeric formulas are lactose free with a caloric density between 1 and 2 cal/mL. The higher-calorie formulas allow for more calories in a smaller volume and are therefore suitable for patients requiring fluid restriction. These high-osmolality formulas often are best initiated at half strength and advanced as tolerated to full strength while maintaining appropriate hydration. Modified, more costly, polymeric formulas have been devised to suit the needs of patients with renal failure, liver failure, and diabetes.

Monomeric formulas require less digestion than do regular foods or polymeric solutions. They are suitable for patients with impaired digestion. These formulas are rarely used in head and neck cancer patients with the exception of extremely low-fat preparations administered in the presence of a chylous fistula. Formulas with medium-chain triglycerides are also useful in the presence of a chyle leak since they enter the systemic circulation without passing through the thoracic duct.

Tube Feeding Administration

With an NG tube, placement in the stomach must be verified prior to each feeding and the head of the patient's bed elevated. Initiation of bolus feeding may range from 100 to 200 mL every 3 to 4 hours, with the final feeding provided to the patient no later than 11:00 pm. Feedings should be administered by gravity drip over 30 to 60 minutes. Advancement of feedings by 50 to 100 mL every shift (8h) or daily are performed as tolerated, and residuals are monitored until the tube feeding volume goal is attained.

Initiation of tube feeding through a surgically placed gastric tube is similar to NG tube feeding except a delay of as much as 48 hours after insertion is required to allow for maturation of the stoma and return of bowel function.

After PEG or percutaneous radiologic gastrostomy feeding tube placement, feeding can often be initiated within 12 hours. If bowel sounds are present, continuous delivery of water at 50 mL/hour can

TABLE 58–7. Comparison and Components of Selected Polymeric Enteral Formulas

Product	*Manufacturer*	*Cal/mL*	*Protein (g/L)*	*Osmolality (moSm/kg)*	*Nonprotein Cal:N ratio*
Isocal	Mead Johnson	1.06	34	270	168
Isocal HN	Mead Johnson	1.06	44	270	125
Isocal HN Plus	Mead Johnson	1.2	54	400	114
Boost	Mead Johnson	1.01	43	610–670	125
Boost Plus	Mead Johnson	1.52	61	630–670	134
Ensure	Ross	1.06	37.2	555	153
Ensure Plus	Ross	1.5	54.9	690	146
Ensure Plus HN	Ross	1.5	62.6	650	125
Isosource	Novartis	1.2	43	490	149
IsosourceHN	Novartis	1.2	53	490	115
Nutren 1.0	Nestle	1.0	40	300–350	134
Nutren 1.5	Nestle	1.5	60	430–530	134
Nutren 2.0	Nestle	2.0	80	720	134
Osmolite	Ross	1.06	37.1	300	153
Osmolite HN	Ross	1.06	37.1	300	153
Osmolite HN Plus	Ross	1.2	55.5	360	110
Resource Plus	Novartis	1.5	55	600	148
Deliver 2.0	Mead Johnson	2.0	75	640	144
Two-Cal HN	Ross	2.0	83.5	690	125
Promote	Ross	1.0	62.5	340	75
Replete	Nestle	1.0	62.4	300–350	75
Boost with Fiber	Mead Johnson	1.06	46	480	120
Ultracal	Mead Johnson	1.06	45	360	124
Ultracal HN Plus	Mead Johnson	1.2	54	370	114
Compleat Modified	Novartis	1.07	43	300	131
Isosource 1.5 Cal	Novartis	1.5	68	650	116
Jevity	Ross	1.06	44.3	300	125
Jevity Plus	Ross	1.2	55.5	450	110
Nutren 1.0 with Fiber	Nestle	1.0	40	310–370	134
Probalance	Nestle	1.2	54	350–450	114
Promote with Fiber	Ross	1.0	62.5	380	75
Replete with Fiber	Nestle	1.0	62.5	310–390	77

begin within a few hours of insertion. If water is tolerated, feeding is initiated using a full-strength, isotonic formula infused at 50 mL/hour. Residuals should be monitored every 4 hours, and feeding is advanced as with an NG or gastric tube.

Feeding into the intestine requires an intermittent or continuous pump-assisted delivery of formula since bolus amounts and hyperosmolarity can lead to diarrhea. After placement of a nasointestinal (nasoduodenal or nasojejunal) tube, feedings are administered over a 12- to 24-hour period and should be started at an isotonic level (300 mOsm) or at a low rate (10 to 15 mL/hour). The rate of infusion may be increased every 8 to 24 hours as tolerated until the desired volume is reached. The rate and concentration of the formula should not be increased simultaneously.

Administration of formula through a feeding jejunostomy tube is identical to a nasointestinal feeding tube except feeding initiation may not begin for 12 to 24 hours after tube placement. With a PEG/jejunostomy, initiation is the same as for a PEG except a continuous infusion of formula is required.

Feeding Tube Complications

Tube Clogging If a feeding tube becomes clogged, any enteral solution remaining in the tube should be withdrawn. To remove any obstruction, the following steps are attempted: (1) Inject 5 mL of warm water into the tube and clamp for 5 minutes. Flush with water until clear; if the tube remains clogged, then (2) inject a bolus (20 to 30 cc) of air to dislodge the clog. If the air seems to free the obstruction, follow it with 30 to 50 mL of warm water, cola, or cranberry juice to cleanse the tube. If the tube remains clogged, (3) crush one 324 mg sodium bicarbonate tablet. Mix the powder with the contents of one pancrelipase (Cotazym or Viokase) capsule and 5 mL of sterile water. Inject the alkalinized enzyme mixture into the tube and clamp it for 5 minutes. Flush the tube with water until clear. If the aforementioned techniques fail, the tube should be replaced.

Formula–Drug Interactions Formula–drug interactions resulting in delayed or diminished drug absorption are outlined in Table 58–8.[25] In addition, physical incompatibilities between enteral formulas and medications may result in tube obstruction, as detailed in Table 58–9.[25]

Diarrhea Diarrhea is defined as an increase in the quantity and/or frequency of bowel movements. Diarrhea has been reported to occur in approximately 30% of patients receiving tube feeding.[26] However, tube feeding itself is often not the cause.[27] Potential causes of diarrhea include (1) an inappropriate rate of formula infusion, (2) impaired functional capacity of the gastrointestinal tract, (3) hypoalbuminemia, (4) the concurrent use of antibiotics and other medications, (5) altered bacterial flora, and (6) enteral formula contamination.

The healthy gastrointestinal tract can handle intermittent feedings of as much as 500 mL given over 10 to 15 minutes. This feeding method has routinely been used with success in patients with dysphagia owing to cancer of the head and neck.[28] If diarrhea occurs, slower delivery of intermittent feedings or a change in the tube feeding product to one that is isotonic or contains fiber may be helpful.

Enteral nutrition-related diarrhea is often secondary to an osmotic effect, which is the result of the small intestine's inability to absorb specific nutrients. This condition is frequently associated with a decline in intestinal blood flow and absorptive capacity. Hypoalbuminemia, which is associated with poor water absorption by the colon, can also contribute to diarrhea. This is observed in patients with edema and fluid overload. Switching the enteral solution to a more calorie-dense product (1.5 to 2.0 cal/mL) will decrease free water intake and may help correct fluid balance. When serum albumin is severely low, intravenous infusion of albumin may be indicated to correct fluid balance and promote protein repletion.

The diarrhea resulting after tolerance to an enteral formula has been achieved may be attributable to bacterial overgrowth. This overgrowth is often caused by long-term antibiotic use. If suspected, stool cultures for *Clostridium difficile* should be performed and the patient treated with oral vancomycin or metronidazole as indicated. The use of a fiber-containing formula has been helpful in correcting loose stools in these cases. Antidiarrhea medications should be used with caution if an infectious cause of diarrhea is suspected.

Another potential cause of diarrhea is fecal impaction in which loose stool can leak around the area of impaction until the bowel is cleared. Medications, especially antacids, can also induce diarrhea and should be considered.

TABLE 58–8. Significant Drug–Nutrient Interactions

Medications	*Interaction*	*Possible Outcome*
Antibiotics		
Tetracycline	Decreased bioavailability with milk and dairy products	Treatment failure
Quinolones (ciprofloxacin/ norfloxacin)	Decreased bioavailability with milk and dairy products	Treatment failure
Azithromycin	Decreased bioavailability with food	Treatment failure
Antifungal		
Itraconazole	Increased bioavailability with food	Possible treatment failure if not administered with meals
Antiviral		
Didanosine Indinavir	Food decreases bioavailability	Treatment failure
Saquinavir	Food increases bioavailability	Increased antiviral activity if administered with meals
Warfarin	Vitamin K–rich foods antagonize the anticoagulant effect	Decreased anticoagulation
Cyclosporin	Food and grapefruit juice increase plasma levels	Possible toxicity, lower doses may be efficacious
Alendronate	Food decreases bioavailability	Treatment failure

Refeeding Syndrome Refeeding syndrome is a serious metabolic disturbance that may occur with the initiation of aggressive nutrition support in a patient with severely depleted nutrition stores. Significant compartmental shifts in phosphorus, magnesium, and potassium can be induced. These shifts, as well as sodium retention with resultant expansion of the extracellular space and associated increased circulatory demands, may prove disastrous.

Hypophosphatemia is potentially the most serious single consequence in refeeding syndrome.[29] In starvation, there is a slowing of the overall metabolic rate and, in particular, glucose oxidation. The need for phosphorus is low. When carbohydrates and glucose are reintroduced, this need dramatically rises and may surpass the rate at which phosphorus can be mobilized from bone. Cardiopulmonary failure and death may occur within days of feeding a previously starved but stable patient.

In patients at risk for refeeding syndrome, normal levels of phosphorus, potassium, and magnesium should be present prior to initiating nutrition support. Calorie repletion begins at less than 50% of need and is advanced over the course of a week to a goal of 30 cal/kg/day or less. If adequate renal and hepatic function exists, 1.5 g of protein/kg/day can also be administered. A daily multivitamin, mineral supplement, thiamin (100 mg/day), and selenium (50 μg/day) are recommended. Careful monitoring of vital signs, weight, and input and output, as well as electrolytes, is necessary during nutrition support of at-risk individuals.

Parenteral Nutrition

Parenteral nutrition is the intravascular administration of carbohydrates, protein, fat, vitamins, and minerals. Total parenteral nutrition involves nutrient administration into the superior vena cava, thereby allowing for the infusion of hypertonic solutions. Hyperalimentation solutions typically contain 25% glucose given at a gradually increasing infusion rate

TABLE 58–9. Mediations Incompatible with Enteral Nutritional Formulas

Product	*Incompatibility/Interaction*
Vitamin B elixir	Forms large particle clumping
Brompheniramine elixir	Forms adhesive gelatinous material
Brompheniramine phenylpropanolamine elixir	Enteral food breakdown
Calcium glubionate syrup	Adhesive gelatinous mass created
Chlorpromazine concentrate	Granulation occurs at point of mixing
Doxepin liquid	Feeding separates slightly with no clogging
Ferrous sulfate elixir	Completely gels and clogs feeding tubes
Guaifenesin expectorant	Viscous precipitate forms
KCl 10% and 20% liquid	Incompatible at mixing interface
KCl syrup	Enteral products become viscous and gelatinous
Lithium syrup	Enteral products form granular film
Medium-chain triglyceride oil	Immiscible with enteral products
Methenamine mandelate suspension forte	Mixture becomes gelatinous
Metoclopramide	Some feedings form thin granular formulation
Opium liquid	Separation, globular particle formation
Phosphosoda	Results in heavy or granular feedings
Pseudoephedrine syrup	Viscous gelatinous mass forms on mixing
Sucralfate	Solidifies formula
Thioridazine solution	Granule formation
Zinc sulfate capsules	May induce gelatinous or hard mass formation

toward a goal of 5 mg glucose/kg/minute. Insulin requirements must be closely monitored. Amino acids are administered in 2.5 to 5.0% solutions in sufficient quantity to ensure positive nitrogen balance. The remainder of noncarbohydrate calories is derived from fat emulsions. Peripheral partial nutrition supplements enteral feeding with the administration of a 10% glucose solution, amino acids, and fat emulsion through a peripheral vein.

Complications from parenteral nutrition are significantly higher than for enteral feeding and include sepsis, hyperglycemia, hypoglycemia, hyperlipidemia, hepatic dysfunction, electrolyte imbalance, and thrombophlebitis. The indications for parenteral nutrition in the head and neck patient are limited to gastrointestinal dysfunction or intractable aspiration of tube feedings. Occasionally, total parenteral nutrition is used to treat a chyle fistula not responsive to medium-chain triglyceride formulas.

Conclusion

Several factors conspire to induce malnutrition in the head and neck patient. It is necessary to identify and quantify nutritional deficits to replete losses adequately. Enteral feeding is desired with a functioning gastrointestinal tract, and various maneuvers exist to use an intact gut better. Nutrition repletion should be carefully monitored to allow the patient the potential benefits of increased treatment tolerance and perhaps improved survival.

REFERENCES

1. van Bokhorst-de van der Schueren MA, van Leeuwen PA, Sauerwein HP, et al. Assessment of malnutrition parameters in head and neck cancer and their relation to postoperative complications. Head Neck 1997;19:419–25.

2. Todorov P, Cariuk P, McDevitt T, et al. Characterization of cancer cachectic factor. Nature 1996; 379:739–42.
3. Allal AS, Maire D, Becker M, Dulguerov P. Feasibility and early results of accelerated radiotherapy for head and neck carcinoma in the elderly. Cancer 2000;88:648–52.
4. Collins MM, Wight RG, Partridge G. Nutritional consequences of radiotherapy in early laryngeal carcinoma. Ann R Coll Surg Engl 1999;81:376–81.
5. Ursule G, Kyle. The patient with head and neck cancer. In: Bloch AS, editor. Nutrition management of the cancer patient. 1990. p. 55.
6. van Bokhorst-de van der Schuer MAE, van Leeuwen PA, Kuik DJ, et al. The impact of nutritional status on the prognoses of patients with advanced head and neck cancer. Cancer 1999;86:519–27.
7. van Bokhorst-De van der Schuer MA, von Blomberg-van der Flier BM, Riezebos RK, et al. Differences in immune status between well-nourished and malnourished head and neck cancer patients. Clin Nutr 1998;17:107–11.
8. Snyderman CH, Kachman K, Molseed L, et al. Reduced postoperative infections with an immune-enhancing nutritional supplement. Laryngoscope 1999;109:915–21.
9. Gottschlich MM, Matarese LE, Shronts EP. Nutrition support dietetics core curriculum. 2nd ed. American Society of Parenteral Nutrition; 1993.
10. Metropolitan Life Insurance Company. Height and frame size vs body weight. January-June: Statistical Bulletin, 1983.
11. Hamwi GJ. Therapy: changing dietary concepts. In: Donowski TS, editor. Diabetes mellitus: diagnosis and treatment. New York: American Diabetes Association; 1964. p. 55.
12. Scheltinga MR. Bioelectrical impedance analysis (BIA): a bedside method for fluid measurement. Studies of the bodies electrical impedance in various models of health and disease [thesis]. Amsterdam: Vrije Universiteit; 1992.
13. Heimburger DC, Weinsier RL. Nutritional assessment. In: Handbook of clinical nutrition. 3rd ed. Mosby-Year Book; 1997. p. 183.
14. National Academy of Sciences. Recommended Dietary Allowances. Washington (DC): US Government Printing Office; 1974.
15. Cerra FB. Surgical nutrition St. Louis: CV Mosby; 1984.
16. Nordenstrom J, Askanazi J, Elwyn DH, et al. Nitrogen balance during total parental nutrition: glucose vs fat. Ann Surg 1983;197:27–33.
17. Gallagher-Allred C. Vitamin and mineral requirements. In: Krey SH, Murray RL, editors. Dynamics of nutrition support. Norwalk (CT): Appleton-Century-Crofts; 1986.
18. Horiuchi H, Nagata I, Komoriya K. Protective effect of vitamin D3 analogues on endotoxin shock in mice. Agents Actions 1991;33:343–8.
19. Mercandante S. Parenteral versus enteral nutrition in cancer patients: indications and practice. Support Care Cancer 1998;6:85.
20. Klein S, Kinney J, Jeejeebhoy K, et al. Nutrition support in clinical practice: review of published data and recommendations for future research directions. Am J Clin Nutr 1997;66:683–706.
21. Loprinzi CL, Kugler JW, Sloan JA, et al. Randomized comparison of megestrol acetate versus dexamethasone versus fluoxymesterone for the treatment of cancer anorexia/cachexia. J Clin Oncol 1999;17: 3299–306.
22. Zogbaum AT, Farkas S, Pease CR, et al. Enteral feedings are associated with improved adherence to radiation treatment prescription and weight maintenance in head and neck cancer patients. J Am Diet Assoc 1996;96 Suppl:A-35.
23. Raynor EM, Williams MF, Martindale RG, Porubsky ES. Timing of percutaneous endoscopic gastrostomy tube placement in head and neck cancer patients. Otolaryngol Head Neck Surg 1999;120:479–82.
24. Medical Nutritionals Product Advisor. Mead Johnson Nutritionals, Mead Johnson and Company, 2000.
25. Huey L, Way J. Medication administration via enteral access tubes. In: Patient care news. Sherwood Medical Company; 1997. p. 2–3.
26. Guenter PA, Settle RG, Perlmutter S, et al. Tube feeding related diarrhea in acutely ill patients. JPEN J Parenter Enteral Nutr 1991;15:277–80.
27. Edes TE, Waek BE, Austin JL. Diarrhea in tube-fed patients: feeding formula not necessarily the cause. Am J Med 1990;88:91–3.
28. Shike M, Berner YN, Gerdes H, et al. Percutaneous endoscopic gastrostomy and jejunostomy for long-term feeding in patients with head and neck cancer. Otolaryngol Head Neck Surg 1989;101:549–54.
29. Crook MA, Collins D, Swaminathan R. Severe hypophosphatemia related to refeeding. Nutrition 1996;12:538–9.

CHAPTER 59

Imaging of the Oral Cavity, Pharynx, Larynx, Trachea, Salivary Glands, and Neck

Mahmood F. Mafee, MD

See Chapter 31 for the theoretical basis of computed tomography (CT) and magnetic resonance imaging (MRI).

ORAL CAVITY, OROPHARYNX, AND UPPER NECK

Imaging Anatomy

The detailed anatomy of the oral cavity and oropharynx can be found in Chapter 43. For the purpose of this chapter, the cross-sectional anatomy visible on CT and MR images will be reviewed and defined. Several CT and MR images of the normal oropharynx and upper neck are shown in Figure 59–1, A through H. For the purpose of uniformity, the scans are all spin density– or proton-weighted MRIs (see Figure 59–1). In axial CT sections or MRIs taken at the level of the floor of the mouth and tongue (see Figure 59–1, A through D), the oropharynx, palatine tonsils, submandibular glands, extrinsic and intrinsic muscles of the tongue, external and internal carotid arteries, internal jugular veins, and parts of the mandibles are among the many structures that are visualized clearly (see Figure 59–1). In axial CT or MRI sections taken at the level of the alveolar process of the maxilla, the soft palate is shown blending with the upper portion of the oropharynx (see Figure 59–1, G and H). The hard palate is limited laterally by the alveolar process of the maxilla. The medial pterygoid muscle is seen as a distinct quadrilateral image occupying a position on the inside of the ramus of the mandible (see Figure 59–1, E through H). Medial to the medial pterygoid muscles, the symmetric low-density (lucent) parapharyngeal spaces are shown. Also visualized are the masseter muscle, retromandibular vessels, parotid gland, styloid process, and retrostyloid neurovascular structures (see Figure 59–1). The differential enhancement of the vascular structures of the neck versus other structures is best demonstrated with the incremental CT dynamic technique. On spin echo MRIs, the vascular structures (arteries) are demonstrated as hypointense images (signal void). On magnetic resonance angiogram (MRA) and magnetic resonance venogram (MRV) without need for contrast injection, the vascular structures are seen as hyperintense images.

ORAL CAVITY, OROPHARYNX, HYPOPHARYNX, LARYNX, AND NECK

Imaging Pathology

Benign and malignant lesions of the oral cavity and oropharynx can be readily visualized with CT and MRI. Figure 59–2 shows a lymphangiohemangioma of the tonsillar region, and Figure 59–3 shows a carcinoma of the tonsil and soft palate with bilateral lymph nodal metastases. With MRI or CT, a primary tumor and metastatic lymph nodal involvement can often be readily identified (Figures 59–3 and 59–4). Figure 59–5 demonstrates the value of MRI, including MRA, in the evaluation of a recurrent tonsillar cancer and thrombosis of the internal jugular vein. Tumors of the tongue and floor of the mouth (Figure 59–6), posterior pharyngeal wall (Figure 59–7), and hypopharynx (Figure 59–8) are readily evaluated with CT and MRI.

Infection, trauma, congenital malformation, and benign and malignant tumors in this region are evaluated adequately by clinical examination, plain films, barium swallow, cine-esophagram, and ultrasonography. The ability of MRI and CT to provide superficial and deep soft and hard tissue information and tissue identification (fat, muscle,

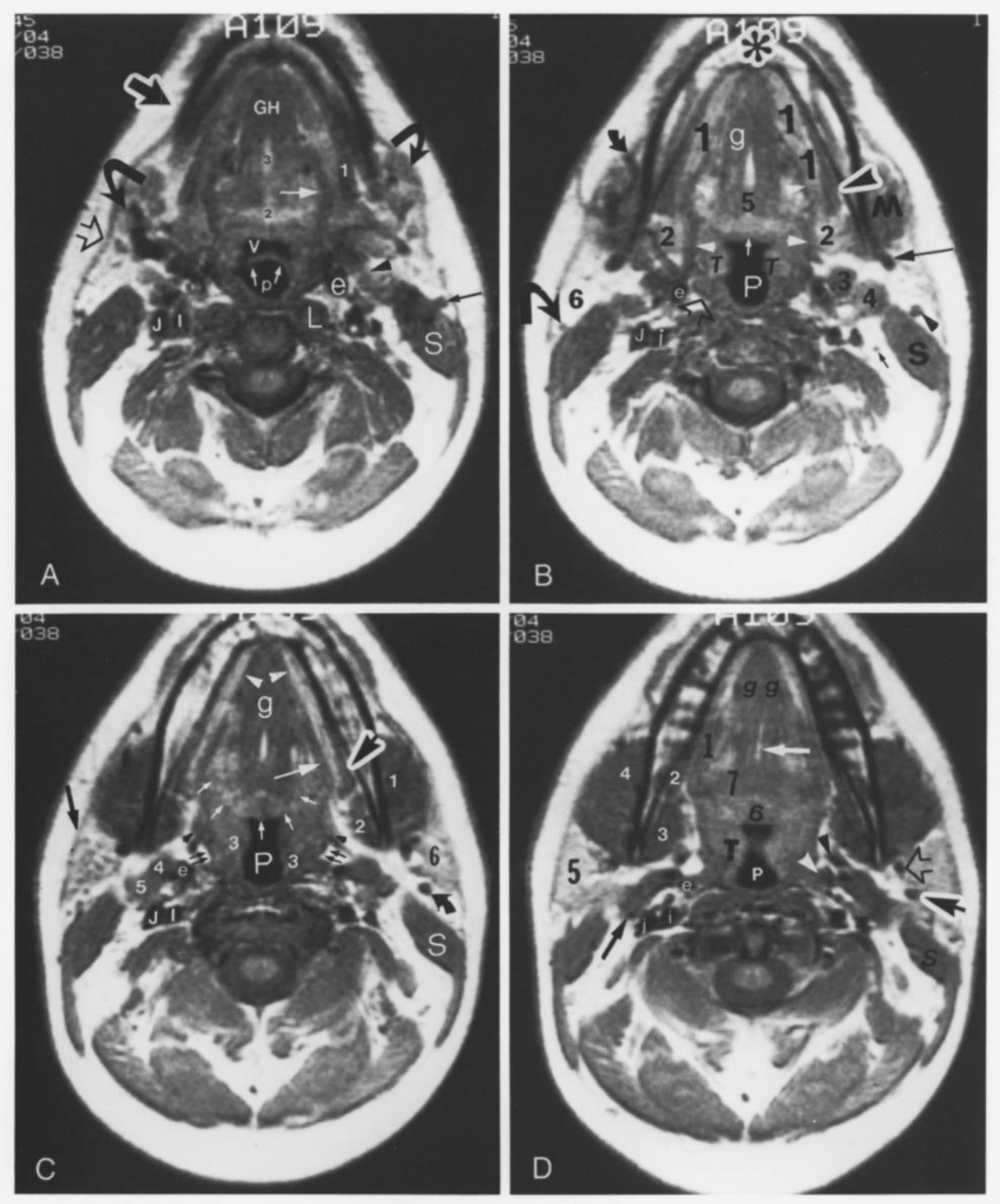
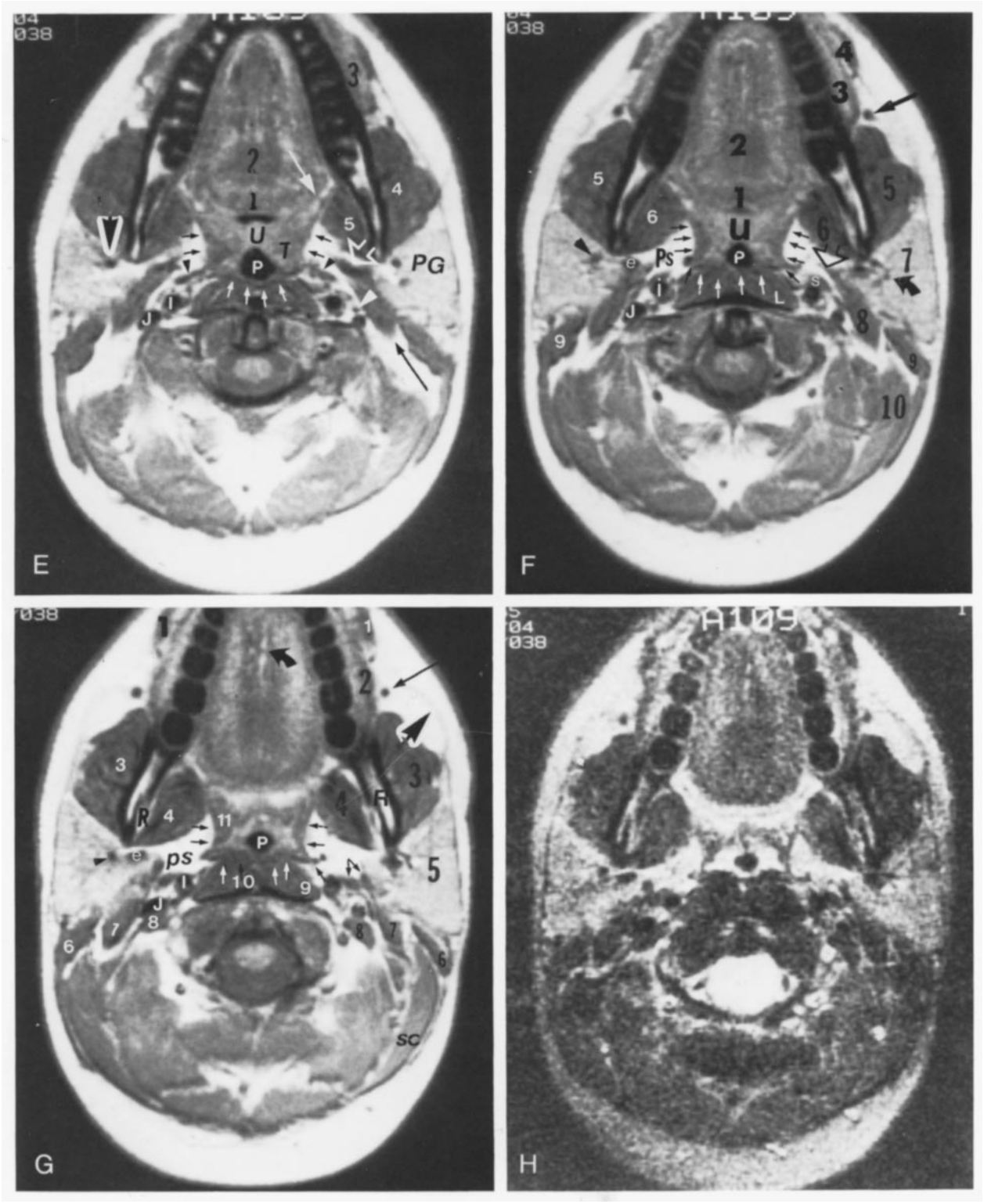


FIGURE 59–1. All except *H* are proton-density 2000/20 magnetic resonance images of the same normal subject. *A*, This section is through the lower margin of the mandible (*black and white arrow*) and through the body of the third cervical vertebra. The section passes through the lower part of the tongue and through the oral portion of the pharynx (p), just above the epiglottis (*small white arrows*). The submandibular glands (*curved arrows*) are cut through their middle portions. The platysma (*open arrow*) and external jugular vein (*black arrow*) are seen. The hyoglossus muscle (*white arrow*),



the stylohyoid muscle (*arrowhead*), and the tendinous portion of the digastric (lateral to the styloid process) muscle are seen medial to the submandibular gland. 1 = mylohyoid muscle; 2 = intrinsic muscle of the tongue with associated intermuscular fat; 3 = median raphe; e = external carotid artery; GH = geniohyoid muscle; I = internal carotid artery; J = internal jugular vein; L = longus colli muscle; S = sternocleidomastoid muscle. *B*, This section passes through the body of the mandible, just above the mental foramen, and through the lower margin of the axis. It passes through the tongue just

above the level of the foramen cecum (*small white arrow*) and the inferior margins of the palatine tonsils (T). Little of the parotid gland (6) is seen in this section. The maximal anteroposterior extensions of the sublingual glands are seen (1), and the submandibular glands (2) are also demonstrated. Note the intrinsic muscle of the tongue (5); the genioglossus muscle (g); the pharynx (P); masseter (M), mylohyoid (*white and black arrowhead*), stylohyoid, and digastric (3) muscles; a possible lymph node and digastric muscle (4); sternocleidomastoid muscle (S); spinal accessory nerve (*small black arrow*); auricular nerve (*long black curved arrow*); external (e) and internal (i) carotid arteries; internal jugular vein (J); retromandibular vein (*black arrow*); anterior facial vein (*short black curved arrow*); styloglossus muscles posterior to the hyoglossus muscles (*white arrowheads*); and pharyngeal constrictor muscles (*open arrow*). *C*, This section passes through the mandible, near the junction of the ramus and the body and the tongue at the level of the upper portions of the sublingual glands (*white arrowheads*). The high-density image medial to the mandible is caused by the sublingual gland and fat in the fascial planes. This section passes through the pharynx (P) at the level of the middle portions of the tonsils (3). The tongue shows the following muscles: genioglossus (g), hyoglossus (*white arrow*), and styloglossus (*black arrowheads*). Note the mylohyoid (*black and white arrowhead*), platysma (*black arrow*), parotid gland (6), digastric muscle (5), stylohyoid muscle (4), external jugular vein (*curved arrow*), and sternocleidomastoid muscle (S). *D*, This section is 3 mm superior to *C* and shows various structures: genioglossus muscles (gg); septum linguae (*white arrow*); stylohyoid muscle (*black arrowhead*); stylopharyngeus muscle (*white arrowhead*); tonsils (T); oropharynx (P); longitudinal (6) and transverse (7) muscle fibers of the tongue; hyoglossus (1), mylohyoid (2), internal pterygoid (3), and masseter (4) muscles; parotid gland (5); posterior belly of the digastric muscle (*black arrow*); sternocleidomastoid muscle(s); retromandibular vein (*open arrow*); external jugular vein (*black and white arrow*); external (e) and internal (i) carotid arteries; and internal jugular vein (j). *E*, This section is 3 mm superior to *D* and passes through the upper portion of the tongue and shows the uvula (U), palatine tonsil (T), oropharynx (P), and posterior belly of the digastric (*black arrow*), stylohyoid (*open arrow*), and stylopharyngeus (*black arrowheads*) muscles. The styloglossus and stylopharyngeus muscles are seen medial to the stylohyoid muscle and lateral to the internal carotid artery (I). Note the palatoglossus muscle (*white arrow*), parotid gland (PG), spinal accessory nerve (*white arrowhead*), superior longitudinal muscle fibers of the tongue (1), transverse muscle fibers of the tongue (2), buccinator muscle (3), masseter muscle (4), internal pterygoid muscle (5), pharyngeal constrictor muscle (*small black and small white arrows*), retromandibular vein (*black and white arrow*), and internal jugular vein (J). *F*, This section is 3 mm superior to *E* and shows the superior longitudinal muscle fibers of the tongue (1); transverse muscle fibers of the tongue (2); buccinator muscle (3); orbicularis oris muscle (4); masseter muscle (5); internal pterygoid muscle (6); parotid gland (7); posterior belly of the digastric (8), sternocleidomastoid (9), splenius capitis (10), and superior pharyngeal constrictor (*small white and small black arrows*) muscles; uvula (u); pharynx (P); parapharyngeal space fat (Ps); external (e) and internal (i) carotid arteries; internal jugular vein (J); anterior facial vein (*black arrow*); intraparotid facial nerve (*curved arrow*); retromandibular vein (*black arrowhead*); and stylopharyngeus (s) and stylohyoid (*open arrow*) muscles. *G*, This section is 3 mm superior to *F* and passes through the lower part of the alveolar process of the maxilla and just below the palatine process of the maxilla. The section passes through the upper part of the ramus of the mandible (R). The following structures are seen: quadratus labii superioris (1), buccinator (2), masseter (3), and internal pterygoid (4) muscles; parotid gland (5); sternocleidomastoid (6), splenius capitis (sc), digastric (7), rectus capitis lateralis (8), rectus capitis anterior (9), longus capitis (10), and palatopharyngeus (11) muscles; superior pharyngeal constrictor muscle (*small white and small black arrows*); styloid muscle group (*joined arrows*); platysma muscle (*white and black arrow*); anterior facial vein (*long black arrow*); retromandibular vein (*black arrowhead*); external (e) and internal (I) carotid arteries; internal jugular vein (J); and the mucous membrane of the hard palate, which shows the medial raphe (*curved arrow*) and numerous mucous glands, which are represented by small white areas. *H*, This is the T_2-weighted magnetic resonance image of *G*. Note the hypointensity of all muscles and hyperintensity of mucosal linings.

lymph node, bone, air, fluid, calcium, and blood) makes them the most complete technique for the diagnosis of pathologic conditions in these regions.[1–13] Branchial cleft cysts (Figure 59–9), ranula (Figure 59–10), cystic hygroma (Figure 59–11), thyroglossal duct cysts (Figure 59–12), dermoid and epidermoid cysts, vascular malformations, cellulitis, and abscesses are all well suited for MRI and CT.[1–11] Deep soft tissue edema and hematoma can be readily recognized in trauma patients. The total extent of deeply seated malignant tumors, associated lymphatic chain involvement, and vascular

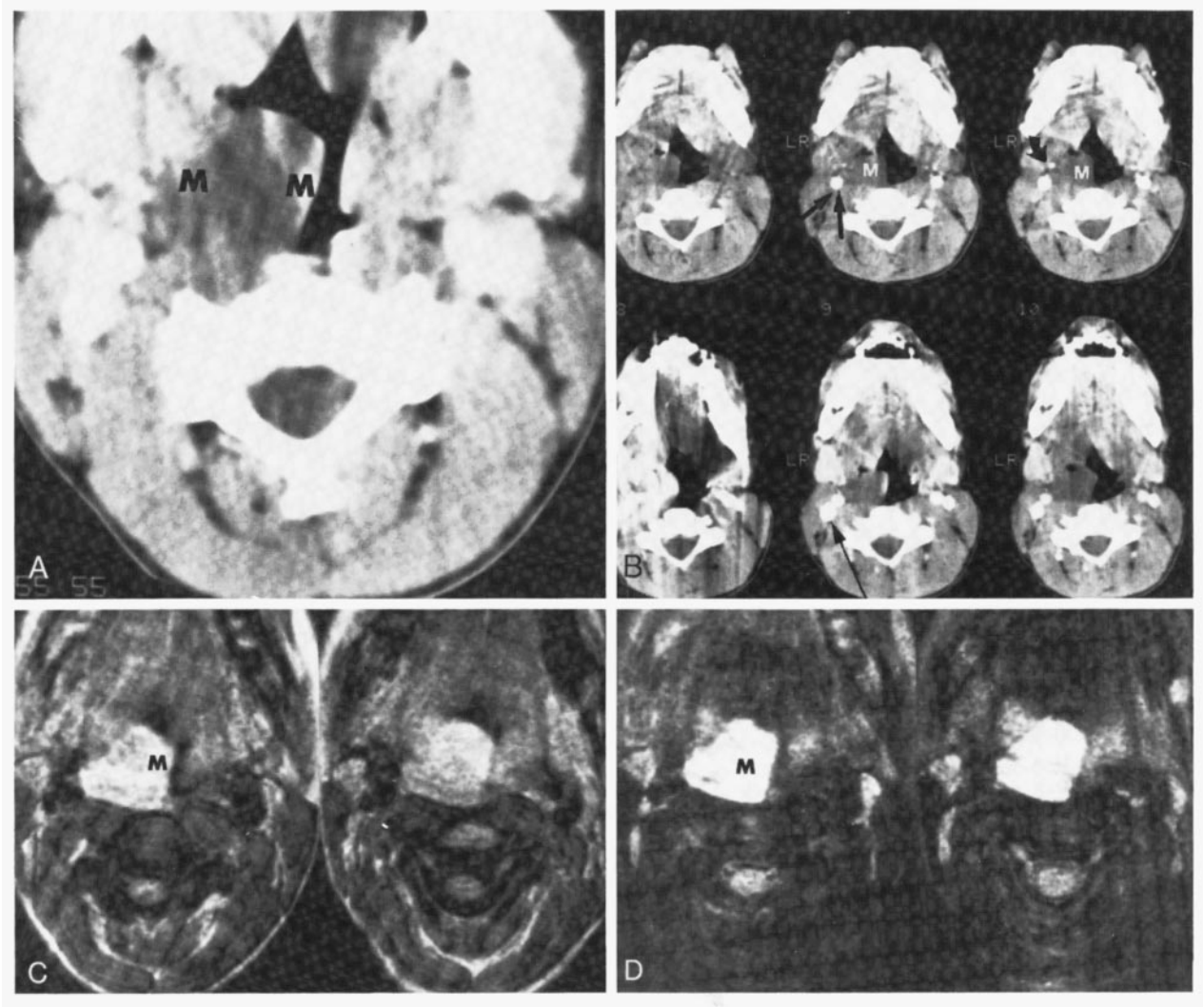


FIGURE 59–2. Cavernous lymphangiohemangioma of the oropharynx. *A*, Axial contrast computed tomographic (CT) scan shows a low-density mass (M) involving the right tonsil and posterolateral wall of the oropharynx. *B*, Serial dynamic CT scans show sequential enhancement of the external (*curved arrow*) and internal (*arrows*) carotid arteries and internal jugular vein (*long arrow*). Note that the mass (M) is not enhanced. Artifacts are seen in the fourth section. *C*, Proton-weighted magnetic resonance images (MRIs) showing the cavernous lymphangiohemangioma (M). *D*, T_2-weighted MRIs, showing the lymphangiohemangioma (M). Reproduced with permission from Mafee MF, Compos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.

encasement can be better determined with contrast-enhanced CT examination (see Figure 59–7).[1,5] The carotid body and glomus vagale tumors and neurogenic tumors, although uncommon, are considered frequently in the differential diagnosis of cervical masses.[8–11] With infusion CT, at times the differential enhancement of these lesions from adjacent muscles and vascular structures may not be sufficiently diagnostic to distinguish them from a vascular tumor (Figures 59–13 through 59–15), lymph node, or other abnormal solid lesion (see Figure 59–13).[7,8] Dynamic CT clearly visualizes the sequential enhancement of the carotid arteries and internal jugular veins, and the

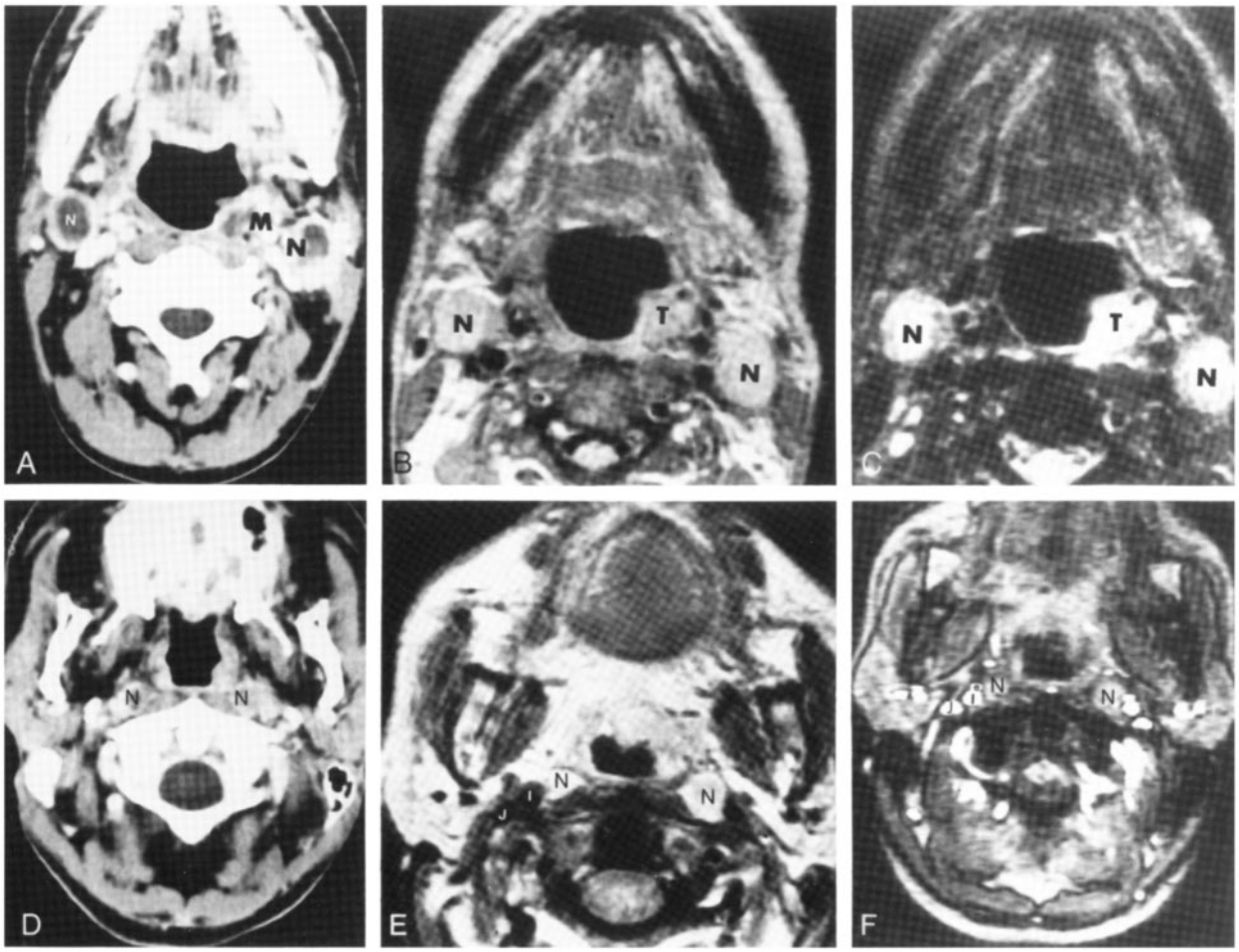


FIGURE 59–3. Squamous cell carcinoma of the left tonsil and soft palate. *A*, Contrast-enhanced computed tomographic (CT) scan shows a left tonsillar mass (M) and bilateral necrotic lymph nodes (N). Proton-weighted (*B*) and T_2-weighted (*C*) magnetic resonance images (MRIs) show tumor (T) and enlarged lymph nodes (N). The necrotic nature of the lymph nodes is better appreciated on CT scan (*A*). Notice that tumor and nodal metastases have identical signal characteristics in this patient. *D*, Axial CT scan obtained superior to *A* showing bilateral retropharyngeal lymph nodes (N). *E*, Proton-weighted MRI shows bilateral retropharyngeal lymph nodes (N). Note hyperintensity of all vascular structures. e = external carotid artery; I = internal carotid artery; J = internal jugular vein. Reproduced with permission from Mafee MF, Compos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.

tumor becomes densely opacified for a short time as the intravenously injected bolus of contrast material traverses the vascular bed of the tumor (see Figures 59–13 and 59–14). With MRI, glomus vagale and carotid body tumors (see Figures 59–14 and 59–15) can be distinguished from other lesions such as neurofibromas (Figure 59–16). Glomus complex tumors are highly vascular (Figure 59–17); therefore, on standard angiography and dynamic CT or MRA, this increased vascularity can be demonstrated, allowing differentiation of them from other lesions (see Figures 59–14 and 59–17).

Larynx and Hypopharynx

Computed tomography and MRI of the neck represent a major advance in laryngology.[14–26] The ability of CT to show accurately mucosa, deep laryngeal tissues, and laryngeal cartilages and the speed of new-generation spiral CT make it an ideal radiologic

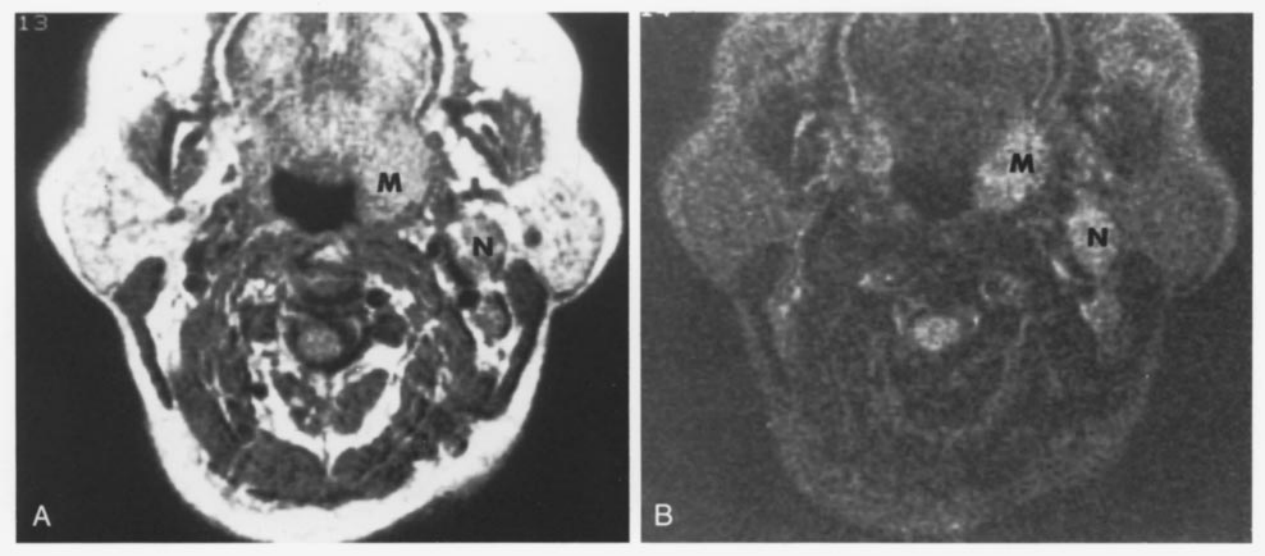


FIGURE 59–4. Non-Hodgkin's lymphoma of the tonsillar region. Proton-weighted (*A*) and T_2-weighted (*B*) axial magnetic resonance images show a tonsillar mass (M) and a deep enlarged lymph node (N). Reproduced with permission from Mafee MF, Compos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.

imaging method for evaluation of the larynx. The characteristic configurations of the hyoid bone and the thyroid, cricoid, and arytenoid cartilages greatly facilitate the orientation of CT scanning.[15,16,22] A section taken at the level of the hyoid bone (Figure 59–18) demonstrates the free portion of the epiglottis, valleculae, glossoepiglottic folds, pharyngoepiglottic folds, and superior portion of the piriform sinuses. In scans taken caudally, the preepiglottic space, epiglottis, aryepiglottic folds, piriform sinuses, thyroid cartilage, false cords, corniculate cartilages, true vocal cords, arytenoid cartilages, cricoid cartilage, cricothyroid articulation, conus elasticus, trachea, thyroid gland, and vascular structures can be visualized (Figures 59–19 through 59–24).

On CT scans, the preepiglottic space (see Figure 59–19) is seen anterior to the epiglottis as a low-density (lucent) image because of its high fat content.[15,16,22] On T_1-weighted MRI, this space is seen as a hyperintense image. The epiglottis is rarely calcified (Figure 59–25). Therefore, in an axial section, it is imaged as a sharply delineated anteriorly convex curved image, separating the preepiglottic space from the laryngeal vestibule (see Figure 59–25). The aryepiglottic folds make the posterolateral limit of the laryngeal vestibule[16] and separate the laryngeal vestibule from the posteriorly and laterally located piriform sinuses lying between the aryepiglottic folds and thyroid cartilage (see Figure 59–19). The ventricular folds (false vocal cords) are two thick folds of mucous membrane that contain both fat and minor salivary gland structures.[16] At the level of the false vocal cords, the airway is pear-shaped (see Figure 59–20). The soft tissues of the false cords are less dense on CT scans than the true vocal cords. The level of the false cords will always be apparent because of the shape of the airway and the fact that the corniculate and arytenoid cartilages (in their upper portions) are usually imaged with the false cords. At this level, the upper portion of the thyroid gland is imaged overlying the thyroid cartilage (see Figure 59–23). The remainder of the gland is seen on CT scans as a triangular, dense image (owing to iodine content) in the lower sections taken down to the upper tracheal rings. The junction of the false and true vocal cords (laryngeal ventricle) reveals a diamond-shaped contour if contiguous thin sections are obtained.[16]

True Vocal Cords and the Rima Glottidis When the transition is made to the glottis, the airway assumes a triangular appearance (see Figure 59–21). The rima glottidis is a fissure between the vocal

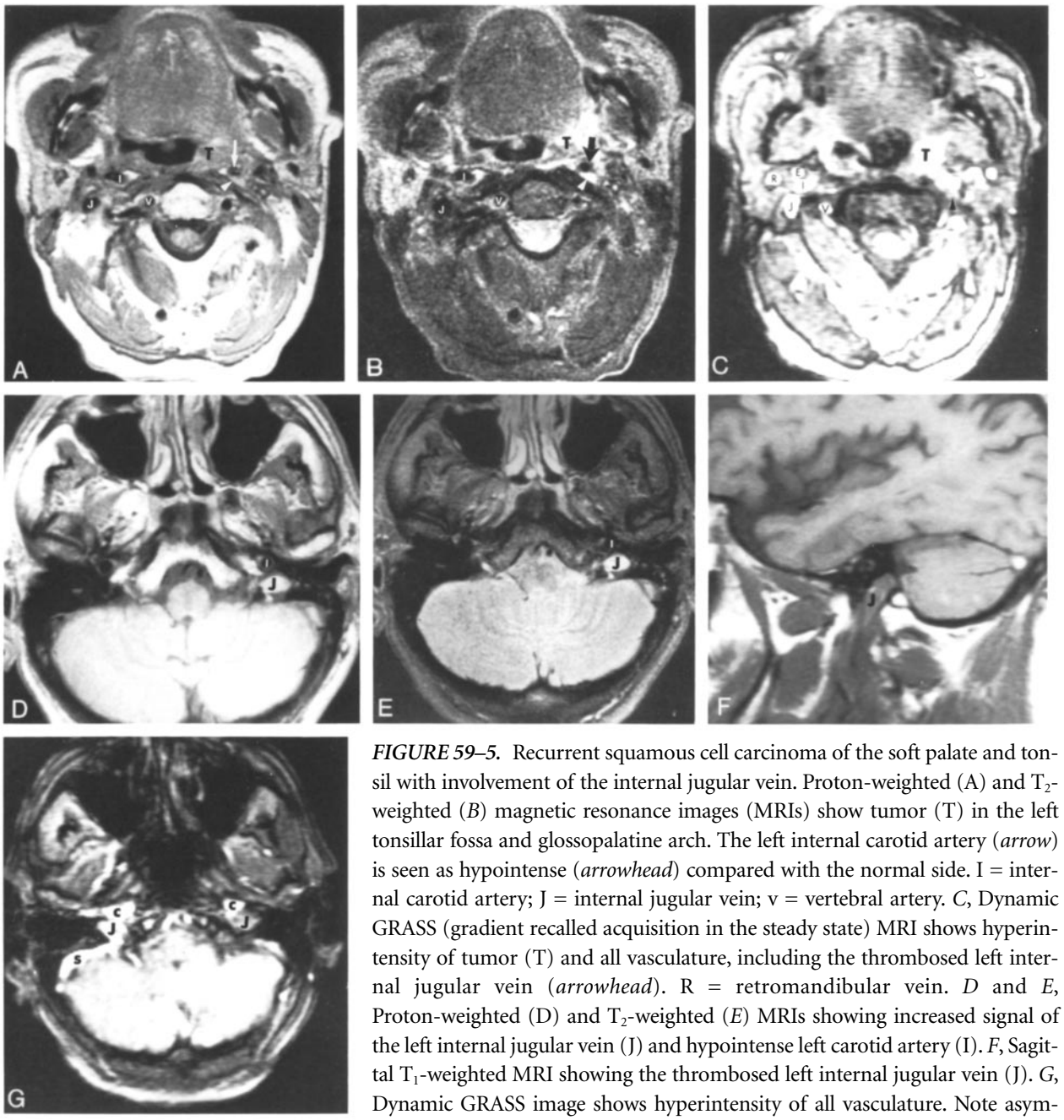


FIGURE 59–5. Recurrent squamous cell carcinoma of the soft palate and tonsil with involvement of the internal jugular vein. Proton-weighted (A) and T_2-weighted (*B*) magnetic resonance images (MRIs) show tumor (T) in the left tonsillar fossa and glossopalatine arch. The left internal carotid artery (*arrow*) is seen as hypointense (*arrowhead*) compared with the normal side. I = internal carotid artery; J = internal jugular vein; v = vertebral artery. *C*, Dynamic GRASS (gradient recalled acquisition in the steady state) MRI shows hyperintensity of tumor (T) and all vasculature, including the thrombosed left internal jugular vein (*arrowhead*). R = retromandibular vein. *D* and *E*, Proton-weighted (D) and T_2-weighted (*E*) MRIs showing increased signal of the left internal jugular vein (J) and hypointense left carotid artery (I). *F*, Sagittal T_1-weighted MRI showing the thrombosed left internal jugular vein (J). *G*, Dynamic GRASS image shows hyperintensity of all vasculature. Note asymmetry of signal of internal jugular vein (J). c = internal carotid artery; s = sigmoid sinus. Reproduced with permission from Mafee MF, Compos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.

cords and includes the vocal process of the arytenoid cartilages. It is limited dorsally by the mucous membrane, passing between the arytenoid cartilages at the level of the true vocal cords (see Figure 59–21).[16,20] The true vocal cords can be seen arising from the vocal process of the arytenoids and extending to the midpoint of the thyroid cartilage as they attach to the anterior commissure region of the larynx immediately subjacent to the point fusion of the paired thyroid alae (see Figure 59–21).[16,20] The true

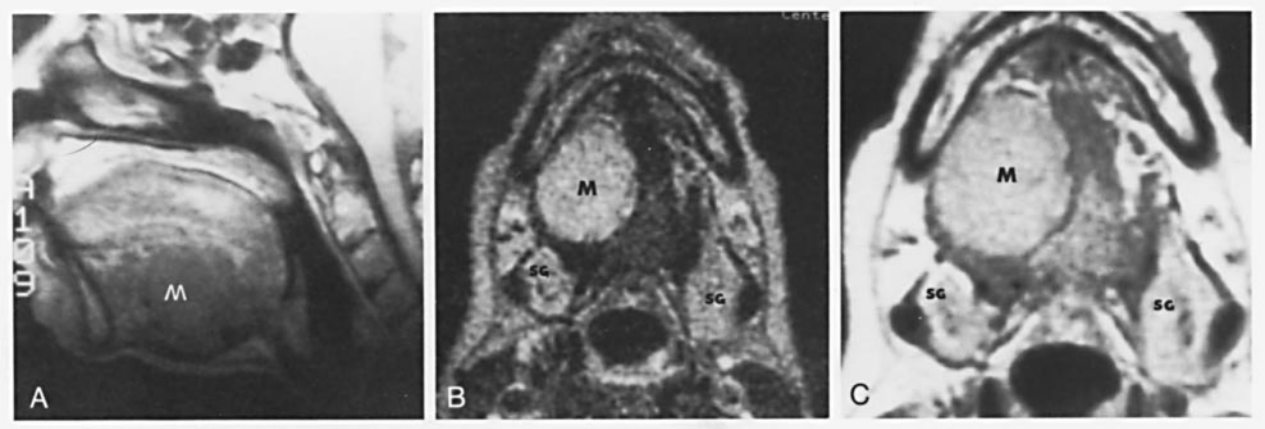


FIGURE 59–6. Pleomorphic adenoma of right submandibular region. *A*, T_1-weighted sagittal scan shows a hypointense mass (M) in the submandibular region. *B*, Proton-weighted magnetic resonance image (MRI) showing a mass (M) of intermediate signal intensity, which is isointense to the submandibular gland (SG). *C*, T_2-weighted MRI showing the mass (M), which is slightly hyperintense to normal submandibular gland (SG). This tumor was thought to arise from a minor salivary gland of the floor of the mouth. Reproduced with permission from Mafee MF, Compos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.

vocal cords have a dense vocal ligament along their free margin (see Figure 59–21, C).[16,20] At its widest, the true vocal cord measures 5 mm in thickness[16] and tapers to a thickness of 1 to 2 mm anteriorly where it meets with the contralateral true vocal cord to form the anterior commissure (see Figure 59–21). If the neck is not positioned properly, slight asymmetry of the true vocal cords will be present (see Figure 59–22).

The anatomy of the rima glottidis and its function can be evaluated by obtaining CT scans and MRIs during quiet breathing, forced inspiration,

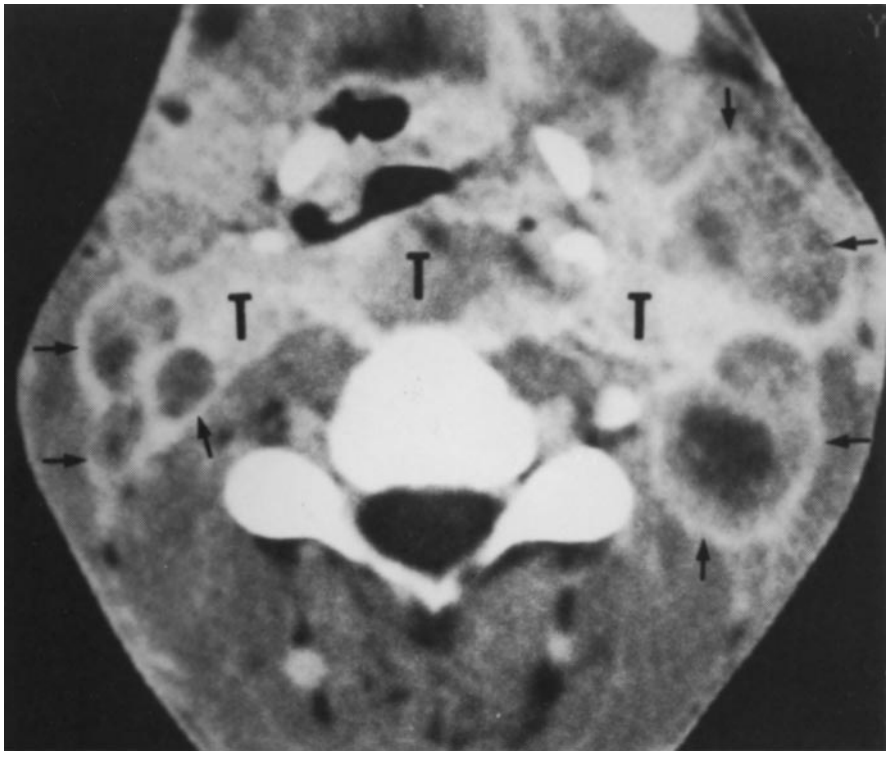


FIGURE 59–7. Pharyngeal carcinoma. Postinfusion contrast-enhanced computed tomographic scan showing a large enhanced, partially necrotic (low-density areas) tumor (T) involving the posterior pharyngeal wall, left tonsillar fossa, left side of the epiglottis, and left side of the base of the tongue, extending beyond the midline, involving the right side of the upper part of the hypopharynx. Note multiple, well-circumscribed ring-like enhancing lesions caused by necrotic lymph nodes (*arrows*). The carotid sheath vessels are engulfed by the metastatic lymph nodes at this level.

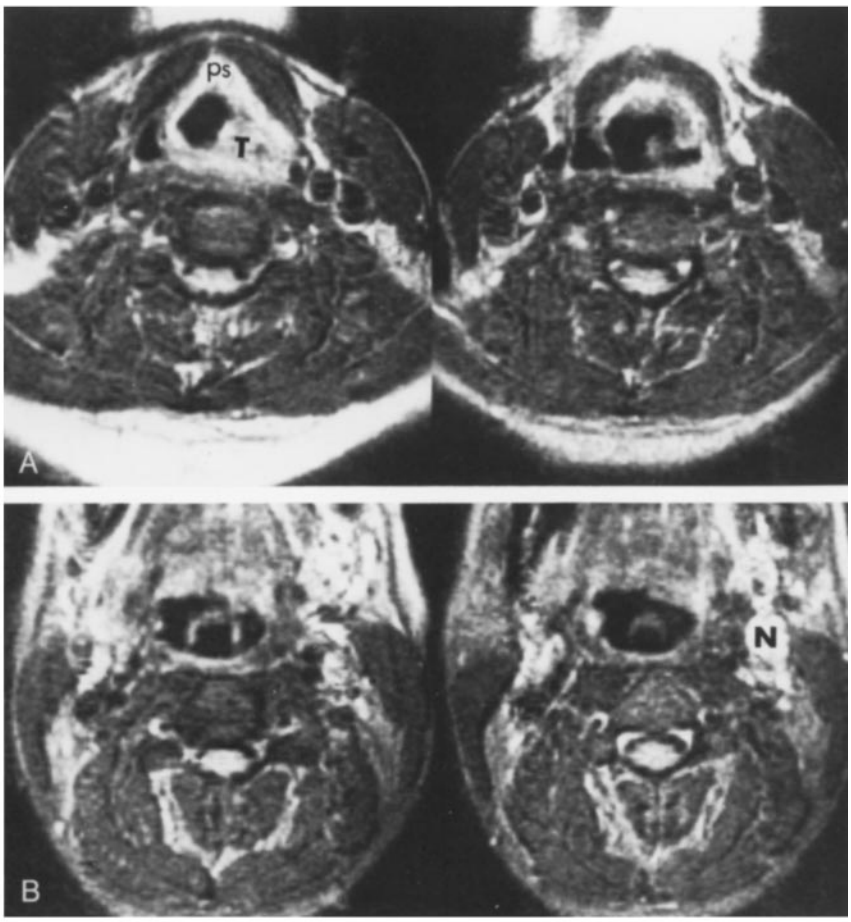


FIGURE 59–8. Squamous cell carcinoma of the piriform sinus. *A*, Proton-weighted magnetic resonance image (MRI) showing the tumor (T) and the fatty-areolar tissue in the preepiglottic space (PS). As seen, the tumor boundaries cannot be distinguished from normal fatty tissues in these MRIs. T_2-weighted MRI is needed for the distinction. *B*, Proton-weighted MRIs obtained superior to *A* showing enlarged lymph node (N).

phonation, and Valsalva's and modified Valsalva's maneuvers.[16,20] The rima glottidis has two parts: intermembranous, the ventral three-fifths between the cords (glottis vocalis), and the intercartilaginous, the dorsal two-fifths between the arytenoid cartilages (glottis respiratorial). It is the narrowest part of the larynx, but its width and shape vary with the movements of the vocal cords and arytenoid cartilages during respiration and phonation.[20] In a state of rest (eg, in quiet respiration), the intermembranous part of the rima is triangular (see Figure 59–21), its apex is ventrally located at the anterior commissure, and its base is represented by a line between the arytenoids (see Figure 59–22).[16,20] The intercartilaginous part is somewhat semicircular in shape (see Figure 59–21) and should be smooth and symmetric on a CT scan and MRI.[16,22] During closure of the rima glottidis (eg, when Valsalva's or modified Valsalva's maneuver is employed), both the true vocal cords and the arytenoid cartilages are adducted (see Figure 59–21, D). During force inspiration, the vocal cords undergo extreme abduction; the arytenoid cartilages are rotated laterally, and their vocal processes move widely apart. During phonation, the intermembranous and intercartilaginous space of the glottis is reduced owing to adduction of the vocal cords and adduction and medial rotation of the arytenoid cartilages (see Figure 59–21).[16,20]

Subglottis and Cricoid Cartilage The subglottis is well imaged by CT and MRI. The undersurface of the true vocal cords forms the superior limits of the subglottis (see Figure 59–22). The cricoid cartilage and the membrane and ligaments that attach to it make up the remaining boundaries of the subglottic space.[16,20] It must be remembered that the cricoid cartilage, which has a signet-ring shape, slopes inferiorly as it is followed anteriorly. The cricothyroid distance approximates 1.5 cm anteriorly, although the cricoid rises to the level of the vocal cords at the cricoarytenoid articulations posteriorly.[16] This is a critical transition point as anterior soft tissue exten-

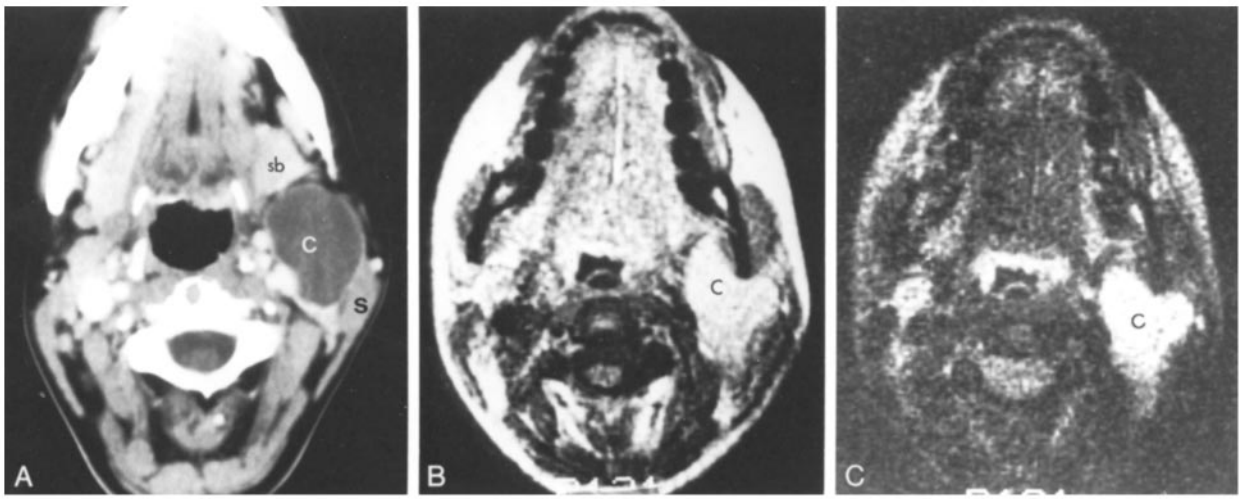


FIGURE 59–9. Second branchial cleft cyst. *A*, Axial computed tomographic scan showing the branchial cyst (c), sternocleidomastoid muscle (S), and submandibular gland (sb). *B*, Proton-weighted axial magnetic resonance image (MRI) showing the cyst (c). *C*, T_2-weighted MRI showing the cyst (c) as a hyperintense image.

sion of glottic carcinoma in the subglottis does not produce cartilage involvement until it has made the full inferior infiltration to the level of the cricoid cartilage. However, posterior extension of glottic cancer provides easy access to the high-rising cricoid signet. When the transition is made to the subglottis, the airway first assumes an ovoid and then a circular appearance (see Figures 59–22 and 59–23). The mucosa in the subglottic region is thin and should be smooth and symmetric in all sections. Increased soft tissue thickening in this area always suggests the presence of some abnormality.

Measurements of the transverse and anteroposterior diameters of the subglottis and the postcricoid-prevertebral, soft tissue thickness in 20 individuals (26 to 73 years of age) who were scanned for possible herniated disk or cervical cord lesion have been made. At the level of the undersurface of the true vocal cords and at the widest area (at the level of the posterior undersurface of the true vocal cords), the transverse diameter measured 1.38 to 1.58 cm, with an average of 1.47 cm. The anteroposterior diameter measured 2.20 to 2.40 cm, with an average of 2.24 cm. The prevertebral soft tissue (from the anterior margin of the vertebra to the posterior margin of the cricoid lamina) measured 0.82 to 0.96 cm, with an average of 0.87 cm. The transverse and anteroposterior diameters of the subglottic region, measured in a section obtained at the cricothyroid articulation, were as follows: the transverse diameter, at the widest area (midportion), measured 1.60 to 1.76 cm with an average of 1.64 cm, and the anteroposterior diameter measured 2.08 to 2.36 cm with an average of 2.12 cm. The prevertebral soft tissue measured 0.70 to 0.98 cm, with an average of 0.85 cm. When the transition is made to the trachea, the airway shows a posterior indentation because of the esophagus, which appears as a round soft tissue image posterior to the trachea (Figure 59–26).

The calcification and ossification of the thyroid cartilage may be occasionally uniform; therefore, on CT, its boundaries are delineated clearly (see Figure 59–21). Frequently, however, its calcification and ossification are irregular and asymmetric. Therefore, care must be taken in evaluating thyroid cartilage involvement by CT.[14–16] The paralaryngeal fat planes marginate the internal cortices of the thyroid cartilage (see Figure 59–22) and are seen as low-density fatty regions that are actually inferior and lateral extensions of the preepiglottic space (see Figure 59–22).[16,20] Obliteration of the paralaryngeal fatty planes usually indicates tumor infiltration.[20] The cricoid cartilage (see Figure 59–23), however, is frequently well calcified and its boundaries clearly delineated (see Figures 59–22 and 59–23). Characteristically, it has a dense cortex and a lucent medullary portion (see Figures 59–22

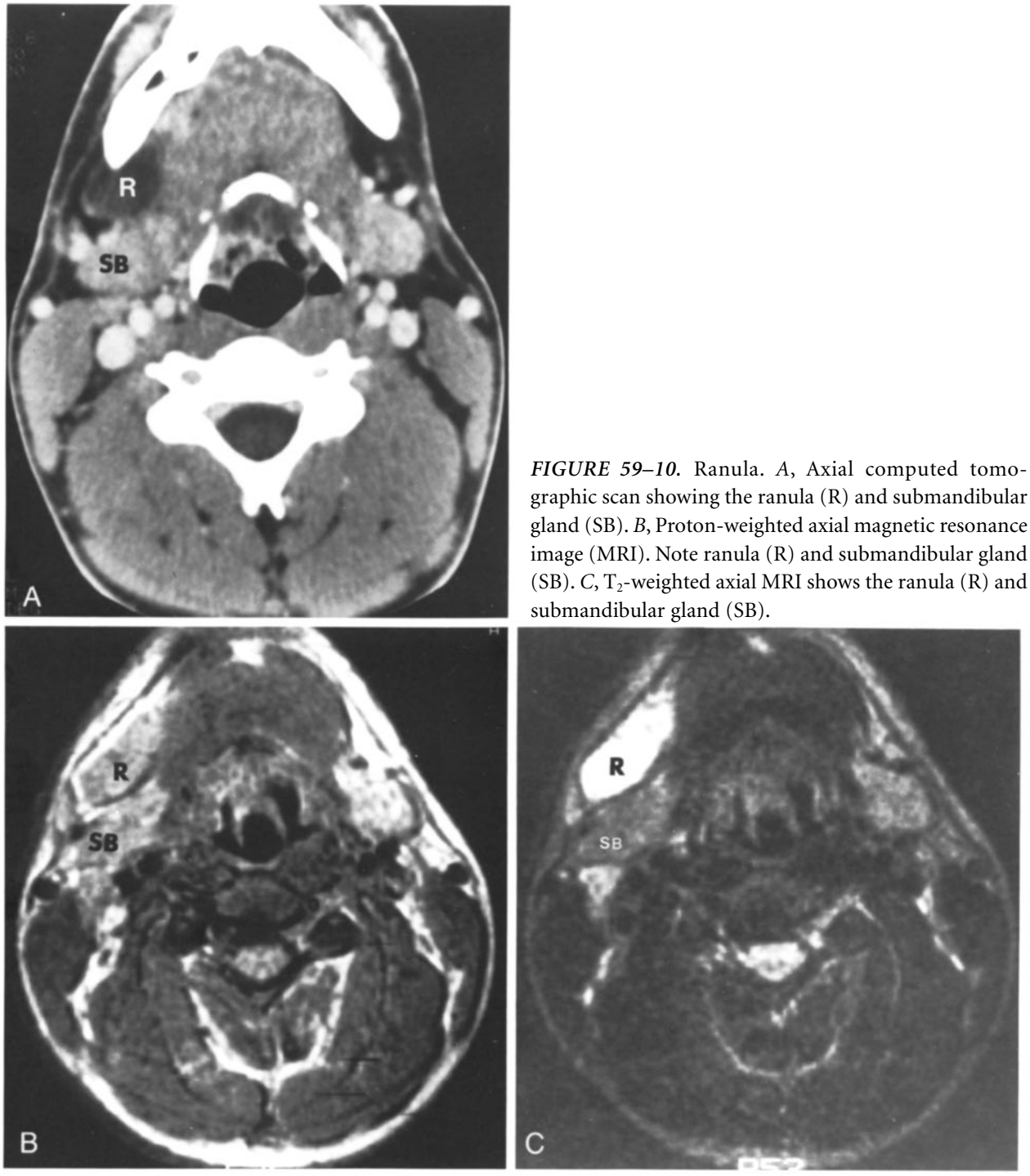


FIGURE 59–10. Ranula. *A*, Axial computed tomographic scan showing the ranula (R) and submandibular gland (SB). *B*, Proton-weighted axial magnetic resonance image (MRI). Note ranula (R) and submandibular gland (SB). *C*, T_2-weighted axial MRI shows the ranula (R) and submandibular gland (SB).

and 59–23). The arytenoid cartilages are frequently calcified and often appear symmetric (see Figure 59–21). Asymmetric calcification of the arytenoid cartilages should not be mistaken for tumor involvement.[16,20,25,26] Increased density (sclerosis) of arytenoid cartilage adjacent to laryngeal carcinoma may be related to tumor involvement; however, it not infrequently represents inflammatory reaction rather than tumor involvement.[25,26]

Magnetic resonance imaging, because of its superior spatial resolution and three-dimensional capability, is superior to CT for delineating the

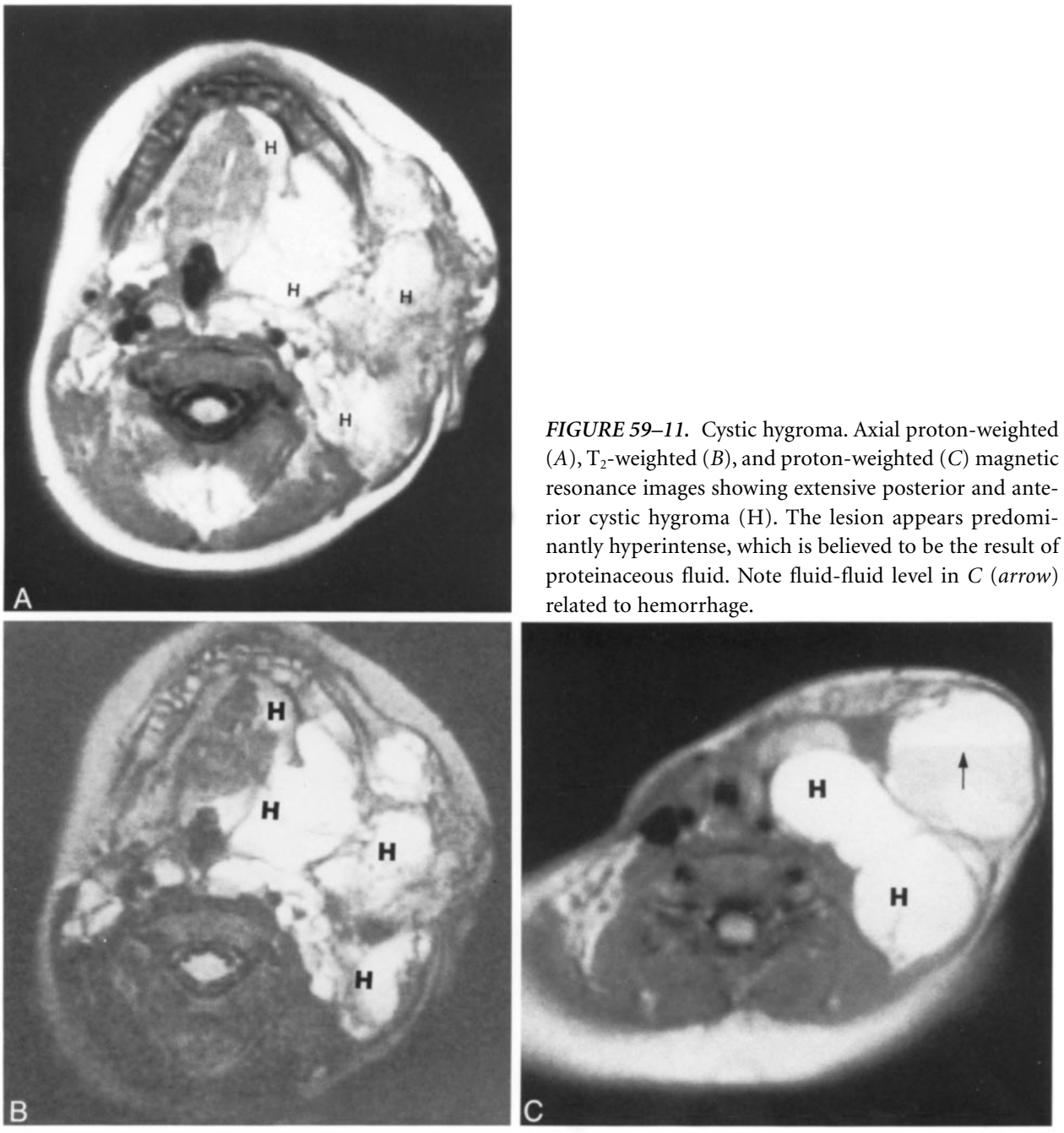


FIGURE 59–11. Cystic hygroma. Axial proton-weighted (*A*), T_2-weighted (*B*), and proton-weighted (*C*) magnetic resonance images showing extensive posterior and anterior cystic hygroma (H). The lesion appears predominantly hyperintense, which is believed to be the result of proteinaceous fluid. Note fluid-fluid level in *C* (*arrow*) related to hemorrhage.

anatomy and pathology of the larynx (see Figure 59–25). Magnetic resonance imaging provides maximum information on the extent of the tumor. The disadvantages of MRI are as follows: (1) MRI frequently requires 45 to 60 minutes scanning time, during which time any patient motion, including heavy breathing, swallowing, and coughing, can degrade the image; (2) MRI is contraindicated in patients with cardiac pacemakers, cochlear implants, and those with ferromagnetic objects in the body; and (3) MRI is, unfortunately, the most expensive of all diagnostic imaging modalities.

Laryngeal Pathology Computed tomography has made a major contribution to the radiologic evaluation of the largynx.[14–23] It is the procedure of choice in the investigation of laryngeal trauma,[19] especially in the acute phase, because it

provides evidence of mucosal disruption, hematoma formation, fracture and cartilage displacement, and an axial display of the airway[19] (Figures 59–27 and 59–28). Benign lesions of the larynx are generally diagnosed by clinical examination rather than radiologic techniques. Laryngeal cysts and laryngoceles are well suited for CT and MRI. Computed tomography is able to define the extent of the laryngocele more precisely than either clinical examination or conventional radiographic techniques. Using CT, a laryngocele appears as a well-circumscribed air-filled enlargement that represents extension and elongation of the laryngeal appendices (saccules) of the laryngeal ventricle (Figure 59–29). When the laryngocele is filled with fluid (complicated laryngocele),

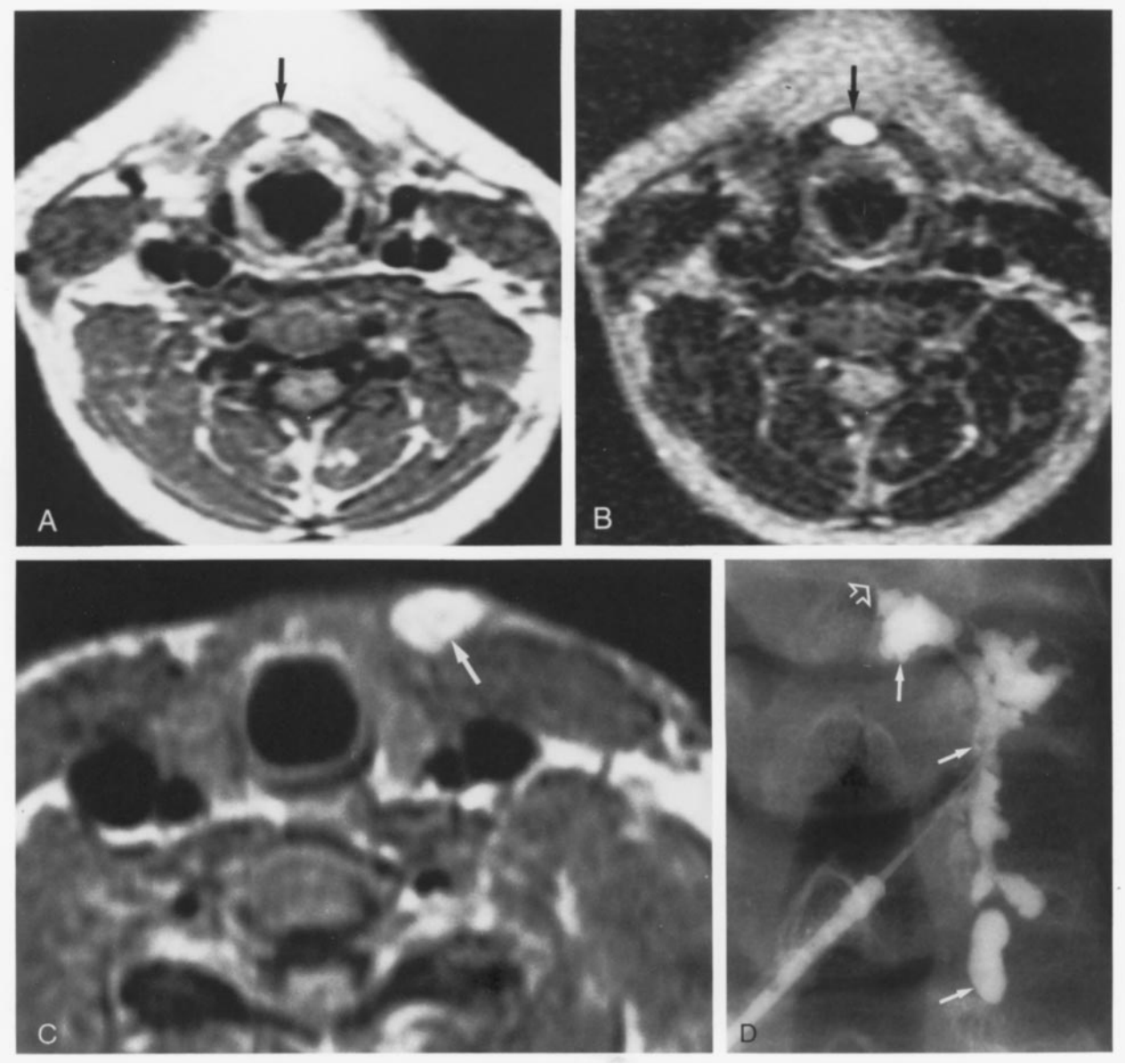


FIGURE 59–12. Thyroglossal duct cyst. Proton-weighted (*A*) and T_2-weighted (*B*) magnetic resonance images (MRIs) at the level of the hyoid bone and (*C*) proton-weighted MRI at the level of the midportion of the cricoid cartilage showing a hyperintense cyst (*arrow*). *D*, Fistulogram of the same patient as in *A* to *C* following injection of iodinated contrast in the fistula demonstrating a branched thyroglossal duct cyst (*arrows*), which extends from the base of the tongue (*hollow arrow*) down to the laryngotracheal level.

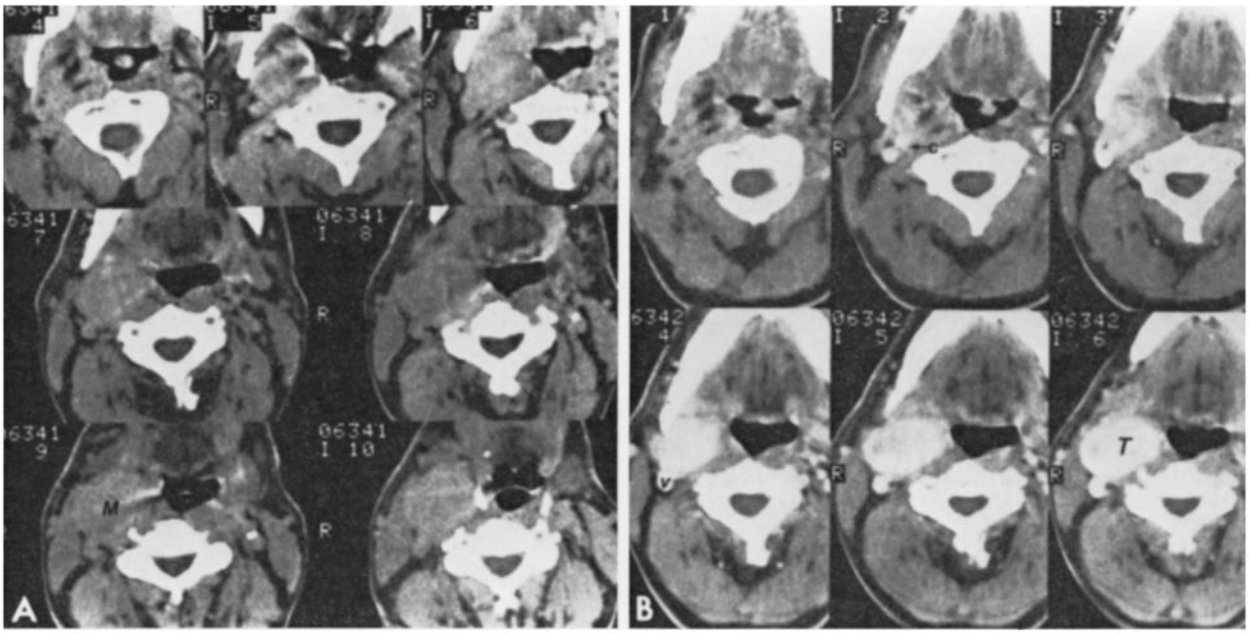


FIGURE 59–13. Carotid body tumor. Contiguous computed tomographic (CT) sections obtained following contrast infusion technique. *A*, A mass (M) is seen behind the mandible in the third section and can be followed down to the last section obtained at the level of the hyoid bone. *B*, Contiguous axial CT sections obtained with dynamic technique. Note identification of the internal carotid artery (c) and internal jugular vein (v) and intense enhancement of the carotid body tumor (T). Reproduced with permission from Mafee MF, Valvassori GE, Shugar MA, et al. High resolution and dynamic sequential computed tomography. Use in the evaluation of glomus complex tumors. Arch Otolaryngol 1983;109:691–6.

it appears on CT and MRI as a smoothly marginated mass along the lateral wall of the larynx, and it may be difficult to differentiate it from an aryepiglottic fold cyst.

Malignant Tumors of the Larynx The vast majority of laryngeal and hypopharyngeal tumors are squamous cell carcinomas. Except for skin cancer, cancer of the larynx is the most common malignant tumor of the head and neck region. Only 5 to 10% of malignant neoplastic lesions of the larynx and hypopharynx are non–squamous cell tumors. Cancers of the larynx and hypopharynx are by far the most common indications for imaging of this region. Involvement of the preepiglottic space or paralaryngeal space and subglottic extension can be detected by both CT and MRI. Detection of cartilage invasion if it is less than 6 mm constitutes a serious challenge. Areas of increased cartilage sclerosis may be seen on CT scans adjacent to the tumor. This is most often related to inflammatory changes in the immediate tumor vicinity rather than tumor invasion. The ability of CT and MRI to demonstrate accurately mucosal, deep laryngeal tissue and associated lymphatic chain involvement makes either of them an ideal imaging method for evaluation of laryngeal tumors (Figures 59–30 through 59–32).[15,20,21,25,26] Computed tomography and MRI have their major value in the pretreatment tumor, node, metastasis staging of laryngeal cancer.[20,25] Tumor involvement of the anterior and posterior commissures (see Figure 59–32, C), subglottic region (see Figures 59–32, E, and 59–32, F), and trachea is best evaluated by CT. Cartilage involvement (see Figure 59–30) and extension of the tumor in the paralaryngeal and paraglottic spaces can be demonstrated in a manner not available before the introduction of CT. Computed tomography is very helpful in the evaluation of tumor extension, specifically in the axial dimension.[16,20] Its primary application is in the evaluation of T3 and T4 laryngeal cancer.[20,21] Magnetic resonance imaging of the larynx can provide valuable information (see Figure 59–32).[25,26] With

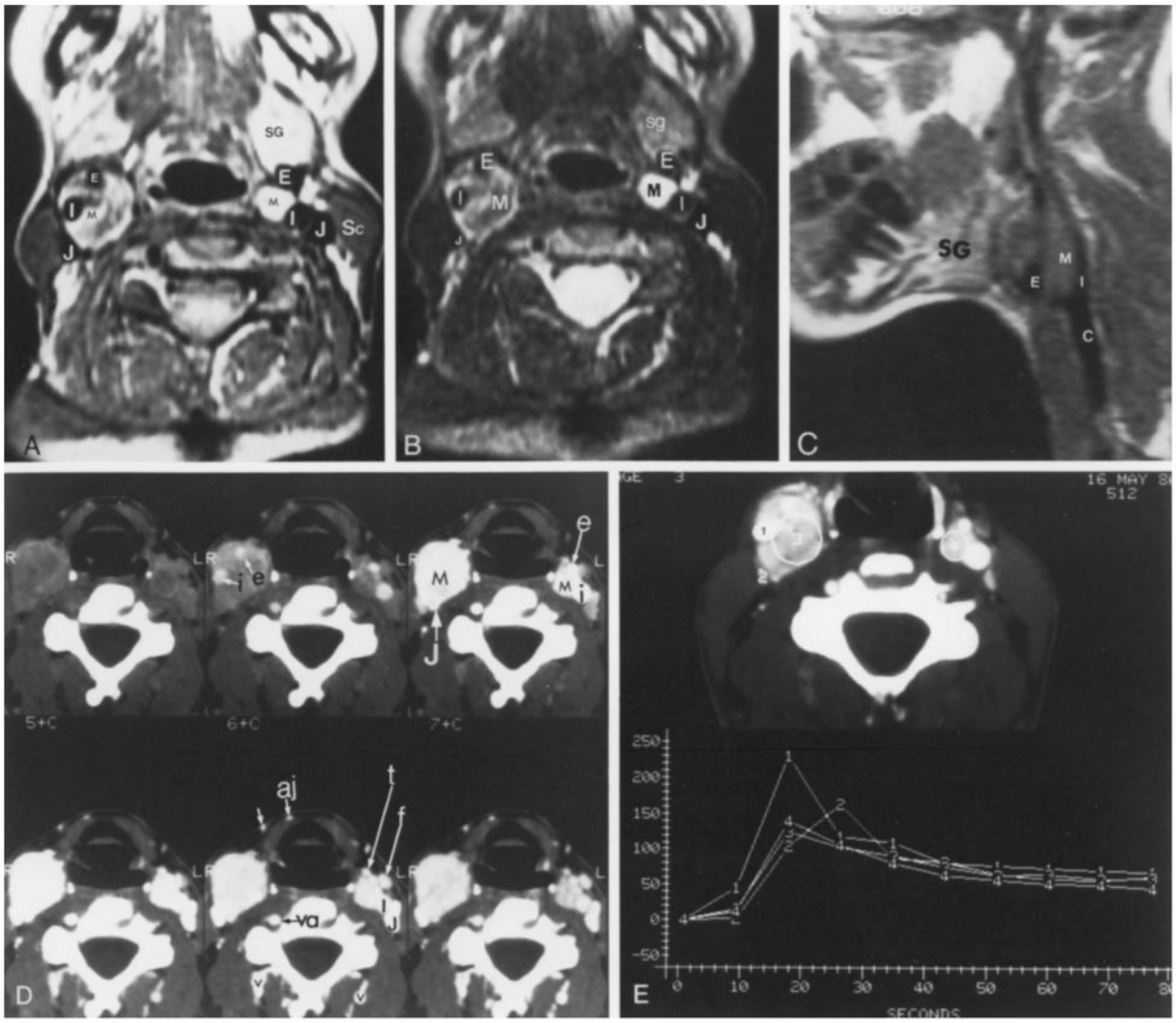


FIGURE 59–14. Bilateral carotid body tumors. Proton-weighted (*A*) and T_2-weighted (*B*) axial magnetic resonance images (MRIs) showing bilateral masses (M), submandibular gland (SG), sternocleidomastoid muscle (Sc), external (E) and internal (I) carotid arteries, and internal jugular vein (J). *C*, T_1-weighted sagittal MRI showing the right carotid body mass (M); common (C), external (E), and internal (I) carotid arteries; and submandibular gland (SG). *D*, Dynamic computed tomographic scans showing the sequential enhancement of external (e) and internal (i) carotid arteries and internal jugular vein (J). Notice intense enhancement of the carotid body tumors (M). Note vertebral artery (va), deep paravertebral venous plexus (v), facial vein (f), superior thyroid artery (t), anterior jugular vein (aj), and anterior facial vein (*white arrow*). *E*, Density-time curve shows characteristic arterial peak of carotid body tumors (3, 4). 1 = common carotid artery; 2 = internal jugular vein. Reproduced with permission from Mafee MF, Compos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.

MRI, however, motion artifact related to respiration, swallowing, and blood flow can cause significant distortion of the images.

Non–Squamous Cell Tumors of the Larynx and Hypopharynx A variety of rare benign and malignant tumors of epithelial, neuroectodermal, and

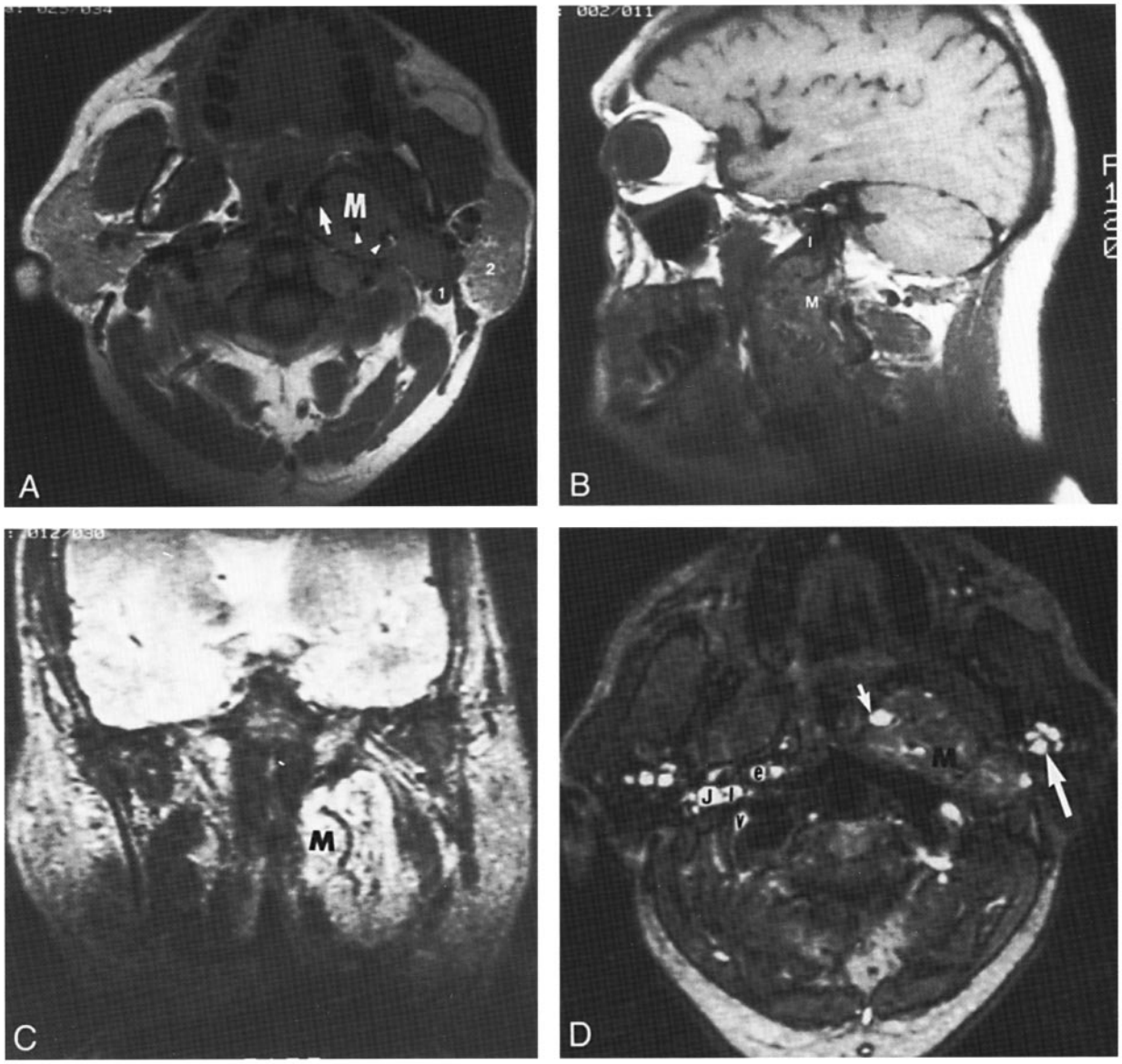


FIGURE 59–15. Carotid body tumor. *A*, Proton-weighted axial magnetic resonance image (MRI) shows a large mass (M) displacing the internal carotid artery (*arrow*) toward the parapharyngeal space. Note prominent vessels within the mass (*arrowheads*). 1 = posterior belly of the digastric muscle; 2 = parotid gland. *B*, Sagittal T_1-weighted MRI showing the carotid body tumor mass (M). Note marked anterior displacement of the internal carotid artery (I) and serpiginous low-density vascular structures within the mass. *C*, T_2-weighted coronal MRI showing hyperintense mass (M), with several tortuous low-density images representing prominent vessels within the mass. The location of the mass as seen in sagittal and coronal MRIs is high, raising the possibility of glomus vagale rather than carotid body tumor. *D*, Dynamic GRASS (gradient recalled acquisition in the steady state) axial MRI showing the tumor (M). Note hyperintensity of arteries and veins in this section. e = external carotid artery; I = internal carotid artery; J = internal jugular vein; v = vertebral artery. Note retromandibular veins (*large arrow*) and displaced left internal carotid artery (small arrow). Reproduced with permission from Mafee MF, Compos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.

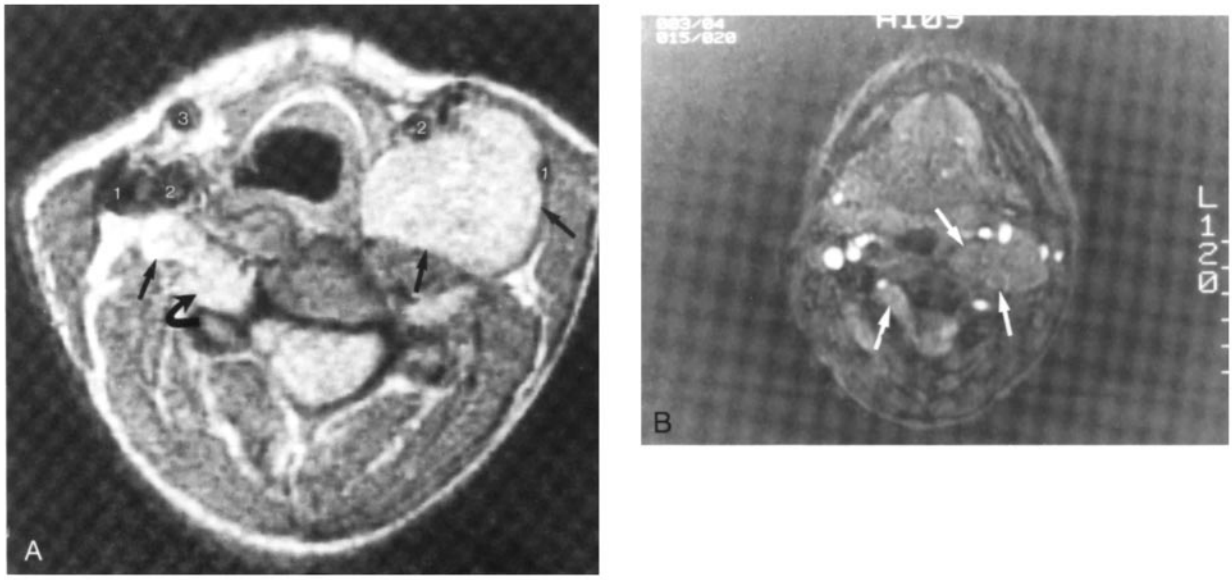


FIGURE 59–16. Neurofibroma. *A*, Proton-weighted magnetic resonance image (MRI) shows bilateral masses (*arrows*) in this patient with neurofibromatosis. On the left side, the tumor has widened the neuroforamen (*curved arrow*). 1 = internal jugular vein; 2 = internal carotid artery; 3 = anterior jugular vein. *B*, Dynamic GRASS (gradient recalled acquisition in the steady state) axial MRI shows the neurofibromas (*arrows*). Notice the hyperintensity of the vascular structures.

mesodermal origin may be seen in larynx and hypopharynx. These unusual tumors comprise about 2 to 5% of all laryngeal tumors.[2] These tumors include hemangiomas (Figure 59–33, A); chondrogenic tumors (chondromas, chondrosarcomas) (Figure 59–34); osteoblastomas; lymphomas, including plasmacytomas; minor salivary origin tumors (pleomorphic adenomas, adenoid cystic carcinomas); tumors of fatty tissue; neurogenic tumors[2,27]; myogenic tumors; fibrohistiocytic tumors; leiomyomas (Figure 59–35); paragangliomas; myxomas; synovial sarcomas; and metastases.

TRACHEA

Subglottic and tracheal airways can be evaluated by standard radiography, CT, and MRI. Subglottic tracheal stenosis can be best evaluated by CT and, in particular, spiral CT. Congenital anomalies and inflammatory, traumatic, and neoplastic conditions of the trachea can be evaluated by plain radiography, CT, and MRI.[28,29] Computed tomography is the study of choice for laryngeal and tracheal trauma. Magnetic resonance imaging is the study of choice for inflammatory and neoplastic conditions of the trachea (Figure 59–36). The extent of tracheal stenosis, tracheal invasion by thyroid carcinoma, and primary tumors of the trachea is best evaluated by MRI (Figure 59–37).

COMPUTED TOMOGRAPHY AND MAGNETIC RESONANCE IMAGING OF NECK LYMPH NODES

Computed tomography and MRI will permit the consistent detection of lymph nodes greater than 1.5 cm in diameter (see Figure 59–32, A and F),[20] and they are almost always tumor bearing. Lymph nodes between 1 and 1.4 cm may represent tumor or reactive hyperplasia.[30,31] Any modality that could show abnormal lymph nodes accurately would have great impact in the management of patients with head and neck cancers.[31,32] The treatment of patients at high risk but without clinical evidence of neck (lymph node) tumors has long been a source of confusion and controversy.[31,32]

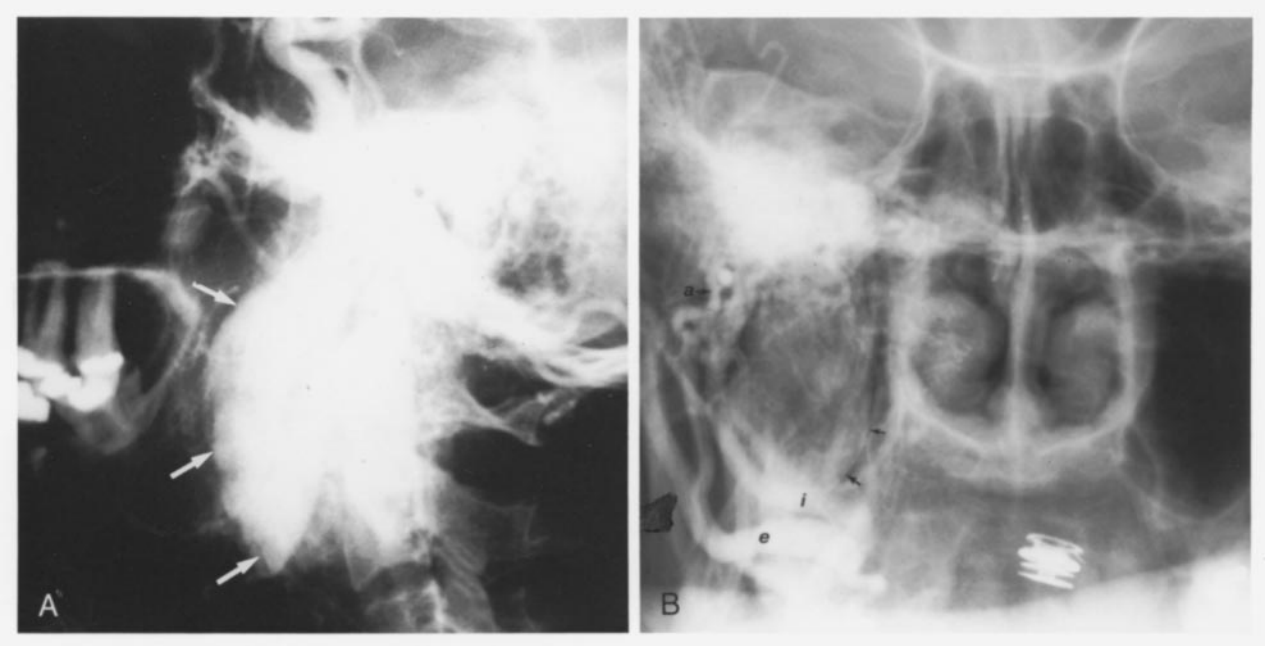


FIGURE 59–17. Carotid body tumor. *A*, Lateral angiogram shows marked tumor blush of a large carotid body tumor (*arrows*). *B*, Frontal angiogram of another patient with a large glomus jugulare tumor showing the tumor blush and involvement of the internal jugular vein (*arrows*). Notice the displaced internal (i) and external (e) carotid arteries. Notice also a large ascending pharyngeal artery (a).

The size of normal lymph nodes varies and may measure up to 15 mm in the upper deep cervical jugular chain, to which most head and neck tumors first metastasize.[32,33] Theoretically, lymph nodes become positive at a nondetectable size. As many as 1 billion cells are necessary to produce a mass of 1 cm.[3,34] The CT and MRI malignancy criteria should include the following: (1) lymph nodes of any size with clear evidence of central low density on CT or abnormal intensity on MRI, evidence of

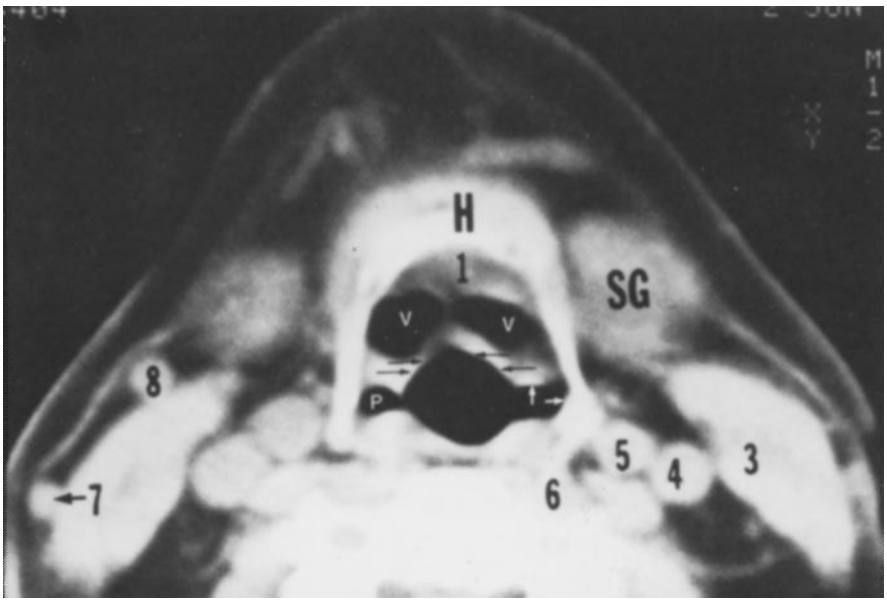


FIGURE 59–18. Axial computed tomographic scan of the oropharynx showing the hyoid bone (H), base of the tongue (1), and, posterior to that, the valleculae (V). Note the epiglottis (*black arrows*), pharyngoepiglottic folds (*white arrows*), upper part of the piriform sinuses (p), submandibular glands (SG), sternocleidomastoid muscles (3), internal jugular veins (4), common carotid arteries (5), longus colli muscles (6), and external jugular veins (7). The platysma muscle is overlying the lateral facial vein (8).

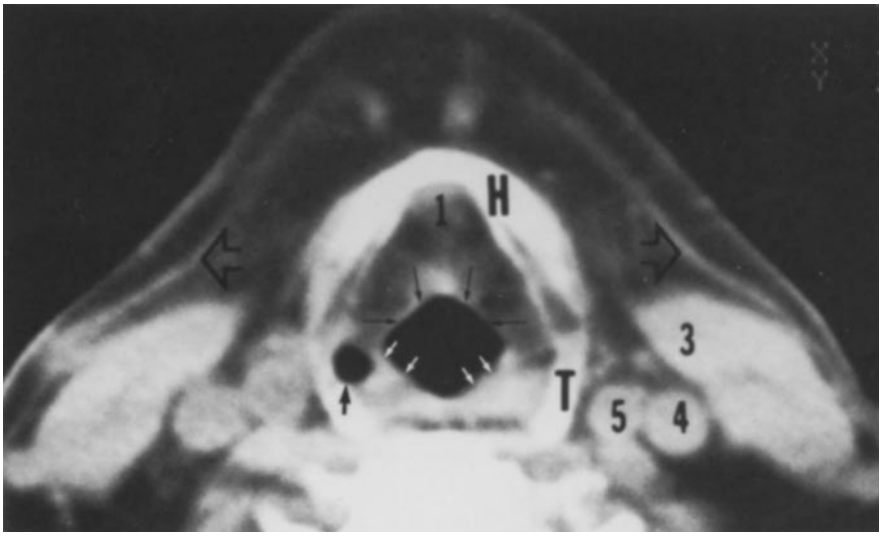


FIGURE 59–19. Axial computed tomographic scan of the larynx and hypopharynx at the level of the superior cornua of the thyroid cartilage (T), showing the inferior part of the hyoid bone (H), preepiglottic space (1), epiglottis (*black arrows*) and aryepiglottic folds (*white arrows*), platysma muscles (*open arrowheads*), sternocleidomastoid muscles (3), internal jugular veins (4), and common carotid arteries (5). Air is seen in the right piriform sinus (*arrow*).

central neurosis; (2) diameter of more than 15 mm for lymph nodes located in levels I and II and more than 10 mm for lymph nodes located in other levels; (3) groups of three or more borderline lymph nodes; (4) loss of tissue planes around any enlarged lymph node; (5) any size lymph node with calcification in a patient with papillary carcinoma of the thyroid gland, which should be considered metastatic unless proven otherwise; (6) any size lymph node in a patient with mucocutaneous malignant melanoma should be considered suspicious for metastases, particularly if it behaves differently on imaging from the contralateral side. Although CT and MRI provide improved clinical staging, they have major limitations. For example, small areas of necrosis may not be depicted clearly, particularly with MRI. In addition, adipose metaplasia in the hilum of a lymph node or an area of hypovascularity can mimic necrosis.[35] More importantly, as Crile pointed out almost 100 years ago, some large lymph nodes may be inflammatory, whereas nonpalpable lymph nodes may be carcinomatous.[36] Although CT and MRI have some limitations in detecting metastatic lymph nodal disease, they are valuable

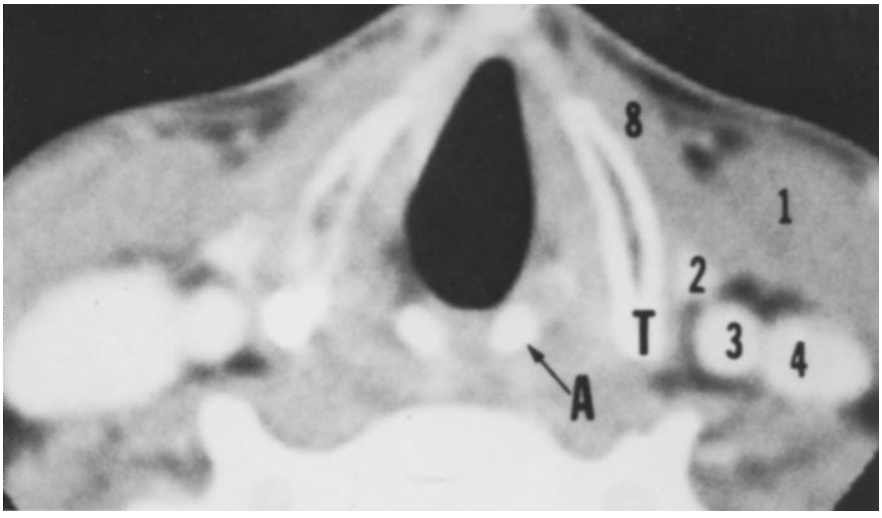


FIGURE 59–20. Axial computed tomographic scan of the larynx at the level of the false vocal cords showing the superior portion of the arytenoid cartilages (A), thyroid cartilage (T), thyroid gland (2), common carotid arteries (3), internal jugular veins (4), sternocleidomastoid muscles (1), and infrahyoid strap muscles (8). Note asymmetry of the internal jugular veins, with the right being larger than the left.

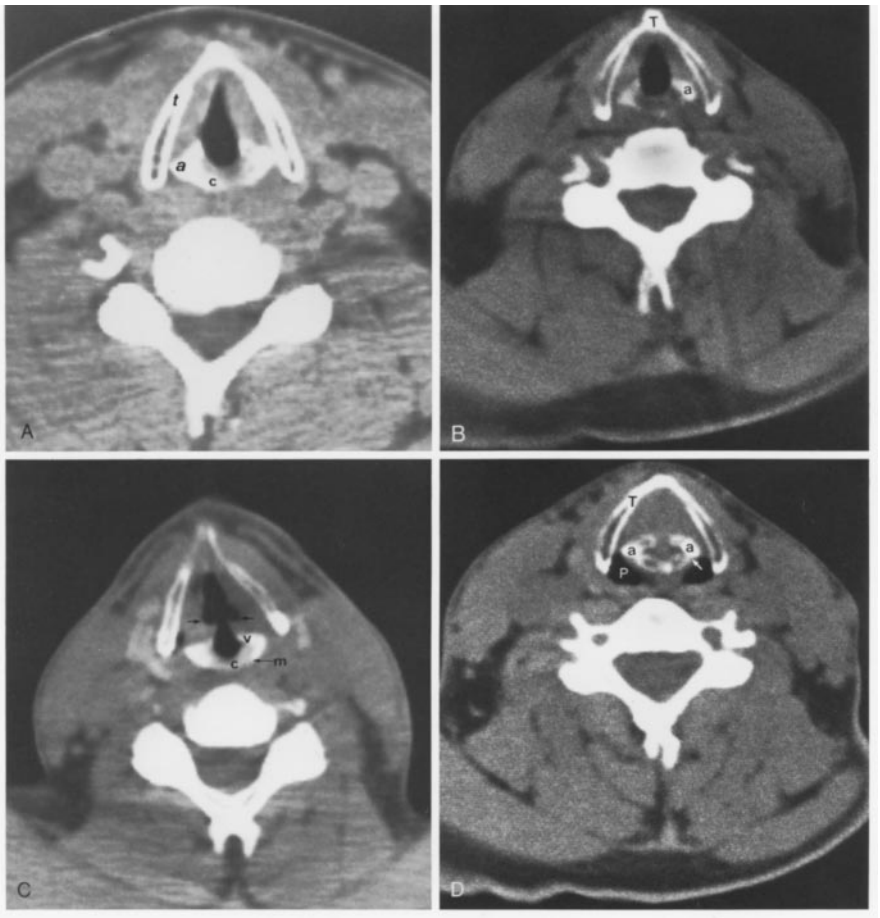


FIGURE 59–21. *A*, Rima glottidis in quiet respiration. Note the symmetry and smooth appearance of the true vocal cords and arytenoid cartilages (a). Notice that the widest part of the glottic aperture is opposite the vocal processes of the arytenoid cartilages. The intermembranous part of the rima glottidis is triangular, and the intercartilaginous part is semicircular in shape. Notice the symmetry in the distances between the thyroid (t) and arytenoid (a) cartilages. c = upper part of the cricoid cartilage. *B*, Normal true vocal cords. Axial computed tomographic (CT) scan obtained during quiet breathing, showing the thyroid cartilage (T) and arytenoid cartilages (a) and their vocal processes. Notice the slight asymmetry of the true vocal cords, which is caused by the asymmetric position of the cords in the section. *C*, Normal true vocal cords. A CT scan of the larynx during expiratory phonation showing the cricoid lamina (c) and the vocal (v) and muscular (m) processes of the arytenoid cartilages. The adducted true vocal cords and partially volumed laryngeal ventricles (*arrows*) are shown. *D*, Closure of the rima glottidis. This is the same patient and the same CT scan level as in *A*; however, during scanning, the patient was requested to hold his breath. Notice closure of the glottic aperture caused by approximation of the true vocal cords. The true vocal cords and arytenoid cartilages (a) are adducted. Notice rotation of the arytenoid cartilages as compared with *A*. The cricoarytenoid articulation (*arrow*) is demonstrated in this breath-holding scan. P = piriform sinus; T = thyroid cartilage. Reproduced with permission from Mafee MF, Schild JA, Valvassori GE, Capek V. Computed tomography of the larynx: correlated anatomy and pathologic macrosection studies. Radiology 1983;147:123–8.

adjuncts to clinical evaluation alone and especially valuable in those head and neck sites that are not accessible to palpation, such as retropharyngeal, intraparotid (deeply seated), parapharyngeal, paratracheal, and mediastinal nodes.[33,37,38] Some lymph nodes, such as the retropharyngeal (adults), intraparotid, suboccipital, periauricular, and superficial lymph nodes, particularly Delphian nodes, are so

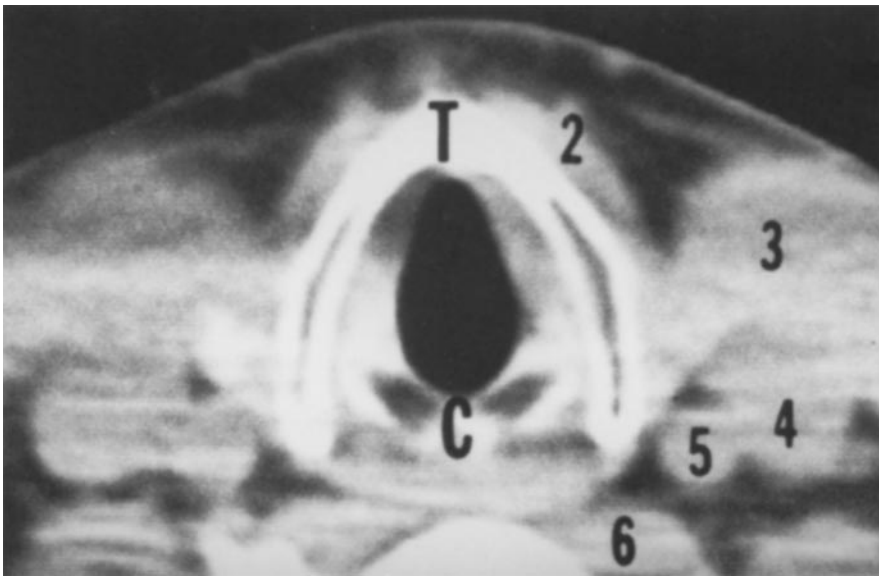


FIGURE 59–22. Normal axial computed tomographic scan at the level of the undersurface of the true vocal cords showing the thyroid cartilage (T), infrahyoid strap muscles (2), sternocleidomastoid muscles (3), internal jugular veins (4), common carotid arteries (5), longus colli muscles (6), and the cricoid cartilage lamina (C).

small that they are rarely detected on CT and MRI. Whenever these lymph nodes are detected by CT and MRI, they should be considered highly suspicious for pathologic lymph nodes.

Newer imaging modalities are currently under investigation and hopefully may further improve the diagnostic accuracy of metastatic lymph nodal disease. Positron emission tomography (PET) has been shown to be sensitive for detecting lymph node metastasis with a higher accuracy rate than MRI.[37–41] The MRI has been widely used for detection of metastatic lymph nodes. The signal intensity of metastatic lymph nodes often does not differ from that of normal lymph nodes. A central necrosis is seen as an area of low signal on T_1-weighted and high intensity on T_2-weighted MRIs. At times, central necrosis may not be appreciated on MRI. It is not unusual to find borderline-sized lymph nodes without central necrosis in CT or MRI. Recent development in MRI contrast material and new contrast agents, such as superparamagnetic iron oxide, has increased the accuracy of MRI.[42,43] Micrometastases are unlikely to be detected by MRI or CT. Localization and imaging of radiolabeled monoclonal antibodies for the detection of cervical metastasis have been successfully investigated in mice.[44]

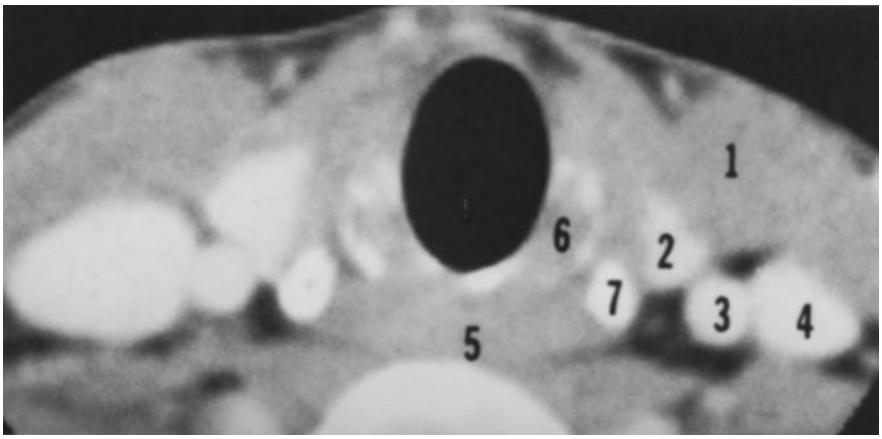


FIGURE 59–23. Normal axial computed tomographic scan at the level of the upper subglottic region. The cricoid cartilage is partially calcified (6). Notice the sternocleidomastoid muscles (1), thyroid gland (2), common carotid arteries (3), internal jugular veins (4), postcricoid region (5), and inferior cornua of the thyroid cartilage (7).

SALIVARY GLANDS

Parotid Gland

The parotid gland, which lies immediately inferior and anterior to the external ear, is divided into a superficial lobe and a deep lobe that curves around the posterior margin of the ramus of the mandible and extends deeply inward toward the lateral parapharyngeal wall (see Figure 59–1, E through H).[45] The main portion of the gland is superficial and is connected to the deep lobe by an isthmus. On a CT scan, the parotid gland is seen as a relatively low-density image with its superficial lobe somewhat flattened and quadrilateral in form in children and young adults. Its density is lower but may be the same as the submandibular gland. It is placed between the ramus of the mandible, mastoid process, and sternocleidomastoid muscle. Not uncommonly, a prominent indentation by the sternocleidomastoid muscle is seen on CT scans or MRIs over the posterior aspect of the gland (see Figure 59–1, E through H). The superficial lobe is broad and can reach nearly to the zygomatic arch. Stensen's duct crosses the masseter muscle, pierces the buccinator muscle and mucous membrane of the mouth, and opens in the cheek opposite the second upper molar tooth. The main duct and intraglandular ductal system can be visualized with standard or CT contrast sialography as well as magnetic resonance sialography. The parotid glands demonstrate an intermediate signal intensity on T_1-weighted and proton-weighted MRIs that enables the clinician to define the gland–muscle interface (see Figure 59–1, E through H).

Submandibular Gland

The submandibular gland is situated in the submandibular triangle, reaching anteriorly to the anterior belly of the digastric muscle and posteriorly to the stylomandibular ligament, which intervenes between it and the parotid gland.[45,46] It extends superiorly deep to the inferior border of the mandible, in the submandibular fossa to the medial side of the body of the mandible, opposite the second and third molar teeth (see Figures 59–1, A, and 59–38). The mylohyoid muscle indents the anterior surface of the gland. The lower part of the gland is covered by the skin, superficial cervical fascia, platysma, and deep cervical fascial. The submandibular duct (Wharton's duct) runs anteriorly between the mylohyoid, hyoglossus, and genioglossus muscles and the sublingual gland to open in the caruncula sublingualis. The entire gland and duct can be visualized with CT sialography (Figure 59–39). Both submandibular and sublingual glands are visualized distinctly on MRI or CT, particularly after intravenous infusion of iodine contrast (Figure 59–40).

Sublingual Gland

The sublingual gland is situated beneath the mucous membrane of the floor of the mouth, at the site of the frenulum of the tongue, in contact with the sublingual depression on the inner surface of the mandible close to the symphysis.[24] Its excretory ducts are multiple. The larger one (Bartholin's duct) opens into the submandibular duct.[45] Sometimes the sublingual gland can be visualized during sialography of the submandibular gland (see Figure 59–39). Because of marked enhancement, both the submandibular and sublingual glands are visualized with CT incremental dynamic study (see Figure 59–40). The parotid, submandibular, and sublingual glands are seen on MRIs with exquisite anatomic details (see Figure 59–1).

Parotid Gland Pathology

Acute infection of the parotid gland manifests itself as diffuse swelling of the gland. An abscess is seen on a CT scan as a discrete low-density (lucent) image (Figure 59–41). The inflammatory processes of adjacent structures, such as an abscess of the masseter and sternocleidomastoid muscles, can be diagnosed and differentiated from a parotid abscess. The diagnosis of chronic parotitis, which results from ductal obstruction because of a stone and recurrent parotitis and chronic punctate sialadenitis (benign lymphoepithelial disease, Mikulicz's syndrome, Sjögren's syndrome) of the parotid or submandibular gland, can be accomplished more readily with conventional sialography.[21,24] Computed tomography of the parotid gland performed with or without contrast infusion (Figure 59–42) and CT contrast sialography (Figure 59–43) can locate parotid masses. Magnetic resonance imaging, similar to CT, is extremely useful and has replaced CT in the evaluation of salivary gland masses. The differentiation of extrinsic mass versus intrinsic parotid tumor is a simple matter if MRI or CT is used,[24,45,46]

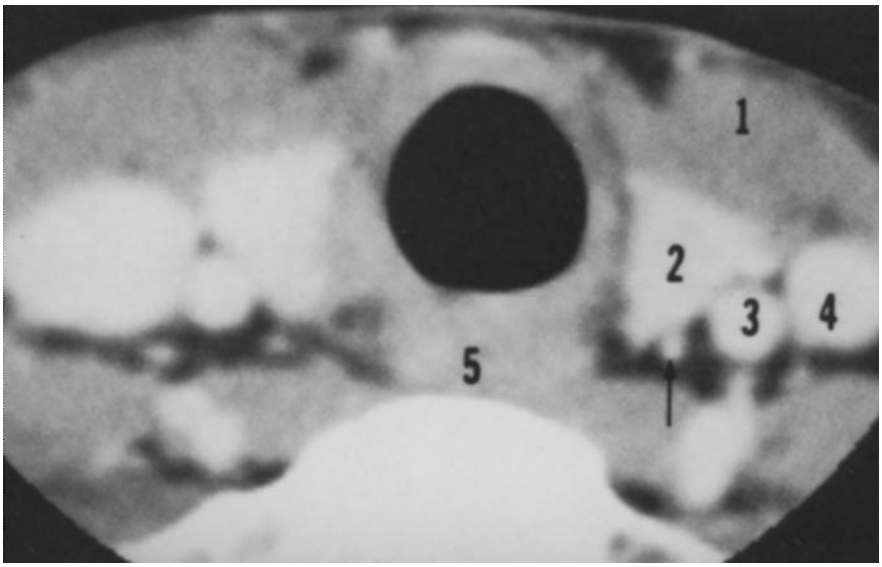


FIGURE 59–24. Normal axial computed tomographic scan at the lower subglottic region showing the sternocleidomastoid muscles (1), thyroid gland (2), common carotid arteries (3), internal jugular veins (4), postcricoid pharyngeal musculature (5), and inferior thyroid artery (*arrow*).

whereas this problem can be extremely difficult to resolve by conventional sialography.[24]

Tumors that are 1 cm in diameter or less are exceedingly difficult to demonstrate by conventional sialography, especially if they are peripherally located.[24] On CT scans, small mass lesions can be seen because they have a different CT appearance in noncontrast and contrast infusion CT as compared with normal parts of the gland.[24] On CT sialography, they appear as filling defects. Magnetic reso-

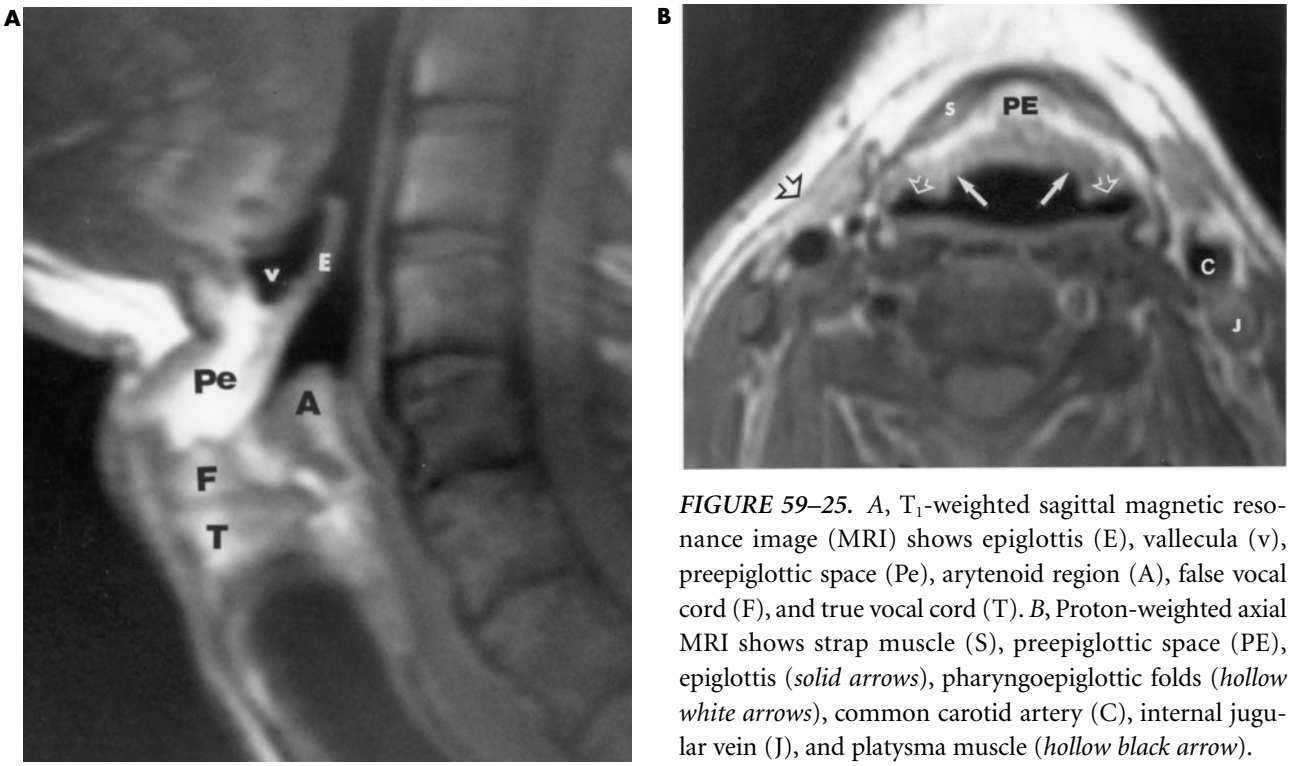


FIGURE 59–25. *A*, T_1-weighted sagittal magnetic resonance image (MRI) shows epiglottis (E), vallecula (v), preepiglottic space (Pe), arytenoid region (A), false vocal cord (F), and true vocal cord (T). *B*, Proton-weighted axial MRI shows strap muscle (S), preepiglottic space (PE), epiglottis (*solid arrows*), pharyngoepiglottic folds (*hollow white arrows*), common carotid artery (C), internal jugular vein (J), and platysma muscle (*hollow black arrow*).

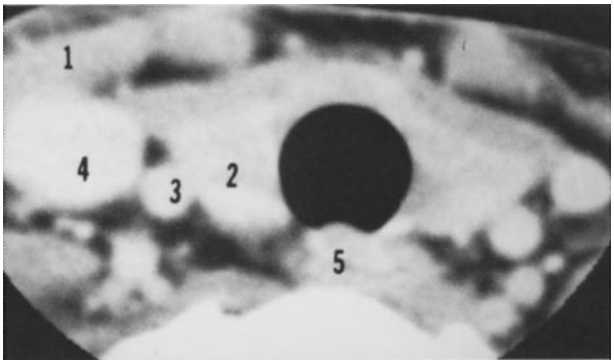


FIGURE 59–26. Normal axial computed tomographic scan at the level of the first tracheal ring showing the sternocleidomastoid muscle (1), thyroid gland (2), common carotid artery (3), internal jugular vein (4), and esophagus (5).

nance imaging, because of its superior contrast resolution, is more sensitive than CT scan and sialography for the evaluation of salivary gland tumors. Conventional sialography is the study of choice for the evaluation of chronic sialopathy, including Sjögren's syndrome.

Tumors of the Parotid Gland The primary neoplastic disorders of the salivary glands may be divided as follows: (1) neoplasms of supporting tissue origin such as hemangioma, lymphangioma, lipoma, schwannoma, fibroma, malignant lymphoma, and various sarcomas and (2) neoplasms of epithelial tissue origin that can be classified as follows:

1. Benign lesions, which include mixed tumors (pleomorphic adenoma), papillary cystadenoma lymphomatosum (Warthin's tumor), and acidophilic cell adenoma (oncocytoma, oncocytosis), and the monomorphic tumors, which include basal cell adenoma, glycogen-rich adenoma, clear cell adenoma, and myoepithelioma.[45,46] Other benign tumors of the salivary gland include papillary ductal adenoma (papilloma), benign lymphoepithelial lesion, sebaceous gland adenoma, and lymphadenoma.[45]
2. Malignant lesions of the salivary glands, which include malignant mixed tumor, malignant

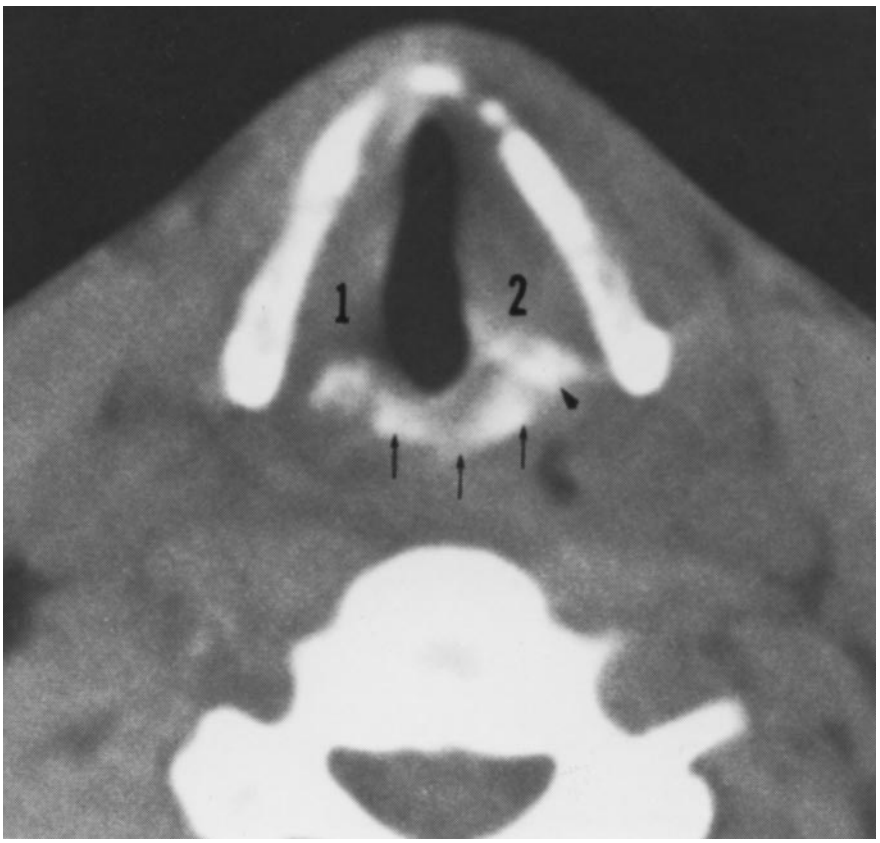


FIGURE 59–27. Computed tomographic scan, showing enlargement (hematoma) of the left true vocal cord (2). The left arytenoid cartilage (*arrowhead*) is rotated medially. The signet portion of the cricoid cartilage (*arrows*) and right true vocal cord (1) appear normal.

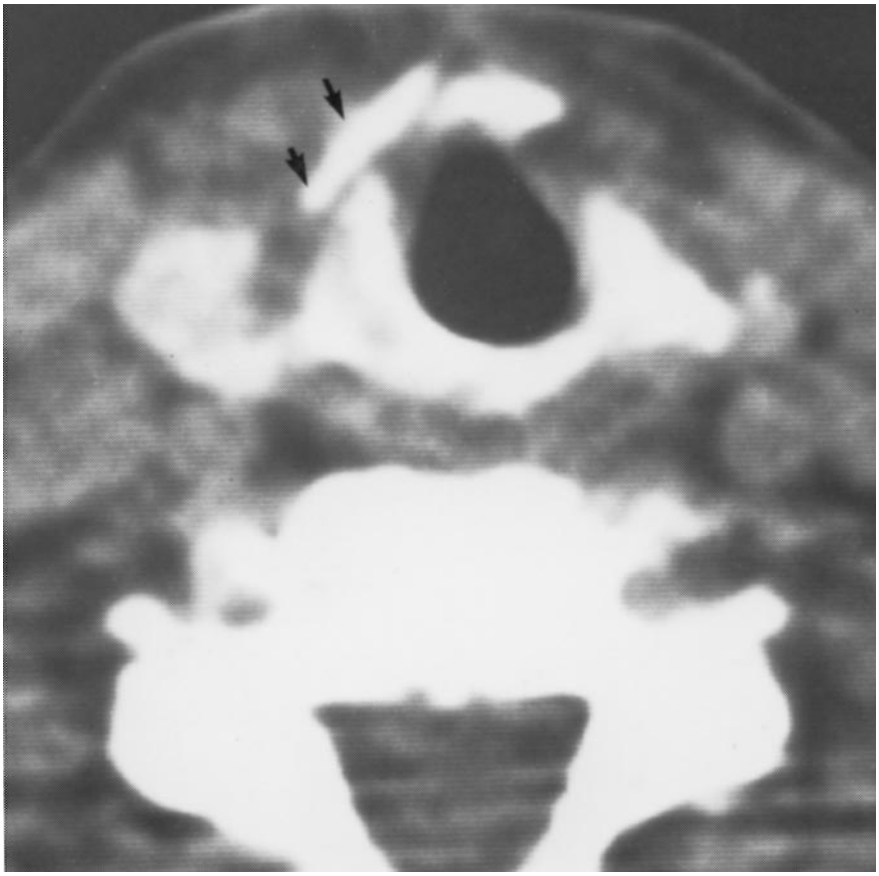


FIGURE 59–28. Computed tomographic scan at the level of the inferior margins of the true vocal cords showing a displaced fragment of the fractured thyroid cartilage (*arrows*).

ex-pleomorphic adenoma (carcinoma arising in a benign mixed tumor), primary squamous cell carcinoma, adenoid cystic adenocarcinoma, mucoepidermoid carcinoma (low grade, high grade), acinic cell (acinous) adenocarcinoma (low grade, high grade), oncocytic carcinoma (malignant oncocytoma), adenocarcinoma (mucus producing), clear cell adenocarcinoma (nonmucinous and glycogen containing or nonglycogen containing), and other adenocarcinomas; and miscellaneous, which include sebaceous adenocarcinoma, Stensen's duct tumors, melanoma, carcinoma lymphoepithelial lesion, metastasis, carcinosarcoma, undifferentiated carcinoma, and unclassified lesions.[45,46]

Cystic lymphoid hyperplasia in acquired immune deficiency syndrome (AIDS) involves the parotid gland lymph nodes. The changes are identical to those of persistent generalized lymphadenopathy in association with gross epithelial cysts in the lymph nodes. The cystic spaces are lined by the tonsil-like, invaginating squamous epithelium.[45]

Computed Tomography and Magnetic Resonance Image Appearance of Salivary Gland Tumors In general, differentiation of benign and malignant tumors by CT or MRI is not the primary goal of CT and MRI. Rather, CT and MRI become important in showing total extent of disease, local lymphadenopathy, and distant metastasis.[45] The CT appearance of cysts is similar to that of cysts in other organs of the body.[45] Generally, cysts have thin walls, are well circumscribed, and contain water-density material, or if they contain proteinaceous fluid, the cyst may appear rather dense. Hemangiomas may have a characteristic CT appearance, demonstrating intense enhancement on dynamic

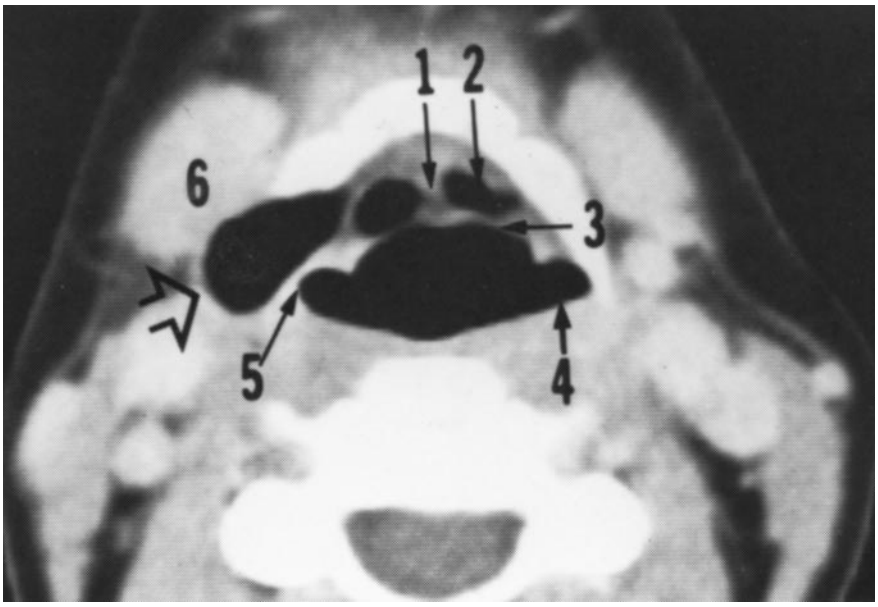


FIGURE 59–29. Laryngocele. The laryngocele bulges outside the larynx laterally (*open arrow*). 1 = median glossoepiglottic fold; 2 = vallecula; 3 = tip of the epiglottis; 4 = upper portion of the piriform sinus; 5 = pharyngoepiglottic fold; 6 = submandibular gland.

CT scans. Benign mixed tumors are typically sharply circumscribed, lying in an otherwise normal gland (Figures 59–44 and 59–45).[45]

Intraparotid facial neuroma may not be differentiated from benign mixed tumor (see Figure 59–42). In the parotid, most tumors are usually of higher density than the general low-density parotid gland. Some parotid glands may be relatively dense; therefore, the mass may be isodense with gland tissue. Identification may then require CT sialography or MRI to outline the tumor mass. Submandibular masses have similar CT appearances to those of the parotid; however, the normally high density of the submandibular gland at times necessitates MRI and

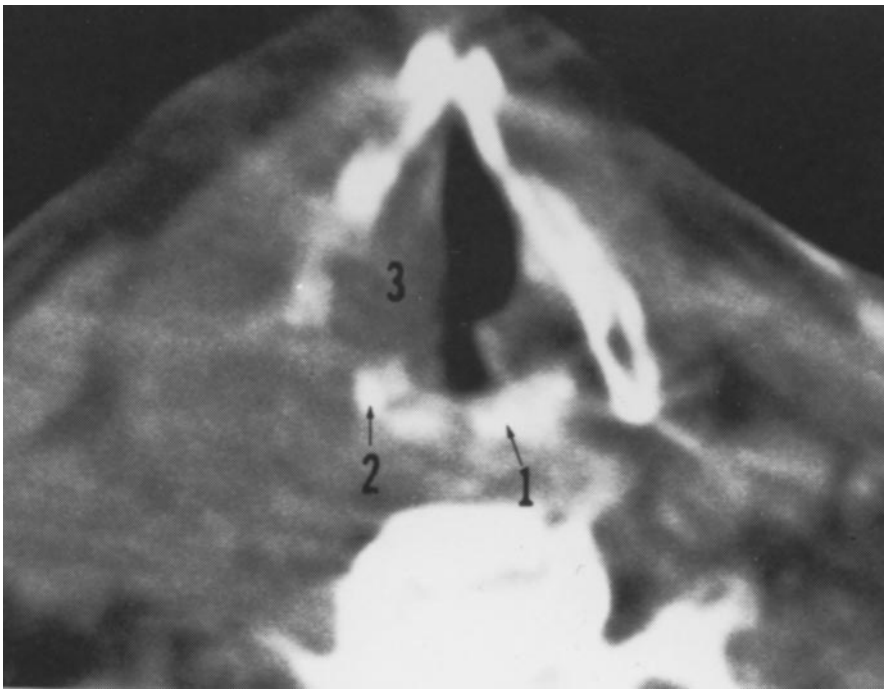


FIGURE 59–30. Computed tomographic scan showing a large tumor involving the right true vocal cord that is fixed in the paramedian position (3). Note extension of the tumor between partially destroyed thyroid cartilage and partially destroyed right arytenoid cartilage (2). Note tumor extension outside of the larynx into the neck. 1= left arytenoid cartilage.

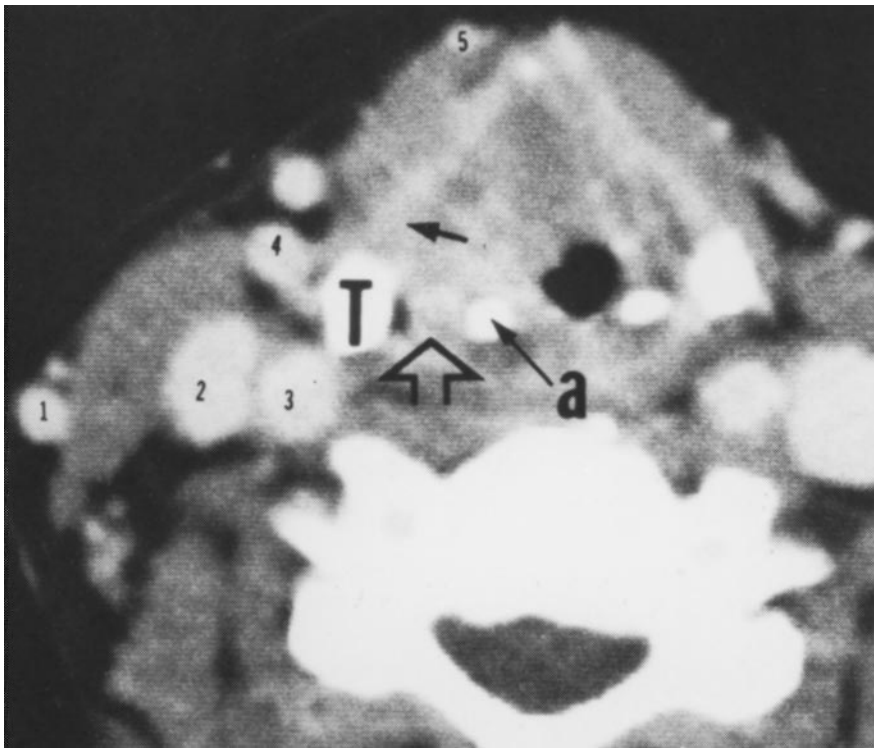


FIGURE 59–31. Postcontrast computed tomographic scan showing a large tumor involving the right false vocal cord and epiglottis and extending beyond the midline involving the left false vocal cord. The thyroid cartilage (T) is partially calcified, and the thyroid alae are seen as faint increased density with the tumor involving the right ala (*arrow*) and extending into the paraglottic space (*open arrowhead*) between the thyroid cartilage and right arytenoid cartilage (a) that is displaced medially. 1 = external jugular vein; 2 = internal jugular vein; 3 = common carotid artery; 4 = posterior facial vein; 5 = anterior jugular vein.

occasionally may necessitate CT sialography. The CT appearance of Warthin's tumor or papillary cystadenoma lymphomatosum, the second most common benign tumor in the parotid gland, is nonspecific, usually appearing as a well-circumscribed mass within or on the surface of the parotid gland.[45] Although, histologically, these tumors often have cystic components and some mucoid fluid within, the CT characteristics are not specific because the cystic spaces are very small.[45] At times, Warthin's tumors tend to have multiple lobules, may originate from relatively small intraglandular components with larger extraglandular components, with or without foci of calcifications, and have the appearance that has been compared to a branch of grapes.[45] Malignant tumors may be similarly sharply circumscribed (Figure 59–46) or may have indistinct margins. Invasive tumors are poorly marginated and extend beyond the glandular tissue into the fat and fascial planes. These tumors invade adjacent anatomic structures and may extend into the external auditory canal, middle ear, base of the skull, lateral pharyngeal space, and infratemporal fossa. Localized inflammatory lesions within the salivary glands may mimic mass lesions. Inflammatory disease usually results in a relatively diffuse, irregular, radiodense lesion in an enlarged gland.[45] It is not possible by CT alone to differentiate a benign lesion from a malignant one whenever a well-defined tumor is present in the salivary glands. Although an infiltrating mass is most likely to be a malignant lesion, an inflammatory reaction around a benign mass can have the same appearance as a malignant tumor. It should also be noted that masses arising in a salivary gland may be entirely or almost entirely surrounded by salivary tissue, or they may extend outside the salivary gland and be only partly buried in the gland.[45] In this situation, the intraglandular component abuts the salivary gland tissue directly without an intervening fat layer.

Magnetic resonance imaging and CT are both useful in the evaluation of salivary gland masses.[45,46] They can help to differentiate intrinsic from extrinsic lesion such as a cervical lymph node and a parapharyngeal mass from a parotid mass. The MRI may give some indication of the relationship of a parotid mass to the facial nerve. Although the MRI, like the CT scan, cannot give a histologic diagnosis, it can

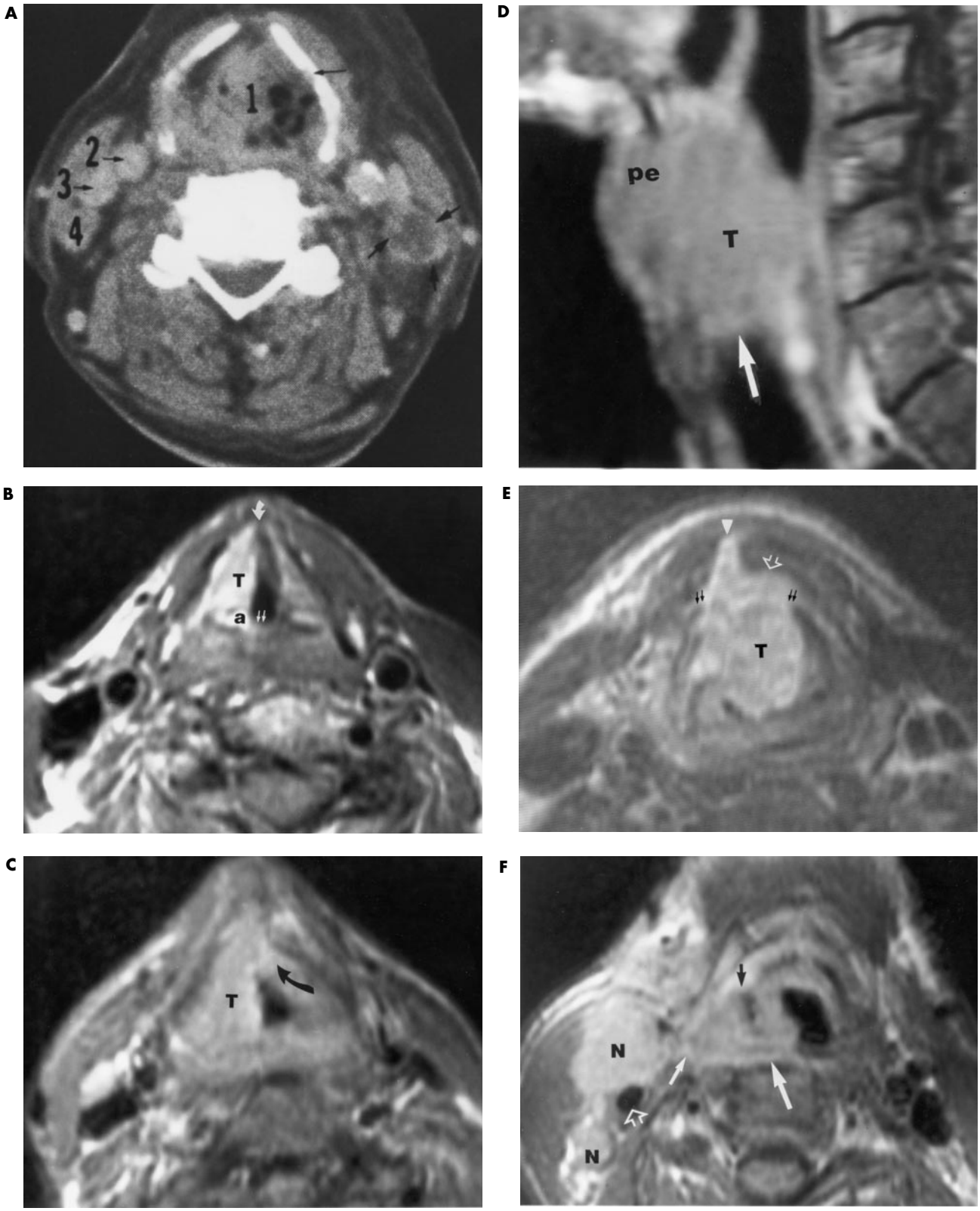


FIGURE 59–32. A, Computed tomographic scan showing a large necrotic exophytic tumor arising from the left aryepiglottic fold (1). Note tumor involvement of the left thyroid ala (*arrow*) and epiglottis and extension beyond the midline with involvement of the right aryepiglottic fold and destruction of the posterior aspect of the right thyroid ala. Note large left necrotic lymph node (*arrows*) and large right lymph node (4). 2 = common carotid artery; 3 = internal

jugular vein. *B*, Glottic carcinoma. Proton-weighted axial magnetic resonance image (MRI) shows tumor (T) involving the entire right true vocal cord. Note normal anterior (*curved arrow*) and posterior (*arrows*) commissures. The arytenoid cartilage (a) is normal. *C*, Supraglottic carcinoma. Proton-weighted axial MRI shows tumor (T) involving the right false vocal cord and extending beyond the midline to involve the left false vocal cord (*curved arrow*). *D*, Supraglottic carcinoma. T_1-weighted sagittal MRI, same patient as in *C*, shows large supraglottic tumor (T) with involvement of the preepiglottic space (pe) and extension into the immediate subglottic space (*arrow*). *E*, Supraglottic carcinoma. T_2-weighted axial MRI shows large tumor (T) with involvement of the thyroid cartilage (*small arrows*), extending into the preepiglottic space (*hollow arrow*) and strap muscles (*arrowhead*). *F*, Carcinoma of the right piriform sinus. Proton-weighted axial MRI shows a large mass (*arrows*) involving the right side of the hypopharynx. Note multiple metastatic lymph nodes (N). The common carotid artery (*hollow arrow*) is intact.

give an indication of the general morphology of the tumor. Rounded, sharply marginated, well-capsulated tumors tend to be benign (see Figure 59–45), whereas irregular and diffuse tumors tend to be malignant. Lipomas, lymphangiomas, and hemangiomas may be specifically differentiated from other salivary tumors. Most salivary gland tumors are seen on MRI as intermediate or low signal intensity on T_1- and proton density–weighted images and as high signal intensity on T_2-weighted images (see Figures 59–42 and 59–45). The tumors, when they are small, have a homogeneous appearance. Mixed signal

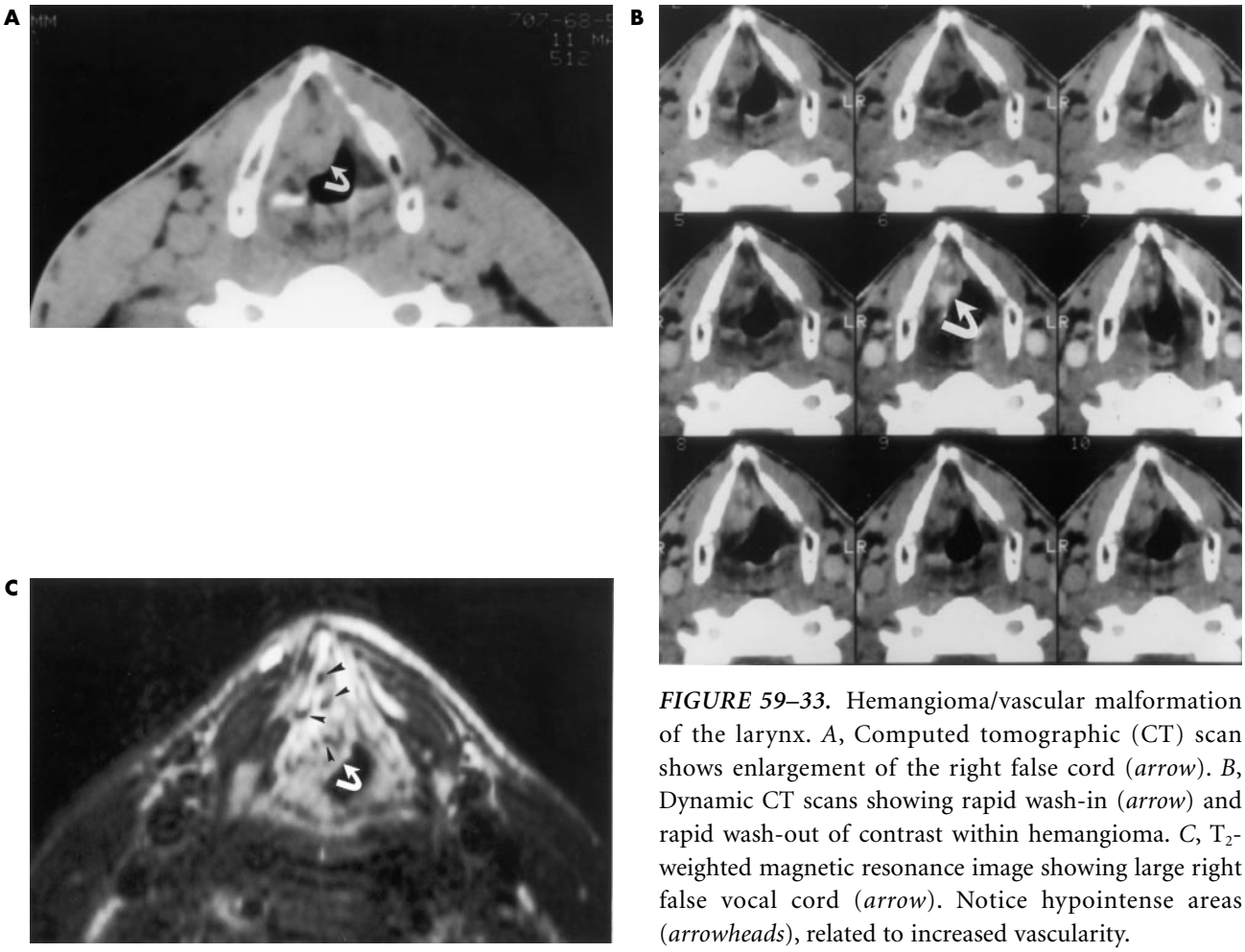


FIGURE 59–33. Hemangioma/vascular malformation of the larynx. *A*, Computed tomographic (CT) scan shows enlargement of the right false cord (*arrow*). *B*, Dynamic CT scans showing rapid wash-in (*arrow*) and rapid wash-out of contrast within hemangioma. *C*, T_2-weighted magnetic resonance image showing large right false vocal cord (*arrow*). Notice hypointense areas (*arrowheads*), related to increased vascularity.

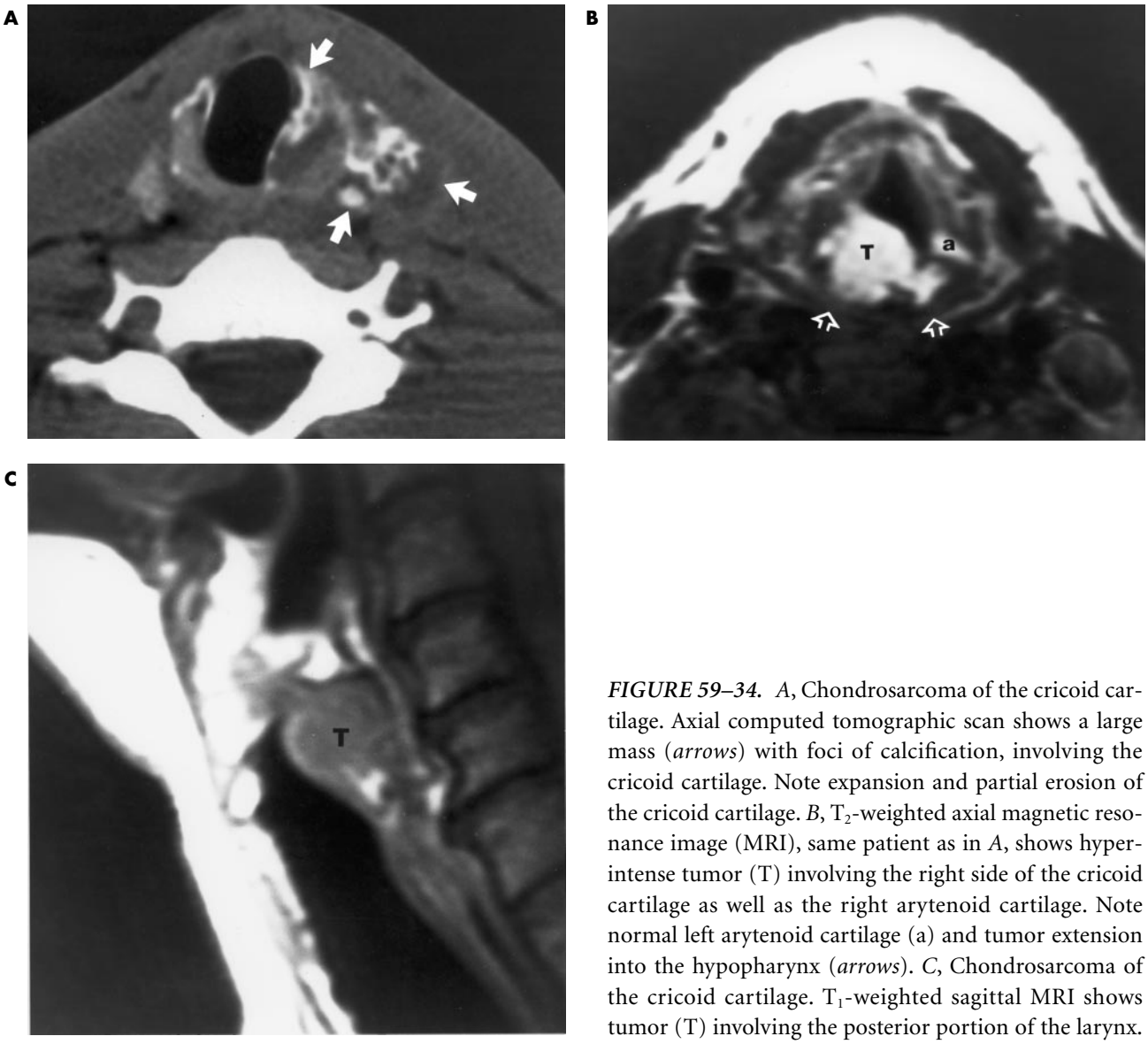


FIGURE 59–34. *A*, Chondrosarcoma of the cricoid cartilage. Axial computed tomographic scan shows a large mass (*arrows*) with foci of calcification, involving the cricoid cartilage. Note expansion and partial erosion of the cricoid cartilage. *B*, T_2-weighted axial magnetic resonance image (MRI), same patient as in *A*, shows hyperintense tumor (T) involving the right side of the cricoid cartilage as well as the right arytenoid cartilage. Note normal left arytenoid cartilage (a) and tumor extension into the hypopharynx (*arrows*). *C*, Chondrosarcoma of the cricoid cartilage. T_1-weighted sagittal MRI shows tumor (T) involving the posterior portion of the larynx.

intensities may be present particularly when they are larger than 3 cm. The serous- and mucus-filled spaces are common in benign mixed tumors.[45] In Warthin's tumor, they are more likely to be related to areas of frank cystic degeneration. In general, mixed signal intensities may be present in inflammatory lesions, benign tumors, and malignant tumors on both T_1- and T_2-weighted images.[45] Malignant salivary gland tumors may show intermediate to low signal intensity on T_2-weighted images. The more cellular and more aggressive a tumor is, the less likely it is that it will be hyperintense on T_2-weighted images. Calcification may be present in Warthin's tumor and acinic cell tumor. Calcification is rare in mixed tumors. Irregular calcifications in a mixed tumor, associated with pain or rapid growth, should raise the possibility of malignant sarcomatous (osteochondrogenic) degeneration.

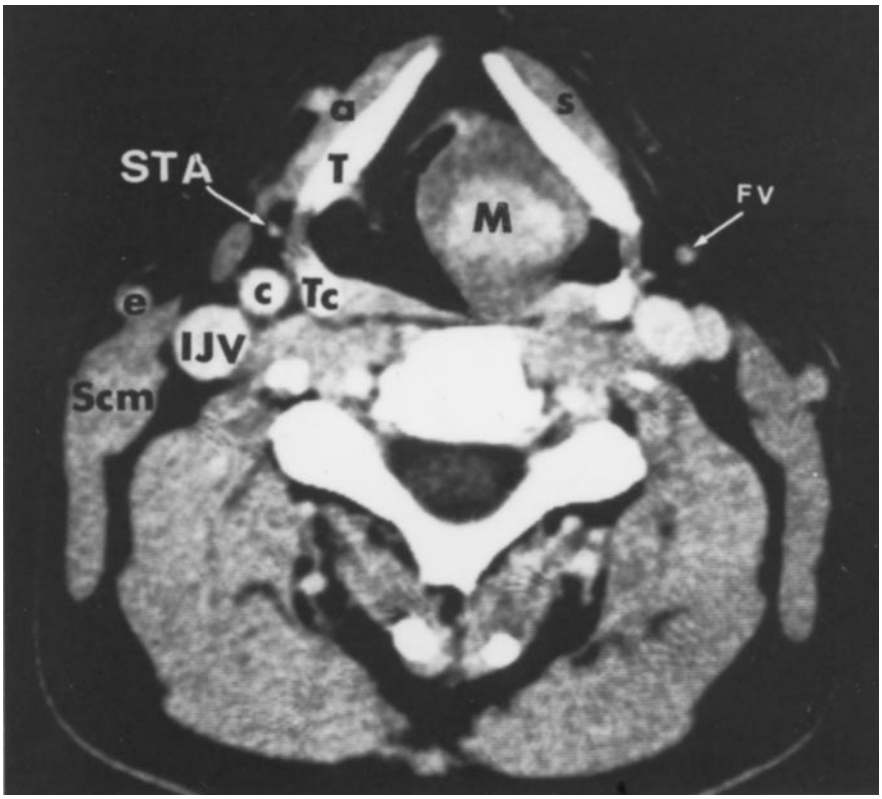


FIGURE 59–35. Leiomyoma of the larynx (M) for comparison with squamous cell carcinoma in Figure 59–32, A. a = anterior jugular vein; c = common carotid artery; e = external jugular vein; FV = facial vein; IJV = internal jugular vein; s = strap muscles; Scm = sternocleidomastoid muscle; STA = superior thyroid artery.

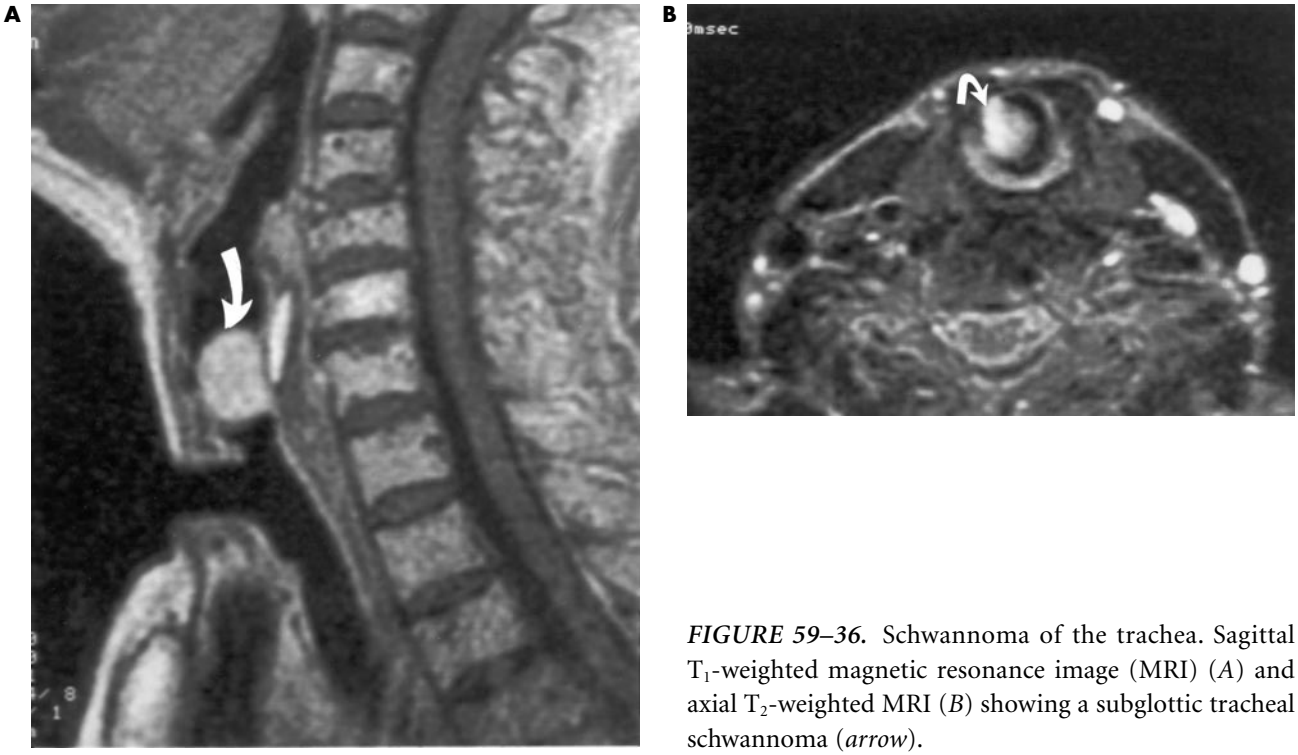


FIGURE 59–36. Schwannoma of the trachea. Sagittal T_1-weighted magnetic resonance image (MRI) (*A*) and axial T_2-weighted MRI (*B*) showing a subglottic tracheal schwannoma (*arrow*).

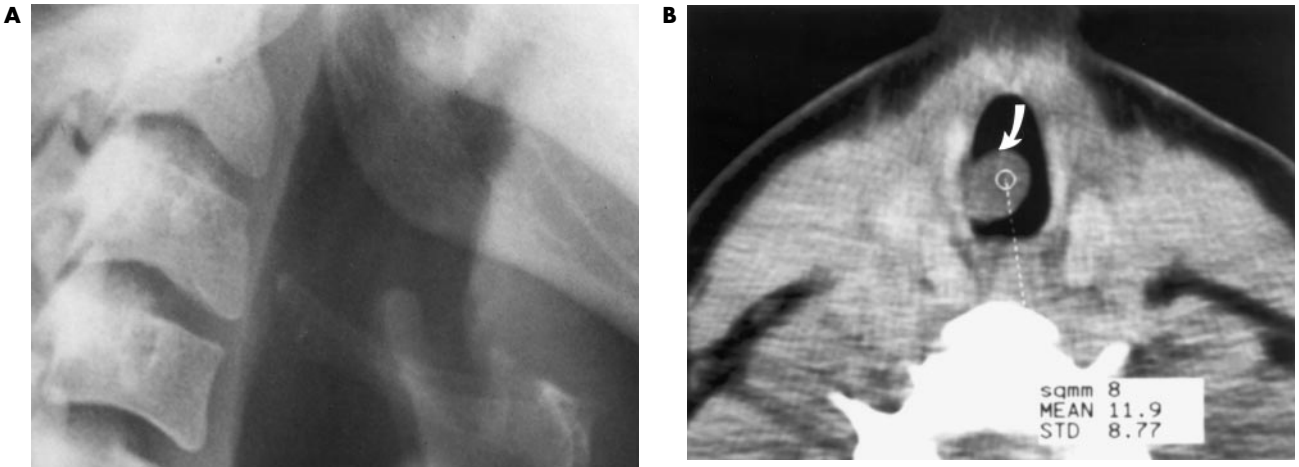


FIGURE 59–37. Inflammatory myofibroblastic tumor of the subglottic-tracheal airway. *A*, Soft tissue lateral x-ray film showing a well-defined intratracheal mass (*arrows*). *B*, Axial computed tomographic scan showing the mass (*arrow*).

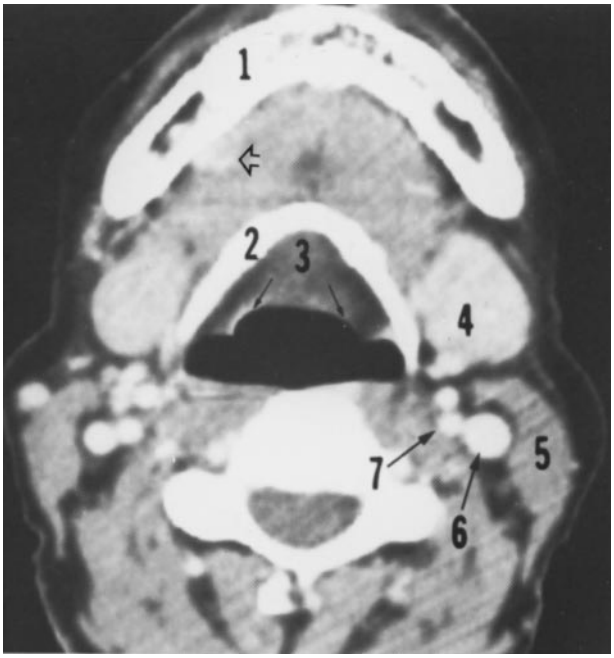


FIGURE 59–38. Axial computed tomographic scan of the oropharynx, showing the mandible (1), hyoid bone (2), base of the tongue (3), epiglottis (*arrows*), submandibular gland (4), sternocleidomastoid muscle (5), internal jugular vein (6), internal carotid artery (7), and, anterior to that, external carotid artery. The submandibular (4) and sublingual (*open arrow*) glands are intensely enhanced during rapid infusion of iodine contrast material.

Benign Mixed Tumors The most common benign tumor of salivary gland origin is the benign mixed tumor. A solitary mass in the parotid glands is a benign mixed tumor in 60 to 65% of the cases.[2,45] In the parotid, they are found most often in the lateral portion but may occur in any portion of it. [45] The CT and MRI appearance of these tumors represents a well-circumscribed mass, lying in an otherwise normal gland (see Figure 59–45). The CT and MRI appearance of an intraparotid facial neuroma (see Figure 59–42) may be identical to a benign mixed tumor. A characteris-

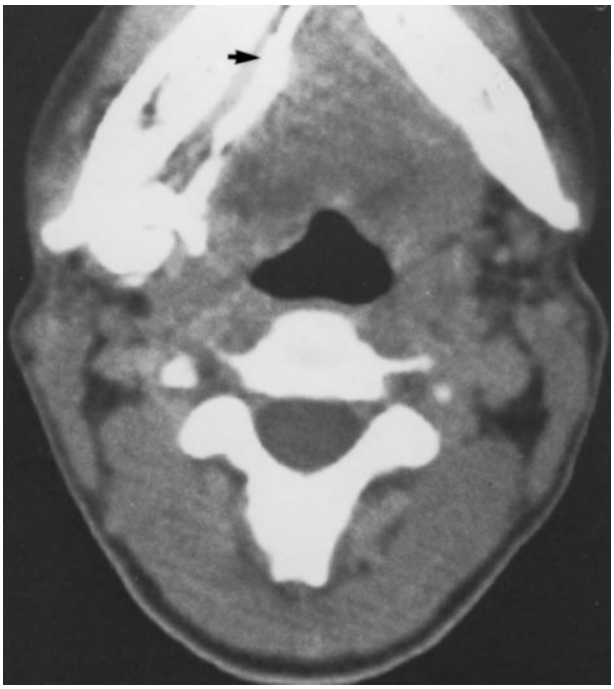


FIGURE 59–39. Computed tomographic sialogram showing the right submandibular gland and Wharton's duct (*arrow*). Note atrophic submandibular gland on the left side.

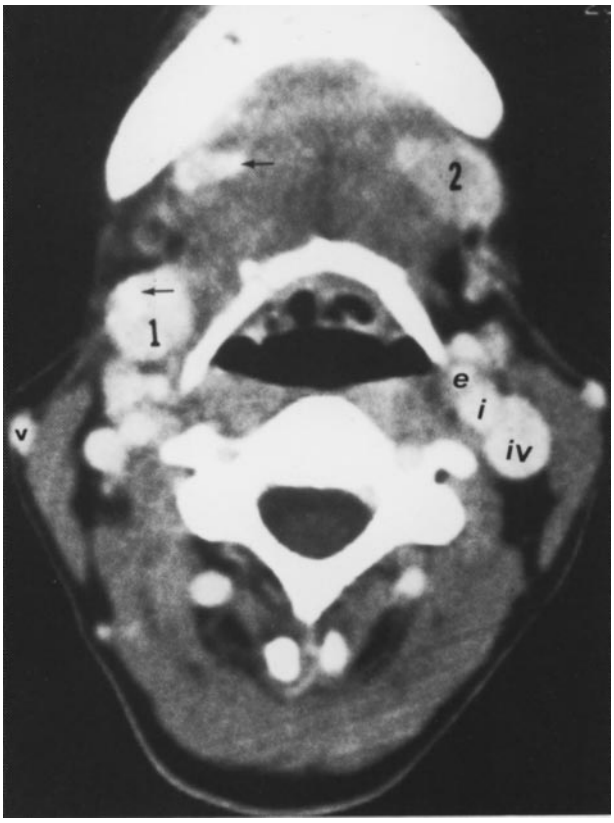


FIGURE 59–40. Axial computed tomographic (CT) scan of the oropharynx showing marked enhancement of the submandibular (1) and sublingual (2) glands. *Arrows* point to the residual pantopaque in the glands from previous CT pantopaque sialography. Note atrophic left submandibular gland (compare with Figure 59–38) and intense enhancement of the external (e) and internal (i) carotid arteries and external (v) and internal (iv) jugular veins.

tic but not pathognomonic feature of benign mixed tumors is that they become more heterogeneous in appearance on postcontrast T_1-weighted MRIs.

Papillary Cystadenoma Lymphomatosum. Papillary cystadenoma lymphomatosum (Warthin's tumor) occurs almost exclusively in the parotid gland and affects males more frequently than females.[2,45,46] It accounts for approximately 6 to 8% of the parotid tumors.[45] The mass usually occurs in the tail of the gland and is characterized by a painless, slow-growing asymptomatic mass. The tumor occasionally occurs in both parotid glands. Histologically, its characteristics include a double-layered epithelium having a tubulopapillary cystic pattern within lymphoid tissue or a lymph node.[45] The CT and MRI appearance of a Warthin's tumor is nonspecific, usually appearing as a well-marginated mass. At times, frank cystic degeneration may be present. There may be foci of fine or coarse calcifications present. These tumors characteristically concentrate technetium 99.[45]

Acidophilic Cell Adenoma. Acidophilic cell adenoma (oncocytoma, oncocytosis) is a benign lesion believed to arise from ductal epithelium.[45] The characteristic cell of the oncocytoma is the oncocyte. This cell is characterized by an excessive number of mitochondria that is best determined with electron

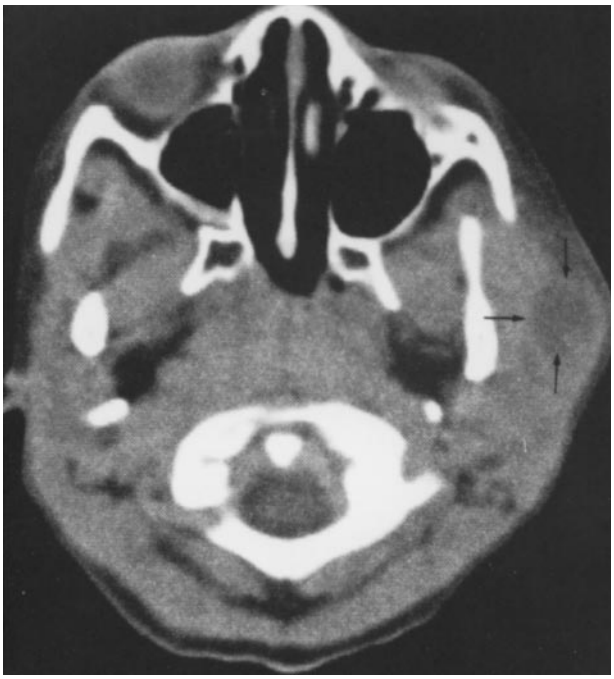


FIGURE 59–41. Parotid gland abscess. Computed tomographic scan of the parotid glands showing a round low-density image (*arrows*) caused by an abscess.

microscopy.[45] Oncocytoma such as Warthin's tumor has the ability to concentrate technetium 99; therefore, a "hot" area on radioisotope technetium 99 will indicate the possible presence of either a Warthin's tumor or an oncocytoma.

Gland Cell Carcinoma (Adenocarcinoma) *Adenoid Cystic Carcinoma.* Adenoid cystic carcinoma is uncommon in the parotid gland and common in the submandibular gland and the minor salivary glands of the oral cavity.[45] This tumor is also variously named adenocystic adenocarcinoma, cylindroma, pseudoadenomatous basal cell carcinoma, and adenocarcinoma.[2] These tumors occur more frequently in women in early middle age.[45] Pain is common and may be a prominent symptom. The mass grows slowly, and encapsulation of the lesion is not common. This tumor is unique in its strong tendency to invade nerves and perineural lymphatics.[45,46]

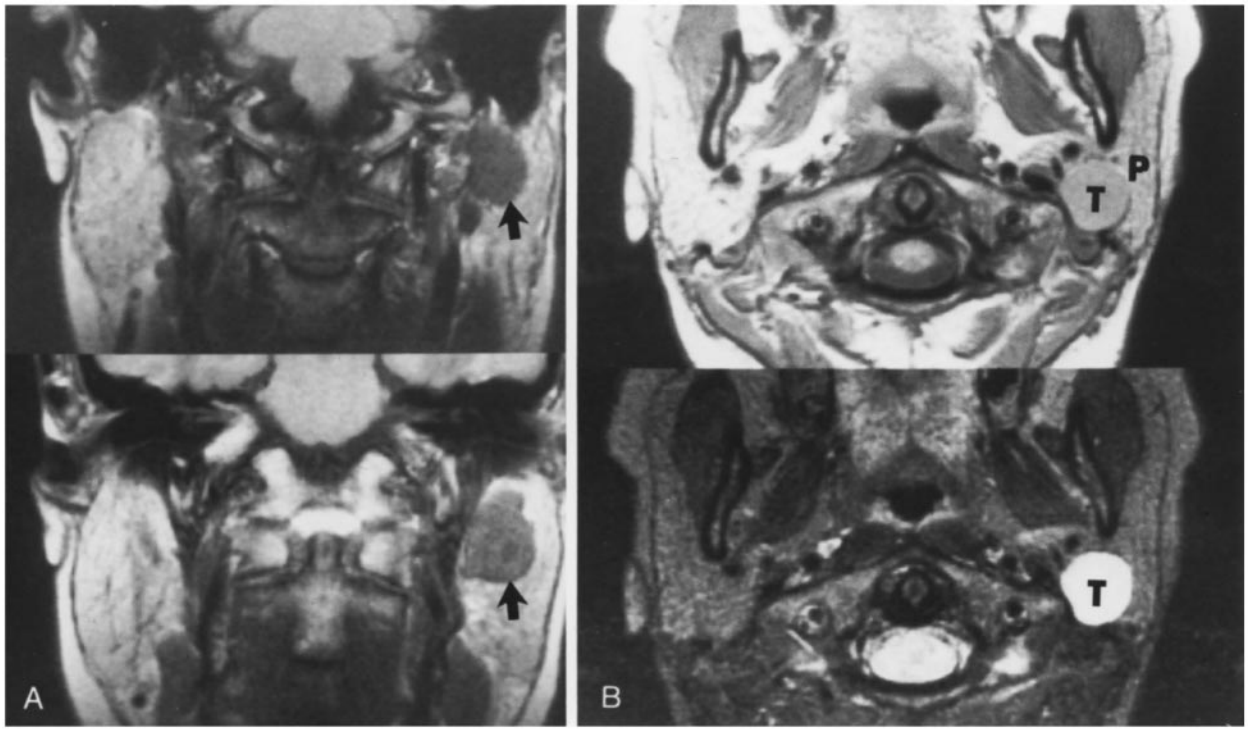


FIGURE 59–42. Facial neuroma arising within the parotid gland. *A*, T_1-weighted magnetic resonance images (MRIs) showing a hypointense mass (*arrows*). *B*, Proton-weighted (*top*) and T_2-weighted (*bottom*) axial MRIs showing the tumor (T) within the parotid gland (P). Notice hyperintensity of the tumor in the T_2-weighted image.

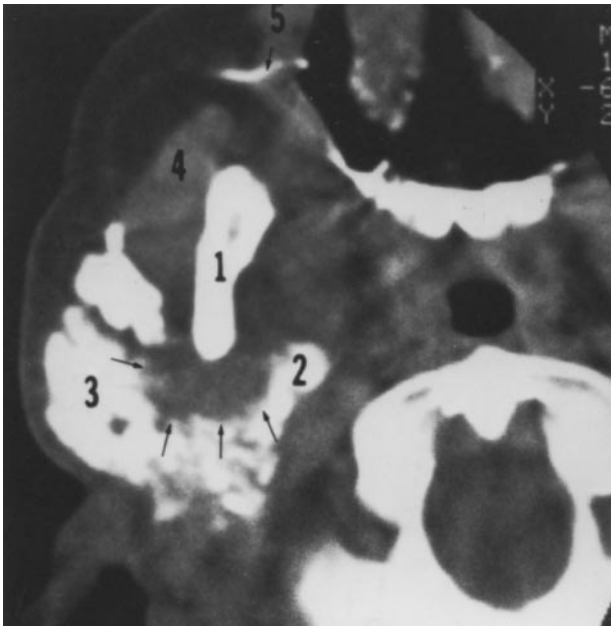


FIGURE 59–43. Axial computed tomographic sialogram of a parotid tumor. A large tumor (*arrows*) is seen involving the superficial and deep lobes of the right parotid gland. The lesion is sharply delineated from the surrounding gland tissue that is filled with contrast. 1 = mandible; 2 = deep lobe of parotid gland; 3 = superficial lobe of the parotid gland; 4 = masseter muscle; 5 = contrast in Stensen's duct.

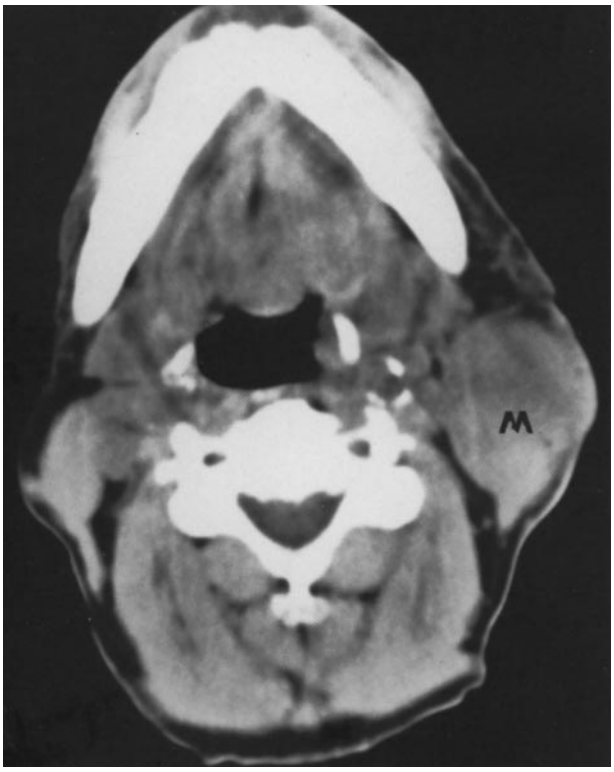


FIGURE 59–44. Parotid tumor. Axial computed tomographic scan shows a large mass (M) involving the left parotid gland.

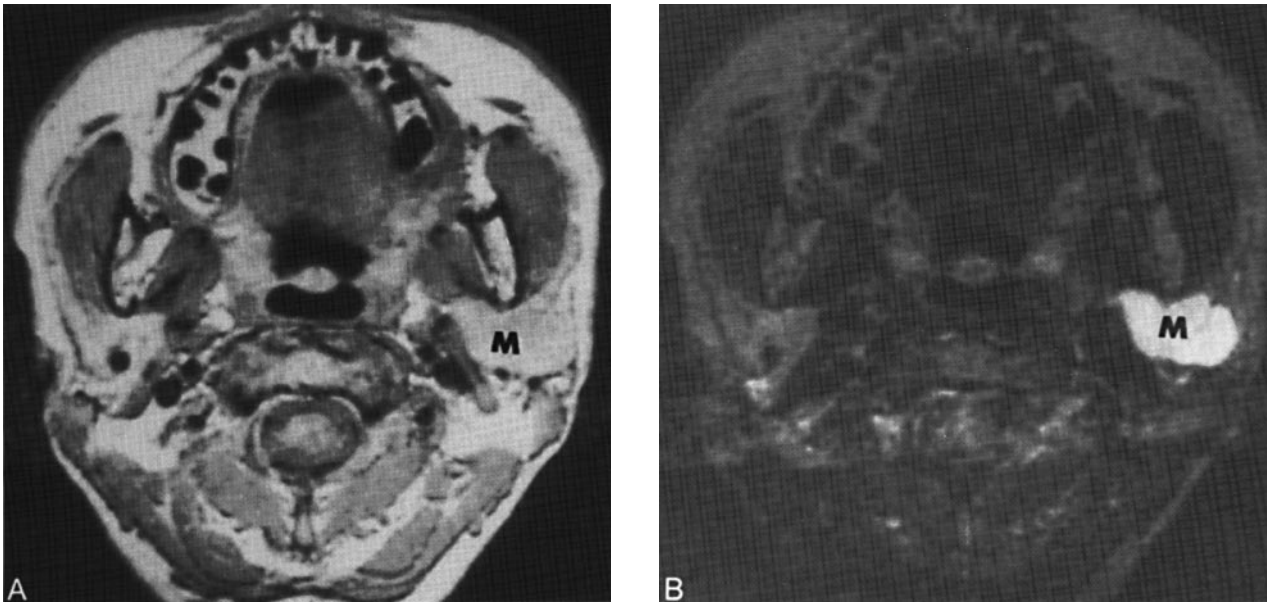


FIGURE 59–45. Pleomorphic adenoma. Proton-weighted (*A*) and T_2-weighted (*B*) axial magnetic resonance images show a well-defined mass (M) involving the left parotid gland.

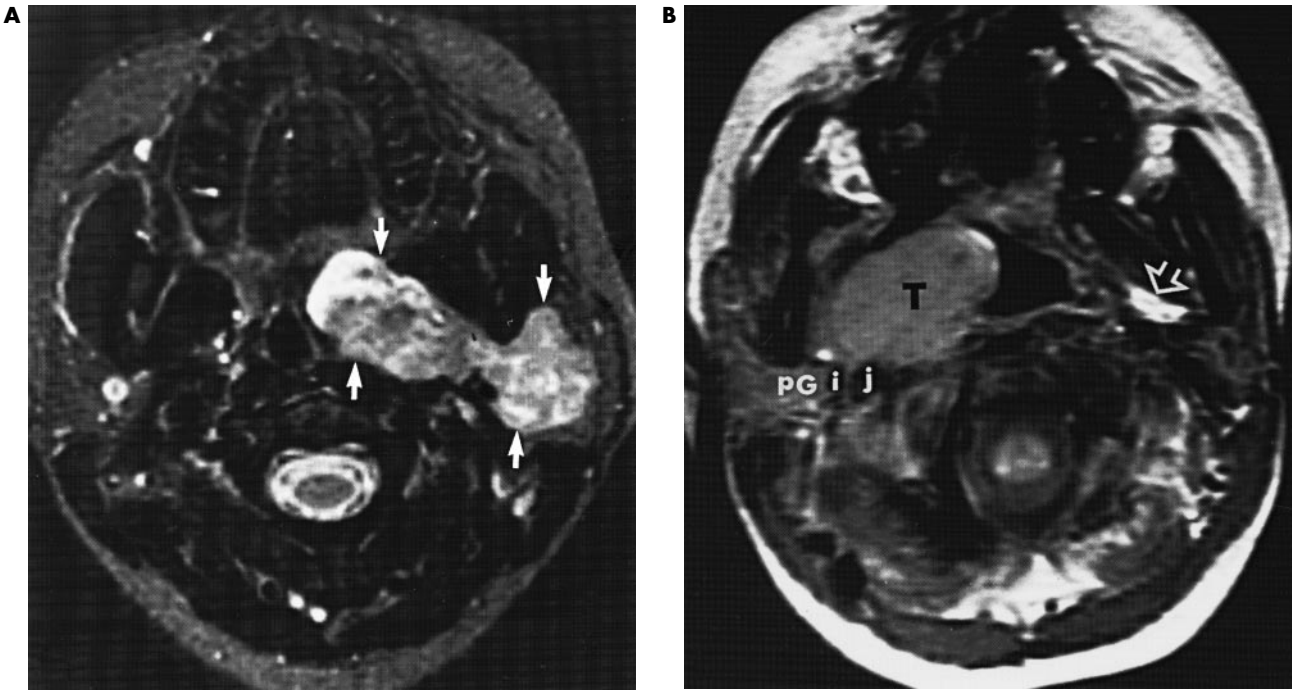


FIGURE 59–46. *A*, Malignant tumor of the parotid gland. T_2-weighted axial magnetic resonance image (MRI) shows a large mass involving the deep lobe of the left parotid gland, extending into the parapharyngeal space (*arrows*). This was reported as being compatible with liposarcoma. *B*, Parapharyngeal adenocarcinoma. Proton-weighted axial MRI shows a large tumor (T) involving the right parapharyngeal space. Notice normal left parapharyngeal space (*arrow*). i = internal carotid artery; j = internal jugular vein; PG = parotid gland.

REFERENCES

1. Mafee MF. Nasopharynx, parapharyngeal space and base of skull. In: Valvassori GE, Mafee MF, Carter BL, editors. Imaging of the head and neck. New York: Georg Thieme Verlag; 1995. p. 332–63.
2. Mafee MF, Rasouli F, Spigos DG, et al. Magnetic resonance imaging in the diagnosis of nonsquamous tumors of the head and neck. Otolaryngol Clin North Am 1986;19:523–36.
3. Mafee MF, Compos M, Raju S, et al. Head and neck: high field magnetic resonance imaging versus computed tomography. Otolaryngol Clin North Am 1988;21:513–46.
4. Pisaneschi MJ, Mafee MF, Samii M. Applications of MR angiography in head and neck pathology. Otolaryngol Clin North Am 1995;28:543–61.
5. Mafee MF. Dynamic CT and its application to otolaryngology-head and neck surgery. J Otolaryngol 1982;11:307–18.
6. Lev S, Lev MH. Imaging of cystic lesions of the neck. Radiol Clin North Am 2000;38:1013–27.
7. Shugar MA, Mafee MF. Diagnosis of carotid body tumors by dynamic computerized tomography. Head Neck Surg 1982;4:518–21.
8. Mafee MF, Valvassori GE, Shugar MA, et al. High resolution and dynamic sequential computed tomography. Use in the evaluation of glomus complex tumors. Arch Otolaryngol 1983;109:691–6.
9. Mafee MF, Raofi B, Kumar A, Muscato C. Glomus faciale, glomus jugulare, glomus tympanicum, glomus vagale, carotid body tumors, and simulating lesions. Role of MR imaging. Radiol Clin North Am 2000;38:1059–76.
10. Weber AI, Montandon C, Robson CD. Neurogenic tumors of the neck. Radiol Clin North Am 2000; 38:1077–90.
11. Eskey CJ, Robson CD, Weber AI. Imaging of benign and malignant soft tissue tumors of the neck. Radiol Clin North Am 2000;38:1091–104.

12. Curtin HD. Separation of the masticator space from the parapharyngeal space. Radiology 1987;163: 195–204.
13. Hanafee WN. Sialography—part V. In: Valvassori GE, Buckingham RA, Carter BL, editors. Radiology of the ear, nose and throat. Stuttgart: Georg Thieme; 1982.
14. Hanafee WN. Radiology of the pharynx and larynx—part IV. In Valvassori GE, Buckingham RA, Carter BL, et al, editors. Radiology of the ear, nose and throat. Stuttgart: Georg Thieme; 1982.
15. Mancuso AA, Hanafee WN, Juillard GJ, et al. The role of computed tomography in the management of cancer of larynx. Radiology 1977;124:243–4.
16. Mafee MF. CT of the normal larynx. Radiol Clin North Am 1984;22:251–64.
17. Ward PH, Hanafee W, Mancuso A, et al. Evaluation of the computerized tomography, cinelaryngoscopy and laryngography in determining the extent of laryngeal disease. Ann Otol Rhinol Laryngol 1979; 88:454–6.
18. Pullen FW. The use of computerized tomography in otolaryngology. Trans Am Acad Ophthalmol Otolaryngol 1977;84:622–7.
19. Mancuso AA, Hanafee WN. Computerized tomography of the injured larynx. Radiology 1979;133:139–44.
20. Mafee MF, Schild JA, Valvassori GE, Capek V. Computed tomography of the larynx: correlated anatomy and pathologic macrosection studies. Radiology 1983;147:123–8.
21. Schild JA, Mafee MF, Valvassori GE, et al. Laryngeal malignancies and computerized tomography—a correlation of tomographic and histopathologic findings. Ann Otol Rhinol Laryngol 1982;91:571–5.
22. Archer CR, Yeager VL. Evaluation of laryngeal cartilages by computed tomography. J Comput Assist Tomogr 1979;3:604–11.
23. Noyek AM, Shulman HS, Steinhardt M. Contemporary laryngeal radiology: a clinical perspective. J Otolaryngol 1982;11:178–85.
24. Valvassori GE, Mafee MF, Carter BL, editors. Imaging of the head and neck. New York: Georg Thieme Verlag; 1995.
25. Becker M, Moulin G, Kurt AM, et al. Non-squamous cell neoplasms of the larynx: radiologic pathologic correlations. Radiographics 1998;18: 1189–209.
26. Becker M, Zbaren P, DeLavelle J, et al. Neoplastic invasion of the laryngeal cartilage: reassessment of criteria for diagnosis at CT. Radiology 1997;203: 521–32.
27. Yousem DM, Oberholzer JC. Neurofibroma of the aryepiglottic fold. AJNR Am J Neuroradiol 1991;12: 1776–8.
28. Walner DL, Loewen MS, Kimura RE. Neonatal subglottic stenosis: incidence and trends. Laryngoscope 2001;111:48–51.
29. Hartl DM, Roger G, Denoyelle F, et al. Extensive lymphangiomas presenting with upper respiratory obstruction. Arch Otolaryngol Head Neck Surg 2000;126:1378–82.
30. Friedman M, Shelton VK, Mafee MF, et al. Metastatic neck disease. Arch Otolaryngol 1984;110:443.
31. Mafee MF. Discussion: criteria for diagnosing lymph node metastasis from squamous cell carcinoma of the oral cavity: a study of the relationship between computed tomography and histologic findings and outcome. J Oral Maxillofac Surg 1998;56:593–5.
32. Som PM. Lymph nodes of the neck. Radiology 1987;165:593–600.
33. Pillsbury HC, Clark M. A rationale for therapy of the N0 neck. Laryngoscope 1997;107:1294–315.
34. Tannock IF. Principles of cell proliferation: cell kinetics. In: Devita VT, editor. Principles and practice of oncology. 2nd ed. Philadelphia: Lippincott; 1988. p. 3–13.
35. Don DM, Anzai Y, Lufkin RB, et al. Evaluation of cervical lymph node metastasis in squamous cell carcinoma of the head and neck. Laryngoscope 1995; 105:669–74.
36. Crile G. Excision of cancer of the head and neck. JAMA 1906;47:1780.
37. Curtin HD, Ischwaran H, Mancuso AA, et al. Comparison of CT and MR imaging in staging of neck metastases. Radiology 1998;207:123–9.
38. Mukherji SK, Holliday RA. Pharynx. In: Som PM, Curtin HD, editors. Head and neck imaging. 3rd ed. St. Louis: Mosby Year Book; 1996. p. 437–72.
39. Laubenbacher C, Saumweber D, Wegner-Manslau C, et al. Comparison of fluorine-18-fluorodeoxyglucose PET, MRI and endoscopy for staging head and neck squamous-cell carcinomas. J Nucl Med 1995;36: 1747–57.
40. Lonneux M, Lawson G, Ide C, et al. Positron emission tomography with fluorodeoxyglucose for suspected head and neck tumor recurrence in the symptomatic patient. Laryngoscope 2000;110:1493–7.
41. Hoffman HT, Quets J, Toshiaki T, et al. Functional magnetic resonance imaging using iron oxide particles in characterizing head and neck adenopathy. Laryngoscope 2000;110:1425–30.

42. Chisin R, McCapinlac HA. The indications of FDG-PET in neck oncology. Radiol Clin North Am 2000; 38:999–1012.
43. Anzai Y, Blackwell KE, Hirschowitz SI, et al. Initial clinical experience with dextran-coated superparamagnetic iron oxide for detection of lymph node metastases in patients with head and neck cancer. Radiology 1994;192:709–15.
44. Quak JJ, Balm AJ, Brakkee JG, et al. Localization and imaging of radiolabelled monoclonal antibody against squamous-cell carcinoma of the head and neck in tumor-bearing nude mice. Int J Cancer 1989; 44:534–8.
45. Mafee MF, Venkatesan TK, Ameli N, et al. Tumors of parotid gland and parapharyngeal space: role of computed tomography and magnetic resonance imaging. Oper Techn Otolaryngol Head Neck Surg 1996;7:348–57.
46. Robinov J. Imaging of salivary gland pathology. Radiol Clin North Am 2000;38:1047–57.

CHAPTER 60

Neoplasms of the Nasopharynx

Frank G. Ondrey, MD, PhD, Simon K. Wright, MD

ANATOMY

Optimal interpretation of the clinical symptoms and natural history of patients with malignant processes in the nasopharynx requires a basic understanding of the anatomy of this structure. The nasopharynx is an open cuboidal chamber that lies beneath the base of the skull at the posterior aspect of the nasal fossa. It measures 4.0 to 5.5 cm transversely and 2.5 to 3.5 cm in the anteroposterior dimension and is roughly 4.0 cm in height. Anteriorly, it borders the posterior nares, within which lie the posterior ends of the middle and inferior turbinates. The roof is a sloping concave surface formed by the posterior body of the sphenoid, basilar component of the occipital bone, and anterior arch of the atlas. The nasopharyngeal tonsil lies in the midline in the roof and posterior wall. The superior pharyngeal constrictor and fascia complete the posterior wall. The floor of the nasopharynx is composed of the soft palate.

The lateral walls contain important structures such as the pharyngotympanic tube, situated 10 to 12 mm behind and slightly below the level of the posterior part of the inferior turbinate. The lateral walls consist of only two layers: the mucous membrane and the pharyngeal aponeurosis. The cartilaginous eustachian tube passes through this aponeurosis, opening into the fossa of Rosenmüller. Lateral to the lateral wall, the mandibular nerve exits the foramen ovale into the infratemporal fossa. Posterior to the eustachian tube is the retroparotid space, which contains the pharyngeal lymph nodes, internal carotid artery, internal jugular vein, and glossopharyngeal, vagus, spinal accessory, and hypoglossal nerves as well as the cervical sympathetic nerve.

Understanding the relative contents and locations of the foramina surrounding the nasopharynx allows the clinician to predict the extent of tumor spread based on cranial nerve examination. Six foramina are adjacent to the wall of the nasopharynx: (1) foramen lacerum, (2) foramen ovale, (3) foramen spongiosum, (4) carotid canal, (5) jugular foramen, and (6) hypoglossal canal. The foramen lacerum and the foramen ovale offer little resistance to tumor spread into the cranium, and their close proximity to the cavernous sinus and cranial nerves II, III, IV, and VI explains the frequency of cranial nerve palsies at diagnosis.

The lymphatics of the nasopharynx drain either through direct efferent channels to the deep lymph nodes of the posterior triangle or first through the lateral pharyngeal wall to the retroparotid or lateral pharyngeal lymph nodes and then on to the upper jugular chain. Some channels may pass directly to the jugulodigastric chain. Lymphatic channels often cross the midline and offer ready access to both sides of the neck.[1]

NASOPHARYNGEAL CARCINOMA

Natural History

Epidemiology Nasopharyngeal carcinoma (NPC) is unique among squamous cell carcinomas of the head and neck in its geographic epidemiologic profile. Although it is a rare tumor in most parts of the world, in regions of southeastern Asia, its incidence soars to 27.3 in 100,000. The highest incidence occurs in Taiwan, where 98% of the population is Chinese, 90% of whom originate from the Guangdong province. There is a decrease in incidence in northern China, falling to 3 in 100,000 in northern provinces. The Japanese, who trace their origins to the Mongoloid region, have an incidence of just 1 in 100,000 for both men and women. In Europe and North America, the incidence is 1 in 100,000. In China, the disease rate rises after age 20 years and falls after age 60; the mean age is 40 to 50 years. The male to female ratio is 3 to 1.

Dietary Associations. Because of the strikingly high incidence of NPC among the "boat people" of southern China (54.7 in 100,000), speculation about their unique environmental exposures led to the examination of salted fish as a risk factor. Salted fish has been shown to contain strong carcinogens and mutagens. Animal studies support the association: 5 of 22 rats fed with low doses of Cantonese-style salted fish developed malignant tumors of the nasal or paranasal sinus cavities, whereas no similar tumors developed in controls.[2] This finding was confirmed by Yu et al in 1989, who demonstrated a dose-dependent relationship between salted fish intake and tumor development in rats.[3] The amounts of salted fish fed to rats were in the range of human consumption.

Yu reviewed seven case-control studies examining the connection between diet and NPC,[4] four of which were conducted in the Cantonese ethnic group. All of these studies demonstrated a significant positive association between salted fish intake and risk of NPC. In two longitudinal studies, childhood exposure, in particular exposure from the time of weaning, was found to be more strongly related to the risk of developing NPC than adult exposure alone. Salted fish is a diet staple in these populations, and the process of preserving fish instills a variety of potential carcinogens into the food. Low levels of several volatile nitrosamines are detected in salted fish, most of which can induce nasal and paranasal tumors in animals. Bacterial mutagens are also present. Substances activating the Epstein-Barr virus (EBV) have also been identified in salted fish. Together, these factors may combine to explain the findings of these case-control studies.

Epstein-Barr Virus Associations. Epstein-Barr virus is a double-stranded DNA virus that, in addition to being the cause of infectious mononucleosis, is implicated in a variety of lymphoid neoplasms: endemic Burkitt's lymphoma, post-transplantation lymphoproliferative disorders, subsets of Hodgkin's disease, and nasal T-cell lymphoma.[5] The connection between EBV and NPC was first made in the late 1960s, when serologic studies demonstrated elevated immunoglobulin (Ig)A and IgG titers against viral antigens. Two major lines of evidence for the role of EBV have been elucidated: serologic studies and nucleic acid studies.

Serologic Studies. Elevated levels of IgA and IgG antibodies directed against the viral capsid antigen and early antigen in patients with NPC have been documented by many studies. Immunoglobulin A can be detected in 80 to 85% of patients with NPC. These profiles are uncommon in patients with carcinomas in other parts of the head and neck. Response to treatment yields a corresponding decrease in levels; increasing titers are associated with progression of disease. Anti-EBV nuclear antigen-1 IgA has also proven to be a sensitive indicator of the presence of NPC: levels are present in 90% of patients with NPC compared with 7% of normal controls and 14% of patients with non-NPC cancers. These serologic tests have implications for screening in high-risk areas and for monitoring the progression of the disease. Although serologic studies paint a convincing picture for the involvement of EBV in NPC, they do not provide direct evidence for EBV in the pathogenesis of the disease.

Nucleic Acid Studies. Direct evidence of the role of EBV in the pathogenesis comes from studies of EBV ribonucleic acid (RNA) in NPC cells, the presence of which indicates infection with the virus. In a study of 120 cases of NPC, including 31 with World Health Organization (WHO) type I histology (keratinizing squamous cell carcinoma), in situ hybridization demonstrated EBV RNA in all samples.[6] Similar studies have also demonstrated EBV-encoded RNA in carcinoma in situ and areas of dysplasia whether or not associated with a primary invasive lesion.[7] These findings are in distinction to patients without NPC. The consistent presence of EBV in preinvasive lesions strongly supports EBV infection as an early event in the malignant transformation of nasopharyngeal epithelial cells.

Clonality. The role of EBV in the pathogenesis of NPC has been proven one step further by the demonstration of a clonal population of EBV DNA in invasive lesions.[7] The mechanism of transformation, however, is not well understood. Epstein-Barr virus elaborates several proteins; one candidate is latent membrane protein (LMP)-1, which is found in 65% of invasive lesions. This protein is required for EBV-mediated transformation of B lymphocytes; transfection of the protein in epithelial cells causes keratin down-regulation and morphologic changes toward spindling of cells. Taken together, the role of EBV in NPC is convincing.

Immunity *Biomarkers.* Epstein-Barr Virus Antigens. Given the close association of EBV and NPC, it is not surprising that considerable attention has been focused on EBV antigens as biomarkers to predict the NPC clinical course. Serologic studies can be useful in this regard: response to therapy is indicated by declines in titers of IgA anti–viral capsid antigen and early antigen. Conversely, disease progression correlates to rising titers of IgA.[8] Epidermal growth factor and its receptor (EGFR) have been associated with malignant progression in squamous cell carcinomas. Additionally, the tumor proliferation marker Ki67 has similar associations with malignant progression. Increased expression of EGFR and elevated numbers of Ki67-positive cells have been correlated with expression of EBV LMP-1 and were more frequently seen in cases of advanced NPC.[9]

Growth Factors Vascular endothelial growth factor (VEGF) is a potent angiogenic growth factor elaborated by many malignant tumors. Elevation of this growth factor has been shown to predict poor prognosis in several tumor types. Serum VEGF was found to be elevated in patients with metastatic NPC but not in patients with nonmetastatic NPC. This correlates with findings of increased tumor microvessel density in tumor from patients with metastatic NPC.[10] Because VEGF is elevated only in metastatic NPC, it is of limited value in early diagnosis of the disease. Other growth factors are currently under investigation and may prove useful as additional therapies are developed for NPC.

Histology

World Health Organization Classification and Outcomes The WHO has classified NPC into three histopathologic groups based on light microscopy. WHO type I is a keratinizing squamous cell carcinoma similar to others of the head and neck. It exhibits abundant keratin formation with intercellular bridges and various degrees of differentiation. It can be subdivided into well-differentiated (G1), moderately differentiated (G2), or poorly differentiated (G3) grades. WHO type II is a nonkeratinizing form with greater pleomorphism, scant keratin formation, and a variety of patterns. It exhibits a pavemented or stratified pattern with clear cell margins. WHO type III is an undifferentiated tumor originally described as lymphoepithelioma by Schmincke and Regaud. It is characterized by greater heterogeneity of cell size, indistinct cell borders, and prominent nucleoli. Types II and III both typically demonstrate lymphocytic infiltration; in contrast, it is less commonly seen in type I. Types II and III are endemic forms of NPC and are classically associated with EBV (Figure 60–1).

The clinical implications of the differential types are considerable. In a National Cancer Data Base study of 5,069 patients diagnosed with NPC in the United States between 1985 and 1989, there were substantial differences between 5-year survivals of keratinizing (37%) versus nonkeratinizing (65%) and undifferentiated (64%) NPC[11] (Figure 60–2). Furthermore, differential survival of various patient origin groups correlated tightly with the respective proportion of keratinizing NPC. Origin groups with the best survival had the highest proportion of radioresponsive nonkeratinizing and undifferentiated NPCs; likewise, worst survival groups had the highest proportion of poorly responsive keratinizing carcinomas. Ironically, the more differentiated unilateral tumors had poorer outcomes. This study was performed in the United States; the translatability of results to endemic areas is not known (Figure 60–3).

Staging Systems

The epidemiologic idiosyncrasy of NPC is reflected in the controversies surrounding the staging of this disease. All staging of NPC begins with a thorough physical examination, including endoscopy. Imaging studies, including computed tomography (CT) and magnetic resonance imaging (MRI), are performed to define soft tissue and bony extension. Ho developed a staging system based on his extensive experience treating NPC in endemic areas of southeastern Asia. His system divides all disease into three tumor categories that do not account for nasopharyngeal subsite involvement. Ho's staging system divides the neck into three zones with increasing node status from superior to inferior.

The fourth edition of the American Joint Committee on Cancer (AJCC) staging was developed and applied in the United States, where NPC is nonendemic. It divides all disease into four tumor categories, including T2 accounting for multiple

nasopharyngeal subsite involvement. The neck is treated similarly to other head and neck cancers.

The fifth edition of the AJCC staging system sought to combine the value of each staging system by including components of the Ho and AJCC systems, producing a system that simplifies the nasopharyngeal subsites and preserves the vertical stratification of lymph node involvement. This staging system was applied to 107 patients in the United States and compared with the two previous staging

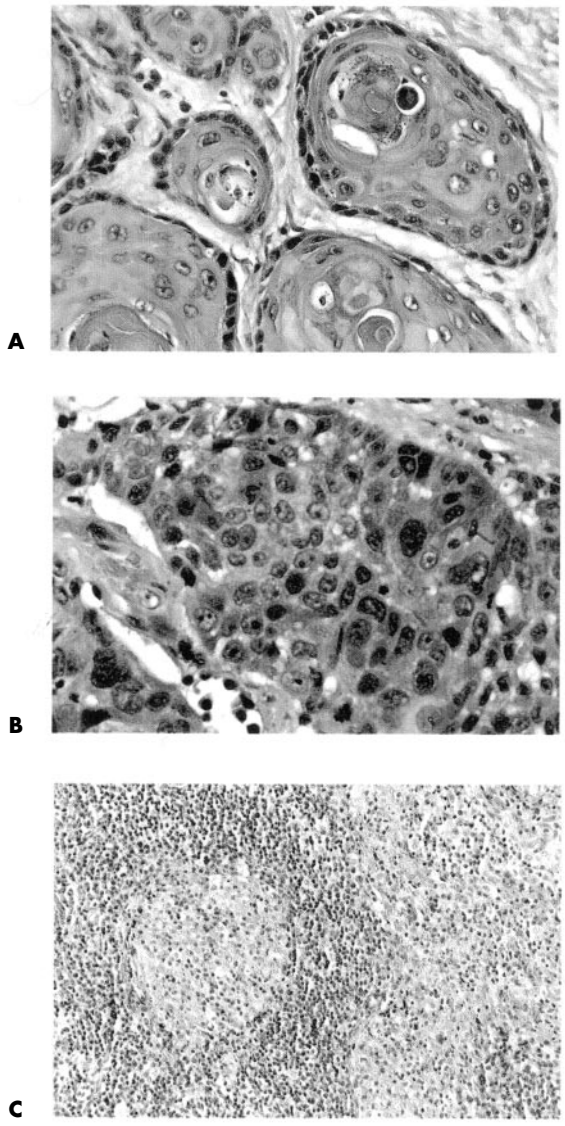


FIGURE 60–1. Nasopharyngeal carcinoma histology. *A*, WHO type I—well-differentiated nasopharyngeal cancer. Note keratin pearls, intracellular bridges, and increased nuclear-to-cytoplasmic ratios but consistent sizes of the nuclei. *B*, WHO type II—Note decreased level of differentiation characterized by increased nuclear pleomorphism. Also note increasing inflammatory infiltrate compared with WHO type I tumors. *C*, WHO type III—classic appearance of a lymphoepithelioma with difficult to distinguish squamous cancer cells in a background of lymphocytes.

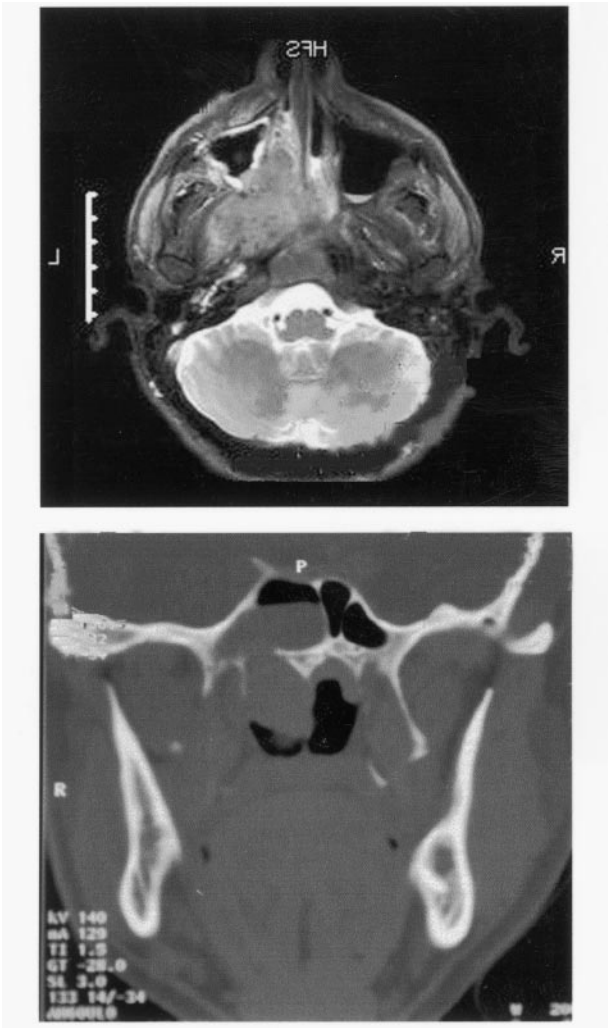


FIGURE 60–2. Magnetic resonance imaging and computed tomographic appearance of invasive nasopharyngeal cancer of the skull base. In the *upper panel*, note the appearance of the tumor in the nasopharynx with extensive invasion into the nasal cavity bilaterally with erosion of the posterior part of the nasal septum and the posterior wall of the maxillary sinus. Note the difference in density of the tumor and the water density of the maxillary sinus mucosa on this T_2-weighted image. Posteriorly, the tumor invades the clivus and area surrounding the brainstem. Laterally, the tumor extends into the pterygoid musculature. In the *lower image*, sphenoid sinus invasion is demonstrated with erosion of the bony floor of the sphenoid sinus.

classifications,[12] and it was found to be superior. A similar study was performed in Singapore on 677 patients with primarily endemic NPC. The revised 1997 AJCC staging was applied and found to be prognostically useful. A direct comparison was not performed.[13] The new staging system should help to standardize staging and facilitate comparison of patient groups undergoing clinical trials around the world (Table 60–1).

Clinical Presentation

The early diagnosis of NPC is impeded by the poor accessibility of the site to routine physical examination. The early symptoms of the disease are often confounding to physicians and empirically treated as non–life-threatening entities for some time before the diagnosis is made. Generally, symptoms fall into one of four general areas of complaint: aural, nasal, neck, and miscellaneous accounted for by cranial nerve involvement (facial pain, diplopia). The classic presentation at the time of diagnosis is a neck mass and conductive hearing loss, often with bloody drainage. Nasopharyngeal carcinoma often arises from the lateral wall of the nasopharynx, near the fossa of Rosenmüller. As the mass enlarges, it obstructs the eustachian tube orifice and induces a serous otitis media.

Site of first symptom in 151 NPC patients was analyzed by Neel et al.[14] The sites were the neck (36%), ear (29%), and nose (21%). Symptoms at the time of admission were compiled by Dickson.[15] A neck mass was present in 60 to 70% of patients, hearing loss in 44 to 53%, nasal obstruction in 27 to 32%, and epistaxis in 34 to 46%. Cranial nerve palsies were present in 3.9 to 15.1 of patients, the most common being cranial nerve VI. Sites of origin in descending order were lateral wall (35.9%), posterior superior wall (27.8%), too extensive to determine (14.4%), and not well described (15.2%).

Facial sensory deficits or facial pain suggests involvement of the foramen lacerum. Invasion of the cavernous sinus is indicated by cranial nerve deficits of cranial nerves III, IV, and VI. Because cranial nerve VI crosses lateral to the carotid artery, this is the first intracavernous cranial nerve encountered. Unilateral deficits of cranial nerves III, IV, VI, and V1 define the superior orbital fissure syndrome, which suggests involvement of the orbit. Deficits of cranial nerves IX, X, and XI indicated extension of tumor through the carotid canal caudad toward the jugular foramen or hypoglossal foramen. Cervical sympathetic involvement may cause Horner's syndrome, ptosis, miosis, and anhidrosis, often seen with lower cranial nerve deficits.[15]

The parapharyngeal space is involved in 75 to 80% of patients by CT. Invasion of local areas is common. The oropharynx is involved in 15% of patients, the nasal cavity in 20%. The bony base of the skull is involved in more than one-quarter of patients, and brain involvement occurs in 3 to 12%. Cranial nerves IV and V are commonly involved; cranial nerve XII may be affected by bulky upper cervical nodes. Cranial nerves IX, X, and XI often reflect jugular foramen involvement; involvement of the cavernous sinus will affect cranial nerves III and IV.[16,17]

Nodal involvement is extremely common in NPC, occurring in 75 to 90% of WHO type II and III histologies at the time of diagnosis. In type III, nodal involvement is bilateral in 60% of patients. Distant dissemination is reported by most major series to be between 5 and 11% at the time of pres-

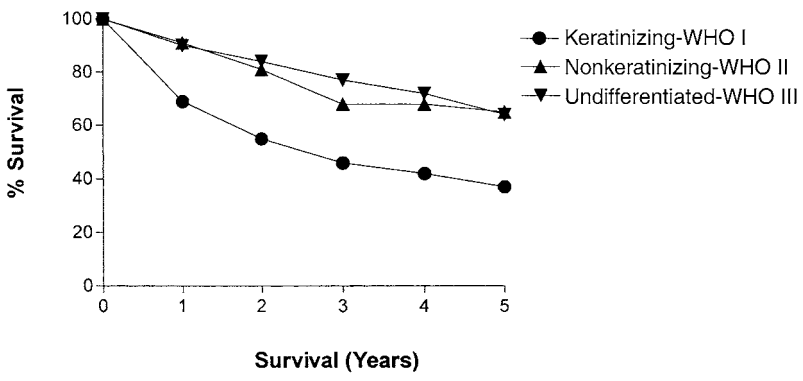


FIGURE 60–3. Five-year survival of patients with nasopharyngeal cancer by WHO classification. Note that the poorer differentiation status portends improved survival.

entation. The most common sites are bone, lung, liver, and extraregional lymph nodes.[18]

Diagnosis

The diagnosis of NPC begins with an accurate history and complete physical examination, including endoscopic visualization of the nasopharynx. This is followed by endoscopic examination of the nasopharynx under general anesthesia for the purpose of tumor staging. Patients who present with neck metastases with unknown primaries should have NPC ruled out by direct biopsy. Imaging studies are done to define the extent of the primary lesion as well as neck metastasis. Magnetic resonance imaging is useful for evaluation of muscle, nerve, or intracranial invasion, whereas CT demonstrates bony involvement more reliably. Olmi et al demonstated that CT can upstage 50% of T2 and T3 lesions.[19] Chest roentgenogram, bone scan, complete blood count, serum chemistries, and liver function tests are recommended. Nutritional support should be instituted when appropriate (Figure 60–4).

TABLE 60–1. Fifth Edition of AJCC Staging Sytem

T1	Tumor confined to the nasopharynx		
T2	Tumor extends to soft tissue of oropharynx and/or nasal fossa T2a: Without parapharyngeal extension T2b: With parapharyngeal extension		
T3	Tumor invades bony structures and/or paranasal sinuses		
T4	Tumor with intracranial extension and/or involvement of cranial nerves, infratemporal fossa, hypopharynx, or orbit		
N0	No regional lymph node metastasis		
N1	Unilateral metastasis in lymph node(s) measuring ≤ 6 cm in greatest dimension above the supraclavicular fossa		
N2	Bilateral metastasis in lymph node(s) measuring ≤ 6 cm in greatest dimension above the supraclavicular fossa		
N3	Metastasis in a lymph node(s): (a): > 6 cm in greatest dimesion (b): Extension to the supraclavicular fossa		
M0	No distant metastasis		
M1	Distant metastasis		
Stage I	T1	N0	M0
Stage IIA	T2a	N0	M0
Stage IIB	T1	N1	M0
	T2b	N0–1	M0
Stage III	T1	N2	M0
	T2	N2	M0
	T3	N0–2	M0
Stage IVA	T4	N0–2	M0
Stage IVB	Any T	N3	M0
Stage IVC	Any T	Any N	M1

AJCC = American Joint Committee on Cancer.

TREATMENT

External Beam Radiotherapy Historically, radiotherapy has been the standard of treatment for locoregionally confined NPC. In such patients, it is administered with the intent to cure. The location of primary disease with respect to vital structures, the propensity for bilateral lymphatic metastases, and the inherent radiosensitivity of NPC have made radiation therapy the primary treatment modality, although surgery is sometimes employed for local recurrences after radiotherapy failure. Neck dissection may be performed to address bulky advanced neck disease.

The entire nasopharynx and bilateral retropharyngeal, jugulodigastric, low neck, posterior chain, and supraclavicular lymph nodes are included in the initial radiotherapy target. At least 45 to 50 Gy at 1.8 to 2.0 Gy per fraction (day) are delivered. The primary tumor dose is usually boosted to 66 to 70 Gy based on tumor stage. Because NPC is generally more radiosensitive than other squamous cell carcinomas of the head and neck, moderate-dose radiotherapy followed by neck dissection is not recommended. However, proven residual disease after full-course radiotherapy should be addressed with neck dissection 6 to 8 weeks following completion of radiotherapy.[18] M. D. Anderson Cancer Center has reported survival on a consecutive series of 378 patients. Respective actuarial survival rates at 5-, 10-, and 20-year follow-up were 48%, 34%, and 18%.[20] Local control rates with radiation alone for T1, T2, T3, and T4 tumors are 85 to 95%, 80 to 90%, 60 to 75%, and 40 to 60%[18]; most failures are isolated, and nearly all occur within 5 years of therapy.

The treatment of the N0 neck is guided by two principles. First, the high probability of neck involvement is unrelated to the stage of the primary disease, and bilateral involvement is frequent. Thus, it is not possible to predict reliably a side of disease or which patients may have subclinical local metastases. Second, although salvage surgery of the untreated neck is effective in 80 to 90% of patients, the appearance of disease in the untreated neck is strongly associated with distant metastasis.[21,22] Because patients without subclinical metastases cannot be distinguished from those with subclinical metastases, coupled with the fact that close surveillance and salvage surgery are associated with adverse outcomes, the elective treatment of the necks of all patients is warranted, even though, inevitably, some patients will be receiving treatment unnecessarily. Regional control rates with radiation alone are good: reported 5-year regional control rates for N0, N1 and N2, and N3 categories are 90 to 100%, 80 to 90%, and 60 to 80%, respectively. It is unclear whether local recurrence increases the risk of distant metastasis.

Timing of radiation delivery is important. It is well established that a prolonged course of radiation

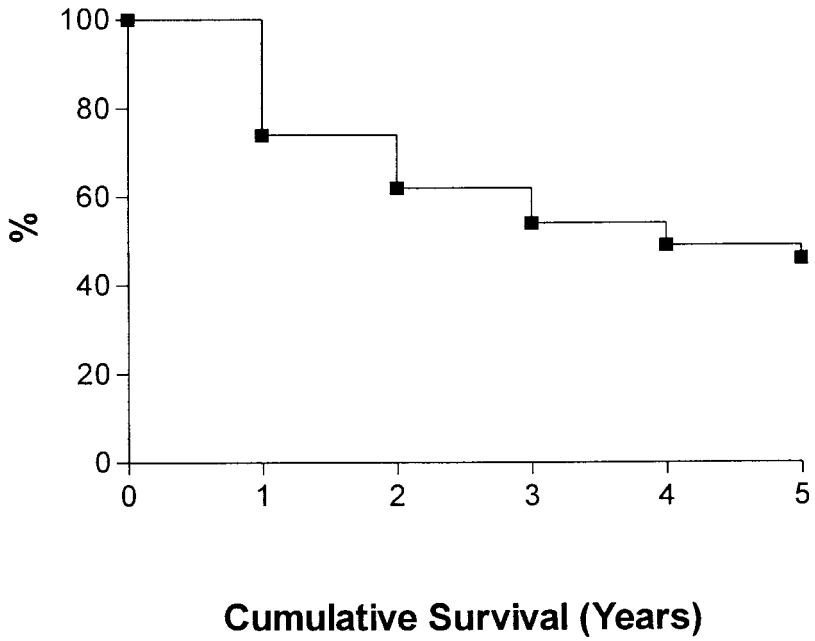


FIGURE 60–4. Five-year survival of non-Hispanic/non-Asian US citizens across all histologic types. All types and stages are shown in this cross-sectional graph.

delivery portends a poorer prognosis. Hyperfractionation involves the use of a smaller dose per fraction given; accelerated fractionation reduces overall treatment time. Hyperfractionation allows a higher total dose to be delivered without increasing the rate of late complications. This is based on the differential response of acute-reacting tissues, such as tumor, compared with late-reacting tissues. This differential reaction is inversely related to fraction size. By decreasing the total duration of treatment, accelerated fractionation reduces the opportunity for tumor clones to proliferate. Phase II studies indicate mixed results. Interestingly, an interval of at least 8 hours between fractions appears to be necessary to prevent brain or spinal cord injury.[18]

The ability to achieve focal exposure and relative ease of access to the nasopharynx have generated interest in endocavitary radiotherapy (brachytherapy) as a means of boosting dose to the primary tumor site following external beam radiotherapy. No phase III studies have been performed to evaluate this modality. However, in a study by Chang et al, the best actuarial survival occurred in a group of patients receiving a total combined external and endocavitary dose of 72.5 to 75 Gy.[23] It was better than the survival of the second experimental arm receiving a > 75 Gy total dose. The authors also concluded that fraction size should be decreased to reduce complications. Fractionated high-dose brachytherapy delivered as a boost to external beam radiotherapy was delivered on an outpatient basis to 29 T1 to T3 patients and 13 T4 patients.[24] The results were compared with matched retrospective controls. Multivariate analysis demonstrated hazard ratios for brachytherapy versus no brachytherapy of 0.24 for time to local failure, 0.35 for time to distant failure, 0.31 for disease-free survival, and 0.44 for cause-specific survival. Doses of 73 to 95 Gy were administered with minimal morbidity. Definite indications for endocavitary radiotherapy await further clinical trials.

Predictors of Radiotherapy Failure. The ability to predict which patients are at greatest risk of failure would make it possible to focus more aggressive or adjuvant therapy on those who have the worst prognosis, while sparing those with a favorable outlook the morbidity of aggressive treatment. As expected, tumor category and node category are well-established predictors of survival. Interestingly, tumor category is unrelated to risk of distant metastasis; its impact on survival can be attributed to increasing risk of local failure. Conversely, node category exerts its influence on survival through increasing risk of distant metastasis, recognizing that, overall, the risk of neck failure is low. Other strong predictors of failure include more differentiated tumors (WHO type I) and cranial nerve involvement (Table 60–2).[18,20]

Management of Radiotherapy Failures. Reirradiation of local failures has been examined. Teo et al treated 103 of 123 isolated local failures with reirradiation using mainly external beams as salvage after failure of contemporary high-dose primary radiotherapy.[25] Twenty were treated with nasopharyngectomy with/without radical neck dissection with/without postoperative radiotherapy. With a 20-month median follow-up, reirradiation to high dose achieved a 5-year overall survival of 7% and 5-year local control of 15%, with significant late complications. Surgical salvage was associated with earlier recurrent tumor categories and better survival and local control than the reirradiation group. This relation persisted even when controlling for recurrent T1 and recurrent T2 size lesions.

Photodynamic therapy (PDT) has been used for the treatment of locally recurrent NPC.[26] In a pilot study, five patients with recurrent NPC with a depth less than 10 mm by CT and with no clinical recurrent neck disease underwent PDT with intravenous hematoporphyrin derivative in four and porfimer sodium in one activated by a 630 nm laser light under local anesthesia. Three of five patients were reported to be disease free at 51, 52, and 60 months; the longest survival time was in a patient who had failed reirradiation to a total dose of 136 Gy. There were no complications related to treatment. At the Prince of Wales Hospital in Hong

TABLE 60–2. Predictors of Local Failure in Nasopharyngeal Carcinoma

Tumor category
Tumor volume
Cranial nerve palsy
Base-of-skull invasion
Parapharyngeal invasion

Kong, where NPC is endemic, Tong et al reported their experience with the same agent and light in the treatment of 12 patients with recurrent disease after radiotherapy. All 12 patients demonstrated dramatic response by CT and MRI at 6 months post-PDT. Of the 8 patients for whom cure was sought, 3 remained disease free at 9 to 12 months following a single treatment; the remaining patients achieved useful palliation. Minimal morbidity occurred. Further study of this modality is warranted.[27]

Complications of Radiotherapy. In modern clinical trials, mucositis is the most frequent radiotherapy acute toxicity, occurring in 18 to 84%. Grade 3 mucositis (< 50%) occurs in 20 to 28% of patients.[20] The most common long-term complications of radiotherapy are skin or subcutaneous fibrosis, osteoradionecrosis, radiation myelitis, brain necrosis, temporomandibular joint ankylosis, hypopituitarism, sensorineural hearing loss, and bone atrophy (Table 60–3). In a cohort of patients undergoing radiotherapy for T4 NPC, severe late complications occurred in 23.5% of patients. Neck fibrosis, massive bleeding, and trismus accounted for 71% of the complications in 42 patients.[28]

Chemotheraphy and Radiotherapy. Since NPC is considered both radiosensitive and chemosensitive, it stands to reason that treatment with chemotherapy and radiotherapy protocols would be tried. Problems with overlapping toxicities had precluded such treatments in the past. However, with advances in radiotherapy delivery, chemotherapeutic agents, and supportive care, chemoradiotherapy protocols have been made possible. Although radiotherapy has been effective for treating early-stage disease, patients with bulky lymphadenopathy or supraclavicular metastases are at greater risk of distant metastases and experience a 5-year survival rate of 10 to 40% with radiotherapy alone. The high incidence of distant metastasis combined with the chemosensitivity of NPC makes this malignancy an ideal candidate for the addition of chemotherapy to the treatment regimen.[29]

TABLE 60–3. Long-Term Complications of Radiotherapy

Skin/subcutaneous fibrosis
Osteoradionecrosis
Radiation myelitis
Brain necrosis
Temporomandibular joint ankylosis
Hypopituitarism
Bone atrophy

The goal of chemoradiotherapy is to enhance local control of tumor and to address distant metastases that are not included in the field of radiation. To obtain optimal local control results, the two modalities should be delivered as close in time as possible to achieve a synergistic effect. Toxicity becomes the limiting factor in this setting, and drugs with minimal overlapping toxicities with radiotherapy and each other are favored. As for control of distant metastases, there is no synergistic effect, and the timing of administration is unrelated to radiotherapy. The earlier the treatment, however, the smaller will be the micrometastatic tumor deposits.

Neoadjuvant Chemoradiotherapy. Radiation for stage I and II NPC generally does well. For patients with locoregionally advanced NPC in whom bulky lymph node metastases, local failure, and distant metastasis rates are high, the 5-year survival ranges from 10 to 40%. Knowing that NPC is chemosensitive, it is logical to add chemotherapy to the treatment regimen in this patient population. Early studies of neoadjuvant chemotherapy using three cycles of cisplatin-containing regimens yielded encouraging results with acceptable toxicities.[29] The M. D. Anderson Cancer Center performed a matched cohort study of neoadjuvant chemotherapy consisting of three cycles of 5-fluorouracil (5-FU) and cisplatin followed by radiotherapy. Sixty-one patients were accrued and matched based on tumor category, node category, histology, and level of cervical lymph metastasis. With a median follow-up of 4.9 years, a significant difference in distant metastasis was observed favoring the chemoradiotherapy group (19%) versus radiotherapy alone (34%). This was reflected in the disease-free survival (64 versus 42%) and the overall survival (69 versus 48%). There were no significant differences in acute toxicity between the two groups.

The Institute Gustave Roussy initiated a phase III international multicenter randomized trial com-

paring neoadjuvant chemotherapy given as three cycles of bleomycin, epirubicin, and cisplatin followed by radiotherapy with radiotherapy alone. A total of 339 patients with undifferentiated NPC with N2 and N3 (AJCC) were accrued and randomized to the treatment arms. With a median follow-up of 49 months, a significant difference in disease-free survival favoring the chemotherapy arm was observed (54.7 versus 32.7%, $p < .01$). An excess of treatment-related deaths occurred in the chemotherapy arm (8 versus 1%). No difference in overall survival was seen; this was ascribed to the better postrelapse survival in the radiotherapy-alone arm.[30] A prospective randomized trial comparing combined neoadjuvant and adjuvant chemoradiotherapy with radiotherapy alone was conducted at the Prince of Wales Hospital in Hong Kong. Eighty-two patients with Ho's N3 staging or any node stage with nodes > 4 cm were randomized to receive two cycles of cisplatin and 5-FU prior to radical radiotherapy followed by four cycles of postradiotherapy chemotherapy. The overall response rate to neoadjuvant chemotherapy was 80% (19% complete response, 62% partial response). With a mean follow-up of 28 months, there were no statistically significant differences between the two arms in terms of acturarial survival, disease-free survival, locoregional relapse rate, distant metastatic rate, and median time to relapse.[31] The absence of difference between the two arms may be explained by the low number of patients enrolled in the study combined with the good overall survival of the radiotherapy control group (80% at 2 years). The low levels of chemotherapy used in this trial may also explain the lack of survival benefit with the experimental arm.

Thus, although neoadjuvant and adjuvant chemoradiotherapy may alter the natural history of locoregionally advanced NPC by prolonging disease-free survival, no randomized trials offer evidence of improving overall or disease-specific survival.

Concurrent Chemoradiotherapy. The use of concurrent chemoradiotherapy was limited by concerns for overlapping toxicity, as experienced in earlier head and neck trials using 5-FU, bleomycin, and mitomycin-C. Two concurrent chemoradiotherapy phase II studies not using cisplatin did not suggest any benefit in overall survival.[32,33] Cisplatin is a good candidate for concomitant treatment because its toxicities do not overlap with those of radiotherapy: myelosuppression is uncommon. Because cisplatin is independently one of the most active agents against NPC, it was employed in a series of 27 patients treated in a concurrent manner resulting in superior disease-free survival, overall survival, and incidence of distant metastasis compared with historical controls.[34]

Based on this, a Head and Neck Intergroup study for patients with stage III/IV NPC was initiated; it randomized patients to a standard radiotherapy control arm (2 Gy/fraction to 70 Gy over 7 weeks) or to a concurrent cisplatin and adjuvant cisplatin and 5-FU in which cisplatin was given every 3 weeks, totaling three doses during radiotherapy. This was followed by cisplatin and 5-FU for three cycles. After the first planned interim analysis with 138 evaluable patients, the trial was closed. About 25% of patients had WHO type I histologies. At a mean follow-up of 2.7 years, 3-year overall survival rates are 78% and 47% for the chemoradiotherapy and radiotherapy arms. The median progression-free survival was 13 months in the radiotherapy arm compared with 52 months in the chemoradiotherapy arm. No fatal toxicity events related to planned treatment occurred; however, grade 3 to 4 toxicity occurred more frequently in the chemoradiotherapy arm (76 versus 50%). Both progression-free survival and 2-year overall survival were statistically significant in favor of concurrent chemoradiotherapy followed by adjuvant chemotherapy.[35] Because of the rarity of this disease, all stage III and IV patients were included, thus representing a heterogeneous group with respect to tumor and node categories. Forty-four percent of patients in the radiotherapy arm and 45% of patients in the chemoradiotherapy arm exhibited type III histologies. The applicability of this heterogeneous study population to treatment of patients in endemic areas, where up to 90% of histologies can be expected to be type III, is unknown. Type 1 and types 2 and 3 differ in their association with EBV as well as in their reported response to radiotherapy,[31] which confound the applicability of this study to populations from endemic areas. A further limitation of this study is the combination of concomitant cisplatin and adjuvant cisplatin and 5-FU. The contribution of concurrent and adjuvant chemotherapy treatment cannot be separated in this study.[35]

Radiotherapy-Chemotherapy Summary. The addition of chemotherapy to standard radiotherapy in

patients with advanced NPC has been supported by clinical studies. Neoadjuvant chemotherapy has been shown to have an impact on preventing distant metastasis,[20] although its efficacy in preventing locoregional recurrence is not clear. To date, no randomized trial has demonstrated definite improvement in overall survival from either neoadjuvant or adjuvant chemotherapy.

Concomitant chemotherapy offers more promising results; however, the only phase III study contains both concomitant and adjuvant chemotherapy in its only experimental arm; thus, definite conclusions regarding the contribution of concomitant chemotherapy cannot be definitively assessed. Further, the heterogeneity of histologic types in this North American study cannot be assumed to be applicable to patient populations in endemic areas. Still, the high response rate to cisplatin-based regimens is encouraging, and reports of patients with distant metastasis achieving long-term disease-free survival warrant further investigation. The benefit of high-dose chemotherapy with autologous bone marrow or peripheral stem cell support has not been investigated.

Overall Treatment Summary. Radiotherapy as a single-modality treatment remains the standard of reference for NPC clinical trials. This may be augmented by brachytherapy in selected patients. Chemotherapy may be added in patients with advanced locoregional disease or recurrent disease or in those with distant metastases in the context of clinical trials, but chemotherapy remains investigational. Other modalities for treatment of recurrent disease include PDT, which remains in its infancy with respect to treatment of NPC.

Further clinical trials are needed to assess the role of chemotherapy in early-stage NPC and to define the optimal time of administration of chemotherapy with respect to radiotherapy. Also, the applicability of North American concomitant chemoradiotherapy trials to patient populations from endemic countries remains to be defined. Additional treatment approaches such as PDT hold promise in selected patients but await clinical evaluation.

Summary

Nasopharyngeal carcinoma is a common malignancy in certain populations but represents less than 1% of the malignancies in the United Stated. Dramatic improvements in survival with the delivery of platinum-based therapies in conjunction with external beam radiotherapy for advanced disease has replaced radiation-only treatment regimens. Advances in imaging modalities of the skull base have allowed for the more accurate staging of this malignancy. Efforts in examining viral serologies have not resulted in predictable biomarkers of survival or response. Efforts in molecular medicine are continuing.

JUVENILE NASOPHARYNGEAL ANGIOFIBROMA

A clinical presentation of epistaxis and nasal obstruction in an adolescent male patient classically represents a juvenile nasopharyngeal angiofibroma (JNA) until proven otherwise. This is a relatively rare sporadic neoplasm that represents approximately 0.05% of all head and neck neoplasms. Clinically, this tumor may be locally invasive, even eroding adjacent skull base bone. This tumor is nonencapsulated and is highly vascular but has not been shown to be metastatic. Incomplete excision of JNA may result in its recurrence, but often microscopic disease is left after excision, and in the majority of individuals, the neoplasms do not recur. Spontaneous regression of the neoplasm has been identified after incomplete gross excision, but untreated angiofibromas have not been shown to regress spontaneously.[36] Several hypotheses regarding the source of this mass have been advanced, but the origin of this tumor as an anomaly of embryology or as part of a genetic syndrome has never been substantiated. The anatomic site of origin in nearly all cases appears to be in the nasopharynx at the superior aspect of the sphenopalatine foramen. The tumor develops and remains submucosal in a complicated anatomic compartment that defies simple surgical extirpation with generous margins.

Natural History

Over time, JNAs increase in size and may involve extensive regions adjacent to the nasopharynx. Anteriorly, the tumor can extend submucosally to involve the posterior aspect of the septum and roof of the nasal cavity at the choanae. The lesion may then extend further anteriorly into the nasal cavity. Superiorly, the tumor will cause flattening of the floor

and posterior aspect of the sphenoid sinus and may enter the sinus and fill it. Tumors usually invade the sphenopalatine foramen and will involve the pterygopalatine fossa. Rarely, anterior invasion into the maxillary sinus will occur after erosion of the pterygoid plates and invasion of the posterior wall of the maxillary sinus. Juvenile nasopharyngeal angiofibroma may also extend laterally into the area of the cheek and extend into the infratemporal fossa. If involvement of the pterygopalatine fossa occurs, tumors may extend into either orbital fissure and gain access to the cavernous sinus. Approximately 10% of tumors will have invaded intracranially at the time of diagnosis, but they typically remain extradural. Intracranial tumor may also be encountered after nasoethmoidal extension up to the level of the cribriform plate.

Differential Diagnosis and Staging

The differential diagnosis of lesions within the nasopharynx in the age group affected by JNA includes both benign and malignant processes. Other benign conditions associated with nasal obstruction include adenoid and turbinate hypertrophy, nasal polyposis, antral choanal polyps, and nasopharyngeal cysts. Other benign neoplasms of the nasopharynx include chordomas, angiomatous polyps, teratomas (eg, dermoids), fibromas, hemangiomas, gliomas, fibrous dysplasia, chondromas, and rhabdomyomas.[37] Soft tissue malignancies in the differential diagnosis include rhabdomyosarcoma, NPC, and lymphomas.

There have been at least three staging systems proposed for JNA since the 1970s, but Radkowski et al proposed the most recent update in 1996.[38] This system appears to reflect a multi-institutional consensus of lesions grouped similarly and also reflects current treatment strategies for this tumor on a stage-dependent basis (Table 60–4).

Clinical Presentation

The average and median ages for the presentation of JNA are 12.5 and 14 years, respectively. Earlier-stage lesions will present with epistaxis and nasal obstruction, and, clearly, symptoms and signs at the initial patient encounter will have predictive value for the anatomic extent of the disease. For example, significant lateral extension of the disease may cause serous otitis media and conductive hearing loss, facial deformity including a cheek mass, and proptosis. Cranial nerve findings are rare but may exist and be attributable to involvement of cranial nerves I to VI.

Diagnostic Workup

Once the clinical presentation suggests a possible differential diagnosis, the radiologic workup is performed. Presently, there is little utility in any plain film analysis of the intricate anatomy of the nasopharynx. Typically, a contrast-enhanced CT and MRI with gadolinium would be the gold standard radiographic evaluations for JNA. The CT scan is advantageous for delineating the extent of any skull base encroachment or invasion that would help the surgeon determine the extent of resection necessary based on bony compartmentalization of the skull base. Additionally, in the differential diagnosis of JNA, other potential benign and malignant processes exist that would have radiographic appearances different from JNA (eg, osteosarcomas or an NPC with extensive bony skull base invasion). The MRI is particularly helpful in identifying soft tissue characteristics of the lesion that would increasingly suggest that it is JNA as opposed to another neoplasm. The lesion will demonstrate contrast enhancement on both studies, and any vascular flow voids within the lesion will be identified on MRI. The usual nearly

TABLE 60–4. Staging of Juvenile Nasopharyngeal Angiofibroma

Stage	*Description*
IA	Limited to nose/nasopharyngeal vault
IB	Extension into one or more sinuses
IIA	Minimal extension into pterygopalatine fossa
IIB	Occupation of whole pterygopalatine fossa with or without orbital bone erosion
IIC	Infratemporal fossa extension with or without posterior extension to pterygoid plates
IIIA	Erosion of skull base, minimal intracranial extension
IIIB	Erosion of skull base, extensive intracranial involvement with/without cavernous sinus involvement

Adapted from Radkowski et al.[38]

uniform enhancement of this lesion radiographically distinguishes it from other vascular entities such as arteriovenous malformations. Additionally, other accompanying features of the soft tissue, including reactive inflammation of the sinus or nasal mucosa or postobstructive sinusitis, will be well delineated on MRI (Figure 60–5).

Because of the characteristic clinical presentation and radiographic appearance, few clinicians would feel compelled to obtain tumor biopsies as part of the workup. Clearly, a biopsy of this lesion could lead to massive hemorrhage and is discouraged. However, there are lesions reported in the literature that resemble JNA by all of the above criteria

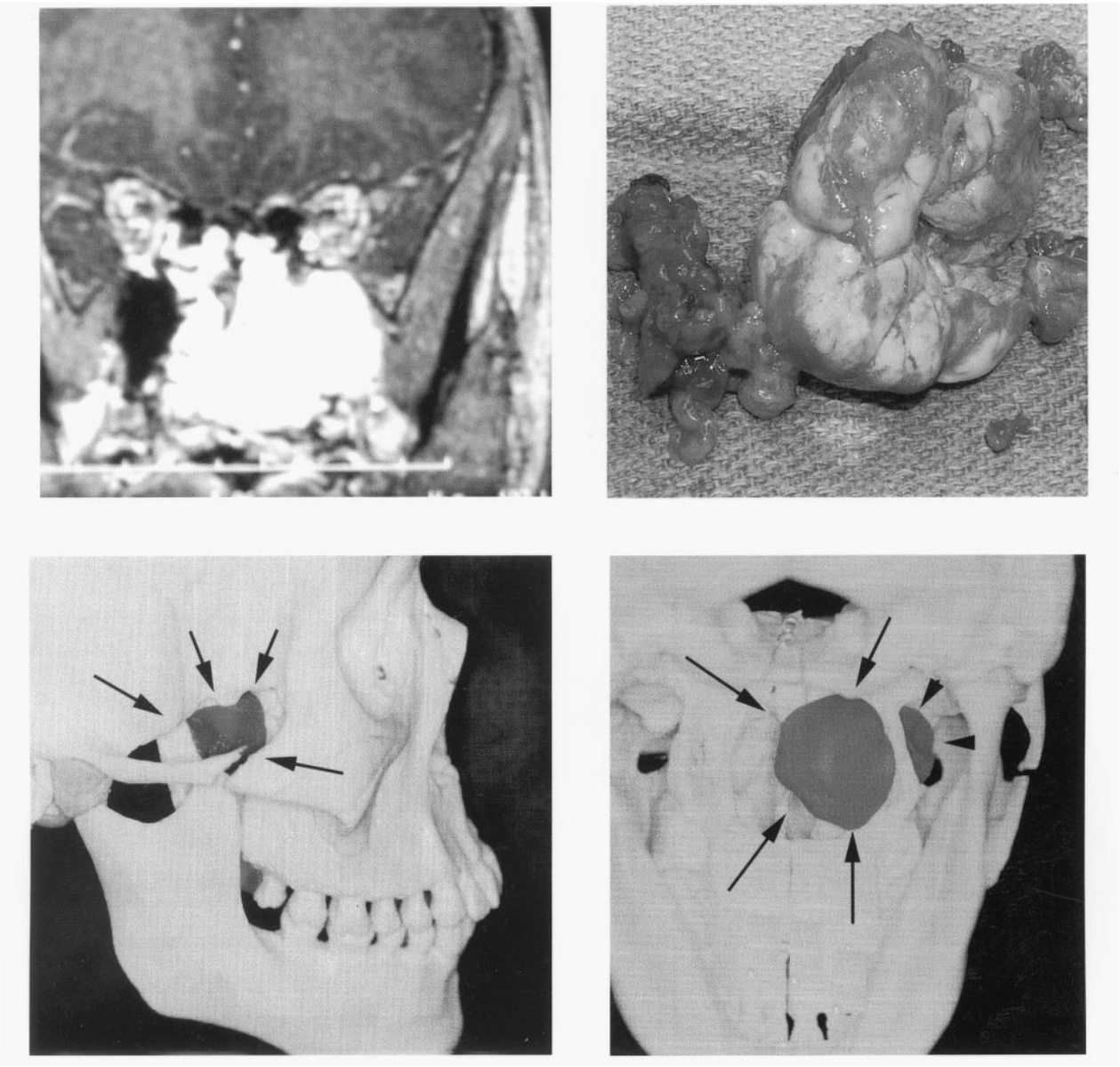


FIGURE 60–5. Juvenile nasopharyngeal angiofibroma. *Upper left*, Coronal magnetic resonance imaging (MRI) appearance of a juvenile nasopharyngeal angiofibroma occupying the nasopharynx with extension into the maxillary and ethmoid sinuses with lateral extension into the infratemporal fossa. *Upper right*, Gross anatomic view of the juvenile nasopharyngeal angiofibroma in the MRI. Note the pseudoencapsulated appearance of the tumor. *Lower left*, Lateral view depicting extension of tumor into infratemporal fossa from the nasopharynx. *Lower right*, Base view of depiction of tumor originating in the nasopharynx with lateral extension into the skull base region with involvement of the pterygoid plates and the infratemporal fossa.

but turn out to be other histologic diagnoses, including rhabdomyosarcoma.[37] Originally, angiography was a diagnostic test of JNA confirmation, but both CT and MRI eliminate the need for this test as part of the workup in most cases. However, lesions currently undergo preoperative embolization with carotid angiography at many major centers. This technique allows for a decrease in blood loss during surgery by 60 to 80%. Typically, embolization agents are directed via the external carotid circulation through the internal maxillary artery, but significant blood supply can be derived directly from the internal carotid artery and ethmoidal arteries.

Treatment

Surgery Because JNAs can involve multiple anatomic sites at the skull base, surgical treatment options can be applied on a more or less stage-specific basis. Additionally, the experience and surgical preference of the otolaryngologist or skull base team may affect which approach is chosen for a particular tumor. Endoscopic approaches have been suggested for small tumors at least for the smallest stage I tumors, and this has become a preferred approach for selected tumors at some institutions.[39] Transpalatal approaches have been used to remove JNA from the nasopharynx when there is limited lateral extent of the tumors since the 1960s. Typically, this resection involves an approach to the soft palate with removal of the third molar, hard palate, and pterygoid plates with a bur or bone punches. This has allowed access to the primary tumor and its lateral extent behind the maxillary sinus. This approach, however, does not allow further lateral access, and excision of the tumor is without a surgical margin. However, several authors agree that residual microscopic disease is associated with involution. Other surgeons approach lesions of limited extent transfacially through a lateral rhinotomy and medial maxillectomy approach. Midfacial degloving techniques may also be used. Lesions with extension into areas outside the nasopharynx will often require a combination of skull base approaches to gain adequate exposure.[40–44] Clearly, lateral approaches to the infratemporal fossa are required in resecting lesions with lateral extensions to that region.[45]

Other Modalities Juvenile nasopharyngeal angiofibromas are known to have recurrence rates up to 25%, and extension of disease into extranasopharyngeal sites correlates with increased rates of recurrence. For these reasons, several investigators have considered additional modalities of treatment for extensive JNA. Some institutions will employ radiation therapy as standard care for intracranial disease extension and when there is significant blood supply to the tumor from the internal cartoid artery system. Typically, 30 to 40 Gy are used in standard fractions to gain control of the lesion.[46] Significant further growth of tumor is not noted, but lesions do not completely regress after treatment. Intensity-modulated radiotherapy (IMRT), using multileaf collimators, has been employed in recurrent JNA. In a limited number of patients, response to 3,400 to 4,500 cGy of IMRT was demonstrated on physical examination and radiographically.[47] Because these tumors are associated with puberty in young males, considerable effort has been expended in hormonal therapies as primary or adjunctive treatment for these tumors. Hormonal therapies for this lesion have included the use of diethylstilbestrol and flutamide.

Summary

Juvenile nasopharyngeal angiofibroma is a benign but locally invasive tumor of the anterior skull base. The natural history of the disease in young males strongly suggests a gonadotropic hormonal contribution to the growth of this neoplasm, but this concept has not been fully exploited in the treatment of this neoplasm. Advances in invasive neuroradiology, imaging, and skull base surgery have allowed for approaches to this neoplasm that have resulted in improved outcomes with decreased morbidity.

REFERENCES

1. Gray's anatomy. Edinburgh: Churchill Livingstone; 1995.
2. Huang DP, Ho JHC, Saw D, et al. Carcinoma of the nasal and paranasal regions in rats fed Cantonese salted marine fish. In: de The G, Ito Y, editors. Nasopharyngeal carcinoma: etiology and control. IRAC Sci Publ No. 20 1978. p. 315–28.
3. Yu MC, Nichols PW, Zou X-N, et al. Induction of malignant nasal cavity tumors in Wistar rats fed Chinese salted fish. Br J Cancer 1989;60:198–201.

4. Yu MC. Diet and nasopharyngeal carcinoma. In: Recent progress in research on nutrition and cancer. 1990. p. 93–105.
5. Weiss LM, Chang KL. Association of Epstein-Barr virus with hematolymphoid neoplasia. Adv Anat Pathol 1996;3:1–15.
6. Hamilton-Dutoit SJ, Pallesen G. Detection of Epstein-Barr virus small RNAs in routine paraffin sections using non-isotopic RNA/RNA in situ hybridization. Histopathology 1994;25:101–11.
7. Pathmanathan R, Prasad U, Sadler R, et al. Clonal proliferations of cells infected with Epstein-Barr virus in pre-invasive lesions related to nasopharyngeal carcinoma. N Engl J Med 1995;333:693–8.
8. Henle W, Ho JHC, Henle G, et al. Nasopharyngeal carcinoma: significance of change in Epstein-Barr virus-related antibody patterns following therapy. Int J Cancer 1977;20:663–72.
9. Sanguineti G, Geara FB, Garden AS, et al. Carcinoma of the nasopharynx treated by radiotherapy alone: determinates of local and regional control. Int J Radiat Oncol Biol Phys 1997;37:985–96.
10. Qian C, Zhang XG, Gou X, et al. Elevation of serum vascular endothelial growth factor in male patients with metastatic nasopharyngeal carcinoma. Cancer 2000;88:255–61.
11. Marks JE, Phillips JL, Menck HR. The National Cancer Data Base report on the relationship of race and national origin to the histology of nasopharyngeal carcinoma. Cancer 1998;83:582–8.
12. Cooper JS, Cohen R, Stevens RE. A comparison of staging systems for nasopharyngeal carcinoma. Cancer 1999;86:213–9.
13. Heng DM, Wee J, Fong K, et al. Prognostic factors in 677 patients in Singapore with nondisseminated nasopharyngeal carcinoma. Cancer 1999;85:1912–20.
14. Neel HB, Pearson GR, Weilain LH, et al. Application of Epstein-Barr virus serology to the diagnosis and staging of North American patients with nasopharyngeal carcinoma. Otolaryngol Head Neck Surg 1983;91:255–62.
15. Dickson RI. Nasopharyngeal carcinoma: an evaluation of 209 patients. Laryngoscope 1981;91:333–53.
16. Altun M, Fandi A, Depuis O, et al. Undifferentiated nasopharyngeal cancer: current diagnostic and therapeutic aspects. Int J Radiat Oncol Biol Phys 1995; 32:859–77.
17. Sham JS, Choy D. Prognostic factors of nasopharyngeal carcinoma: a review of 759 patients. Br J Radiol 1990;63:51–8.
18. Sanguineti G, Corvo R. Treatment of nasopharyngeal carcinoma: state of the art and new perspectives. Oncol Rep 1999;6:377–91.
19. Olmi P, Cellai E, Chiavacci A, et al. Computed tomography in nasopharyngeal carcinoma: part I: T-stage conversion with CT-staging. Int J Radiat Oncol Biol Phys 1990;19:1171–5.
20. Geara FB, Sanguineti G, Tucker SL, et al. Carcinoma of the nasopharynx treated by radiotherapy alone: determinants of distant metastasis and survival. Radiother Oncol 1997;43:53–61.
21. Lee AW, Sham JS, Poon YF, Ho JH. Treatment of stage I nasopharyngeal carcinoma: analysis of patterns of relapse and the results of withholding elective neck radiation. Int J Radiat Oncol Biol Phys 1989;17:1183–90.
22. Ho JH. An epidemiologic and clinical study of nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys 1978;4:182–90.
23. Chang JT, See LC, Tang SC, et al. The role of brachytherapy in early-stage nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys 1996;36: 1019–24.
24. Levendag PC, Schmitz PI, Jansen PP, et al. Fractionated high-dose-rate brachytherapy in primary carcinoma of the nasopharynx. J Clin Oncol 1998; 16:2213–20.
25. Teo PM, Kwan WH, Chan AT, et al. How successful is high-dose (> or = 60 Gy) re-irradiation using mainly external beams in salvaging local failures of nasopharyngeal carcinoma? Int J Radiat Oncol Biol Phys 1998;40:897–913.
26. Lofgren LA, Hallgren S, Nilsson E, et al. Photodynamic therapy for recurrent nasopharyngeal cancer. Arch Otolaryngol Head Neck Surg 1995; 121:997–1002.
27. Tong MC, van Hasselt CA, Woo JK. Preliminary results of photodynamic therapy for recurrent nasopharyngeal carcinoma. Eur Arch Otorhinolaryngol 1996;253:189–92.
28. Lin JC, Jan LS. Locally advanced nasopharyngeal cancer: long term outcomes of radiation therapy. Radiology 1999;211:513–8.
29. Chan AR, Teo PM, Leumg TW, Johnson PJ. The role of chemotherapy in the management of nasopharyngeal carcinoma. Cancer 1998; 82:1003–12.
30. International Nasopharynx Cancer Study Group: VUMVAI Trial. Preliminary results of a randomized trial comparing neoadjuvant chemotherapy (cisplatin, epirubicin, bleomycin) plus radiotherapy vs.

radiotherapy alone in stage IV (> N2,M0) undifferentiated nasopharygeal carcinoma: a positive effect on progression-free survival. Int J Radiat Oncol Biol Phys 1996;35:463–9.
31. Chan AT, Teo PM, Leung TW, et al. A prospective randomized study of chemotherapy adjunctive to definitive radiotherapy in advanced nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys 1995; 33:569–77.
32. Turner SL, Tiver KW. Synchronous radiotherapy and chemotherapy in the treatment of nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys 1993; 27:371–7.
33. Souhami L, Rabinowits M. Combined treatment in carcinoma of the nasopharynx. Laryngoscope 1988;98:881–3.
34. Al-Sarraf M, Pajak TF, Cooper JS, et al. Chemoradiotherapy in patients with locally advanced nasopharyngeal carcinoma: a Radiation Therapy Oncology Group study. J Clin Oncol 1990;8:1342–51.
35. Al-Sarraf M, LeBlanc M, Giri PG, et al. Chemoradiotherapy versus radiotherapy in patients with advanced nasopharyngeal cancer: phase III randomized Intergroup study 0099. J Clin Oncol 1998; 16:1310–7.
36. Neel HB, Whicker J, Devine KD. Juvenile angiofibroma: review of 120 cases. Am J Surg 1973;126: 547–56.
37. Burkey B, Koopman C, Brunberg J. The use of biopsy in the evaluation of pediatric nasopharyngeal masses. Int J Pediatr Otorhinolaryngol 1990;20:169–79.
38. Radkowski D, McGill T, Healy G, et al. Angiofibroma. Changes in staging and treatment. Arch Otolaryngol Head Neck Surg 1996;122:122–9.
39. Fagan J, Snyderman C, Carrau R, Janecka I. Nasopharyngeal angiofibroma: selecting a surgical approach. Head Neck 1997;19:391–9.
40. Standefer J, Holt GR, Brown W, Gates G. Combined intracranial and extracranial excision of nasopharyngeal angiofibroma. Laryngoscope 1983;93: 772–9.
41. Gullane P, Davidson J, O'Dwyer T, Forte V. Juvenile angiofibroma: a review of the literature and a case series report. Laryngoscope 1992;102:928–33.
42. Gantz B, Seid A, Weber R. Nasopharyngeal angiofibroma. Head Neck 1992;14:67–71.
43. Ungkanont K, Byers RM, Weber RS, et al. Juvenile nasopharyngeal angiofibroma: an update of therapeutic management. Head Neck 1996;118:60–6.
44. Herman P, Guillaume L, Chapot R, et al. Long-term follow-up of juvenile nasopharyngeal angiofibromas: analysis of recurrences. Laryngoscope 1999; 109:140–7.
45. Haines S, Duvall A. Transzygomatic and transpalatal excision of juvenile nasopharyngeal angiofibroma with intracranial extension. In: Sekhar L, Janecka I, editors. Surgery of cranial base tumors. New York: Raven Press; 1993.
46. Wiatrak BJ, Koopman CF, Turrisi AT. Radiation therapy as an alternative to surgery in the management of intracranial juvenile nasopharyngeal angiofibroma. Int J Pediatr Otorhinolaryngol 1993;28: 51–61.
47. Kuppersmith RB, Teh BS, Donovan DT, et al. The use of intensity modulated radiotherapy for the treatment of extensive and recurrent juvenile angiofibroma. Int J Pediatr Otorhinolaryngol 2000;52: 261–8.

CHAPTER 61

Neoplasms of the Oral Cavity and Oropharynx

Dennis H. Kraus, MD, John K. Joe, MD

The list of neoplasms potentially involving the oral cavity and oropharynx is extensive and diverse. The focus of this chapter is on squamous cell carcinoma (SCC), also called epidermoid carcinoma, the most common malignant neoplasm of the oral cavity and oropharynx. Tumors affecting these regions have significant effects on respiration, deglutition, and speech. The propensity for locoregional recurrence of advanced cancers of the oral cavity or oropharynx necessitates combined therapy, usually surgery and adjuvant radiotherapy. Appropriate treatment planning, therefore, is most effective using a multidisciplinary approach, encompassing the head and neck surgeon, radiotherapist, medical oncologist, prosthodontist, and speech pathologist.

ANATOMY

Both the oral cavity and the oropharynx are composed of distinct anatomic subsites.[1] Malignancy arising at different subsites may vary in behavior, so recognition of this anatomic classification allows for prognostic evaluation and appropriate treatment planning.

Oral Cavity

The oral cavity is defined as the region from the skin–vermilion junction of the lips to the junction of the hard and soft palate above and to the line of the circumvallate papillae below. As illustrated in Figure 61–1, the oral cavity includes the lips, buccal mucosa, upper and lower alveolar ridges, retromolar trigone, hard palate, floor of the mouth, and anterior two-thirds of the tongue (oral tongue).

The lip begins at the junction of the vermilion border with the skin and includes only the vermilion surface or that portion of the lip that comes into contact with the opposing lip. The upper and lower lips form the anterior boundary of the oral cavity and are joined at each lateral end at the oral commissure. The mental branch of the mandibular nerve (V3) exits from the mandible at the mental foramen to provide sensory innervation to the lower lip and may serve as an important pathway for cancer spread into the mandible.

The buccal mucosa includes the membranous lining of the inner surface of the cheeks and lips from the line of contact of the opposing lips to the line of attachment of mucosa of the alveolar ridge (upper and lower) and pterygomandibular raphe. The buccal mucosa forms the lateral wall of the oral cavity. Tumors of the buccal surface may penetrate through the mucosa to involve deeper structures such as the buccal fat pad or subcutaneous tissues of the cheek.

The upper alveolar ridge refers to the mucosa-covered alveolar process of the maxilla, which extends from the line of attachment of mucosa in the upper gingival buccal gutter to the junction with the hard palate. Its posterior margin is the upper end of the pterygopalatine arch. Cancer of the upper alveolar ridge may extend superiorly to involve the floor of the nasal cavity or the antrum of the maxilla.

The lower alveolar ridge refers to the mucosa-covered alveolar process of the mandible, which extends from the line of attachment of mucosa in the buccal gutter to the line of free mucosa of the floor of the mouth. Posteriorly, it extends to the ascending ramus of the mandible. The inferior alveolar nerve is a branch of the mandibular nerve (V3) providing sensory innervation to the gingiva of the lower alveolar ridge. The course of the mandibular nerve may serve as a route of spread for tumors of the lower alveolar ridge to the skull base.

The retromolar trigone (pterygomandibular raphe) is the triangular region from the level of the posterior surface of the last molar tooth to the apex superiorly, where it is attached to the pterygoid

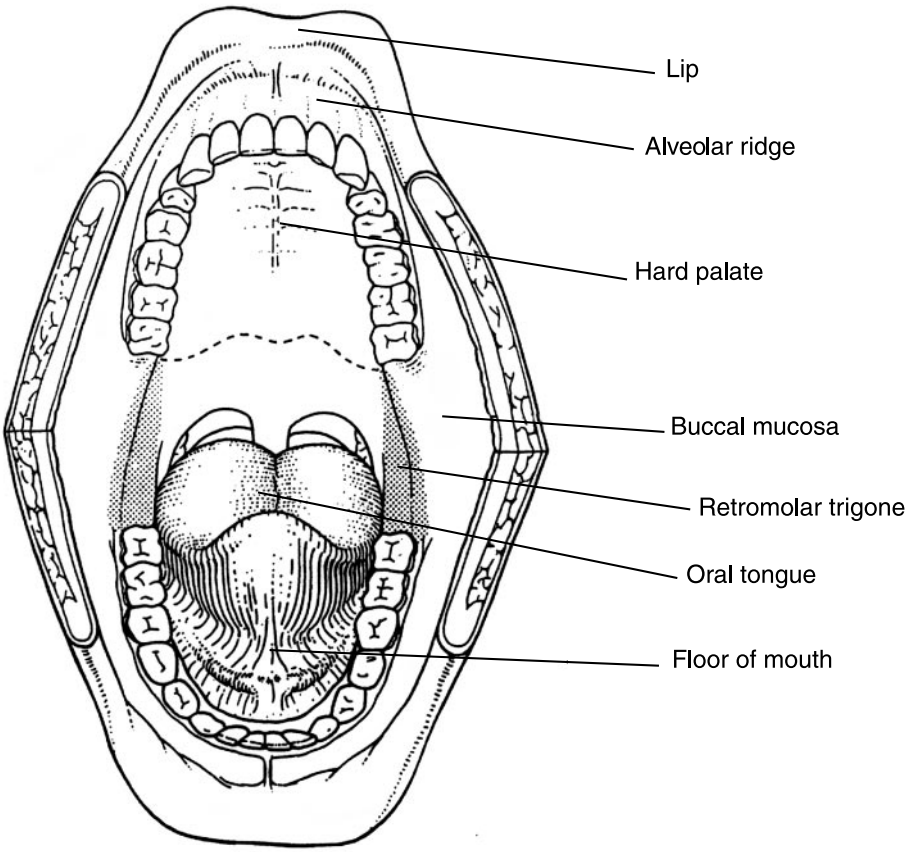


FIGURE 61–1. Anatomic sites of the oral cavity. Adapted from Alvi A, Myers EN, Johnson JT. Cancer of the oral cavity. In: Myers EN, Suen JY, editors. Cancer of the head and neck. 3rd ed. Philadelphia: WB Saunders; 1996. p. 321.

hamulus of the sphenoid bone. Mucosa of the retromolar trigone is tightly adherent to the underlying tendon between the buccinator and superior pharyngeal constrictor muscles and is separated from the periosteum of the ascending ramus of the mandible by the buccal fat pad. Tumors in this region frequently invade the mandible. The site of entry of the inferior alveolar nerve into the mandibular foramen places it at risk for involvement by tumors of the retromolar trigone. Branches from the glossopharyngeal (IX) nerve provide sensory innervation to the retromolar trigone, so cancer of this site often presents with referred pain in the ipsilateral ear.

The hard palate is the semilunar area within the upper alveolar ridge and includes the mucous membrane–covered palatine processes of the maxillary bones and the horizontal parts of the palatine bones. It extends from the inner surface of the superior alveolar ridge to the posterior edge of the palatine bone, forming the roof of the oral cavity. The palatine foramina are located near the junction of the hard palate and soft palate and contain the descending palatine arteries and the anterior, middle, and posterior palatine nerves. Although the mucosa and the periosteum of the hard palate are tightly adherent, the palatine foramina serve as potential routes of spread of cancer from the hard palate to the pterygopalatine fossae and the skull base. In a similar fashion, the incisive canals transmit the nasopalatine nerves and branches of the maxillary nerves (V2) and allow for tumor extension from the hard palate into the nasal cavities.

The floor of the mouth is the semilunar space over the mylohyoid and hyoglossus muscles, extending from the inner surface of the lower alveolar ridge to the undersurface of the anterior two-thirds of the tongue. Its posterior boundaries are the bases of the anterior pillars of the tonsils. The floor of the oral cavity is divided into two sides by the frenulum of the tongue and contains the openings of the ducts of the submandibular and sublingual salivary glands. The lingual nerve provides

sensory innervation to the floor of the mouth and may provide a conduit for tumor extension. The lingual nerve is a branch of the mandibular nerve (V3), as is the auriculotemporal nerve, which provides sensory innervation to the auricle, external auditory canal, and tympanic membrane. Therefore, cancer of the floor of the mouth may present with referred pain in the ipsilateral ear. Likewise, the submandibular duct may serve as a potential route for tumor spread. Cancer of the floor of the mouth may involve the mandible or tongue by direct extension.

The anterior two-thirds of the tongue are referred to as the oral tongue and include the freely mobile portion of the tongue that extends anteriorly from the line of the circumvallate papillae to the undersurface of the tongue at the junction of the floor of the mouth. It is composed of four areas: the tip, lateral borders, dorsum, and undersurface (nonvillous ventral surface of the tongue). The anterior two-thirds of the tongue or oral tongue are part of the oral cavity, and the posterior one-third or base of the tongue is part of the oropharynx. The hypoglossal nerve (XII) provides motor innervation to the tongue. The oral tongue receives sensory innervation from the lingual branch of the mandibular nerve (V3) and taste sensation from the chorda tympani branch of the facial nerve (VII). As with the floor of the mouth, cancers of the oral tongue may present with referred pain in the ipsilateral ear owing to involvement of the mandibular nerve (V3).

Lymphatics As illustrated in Figure 61–2, regional lymph node groups in the neck are grouped into various levels for ease of description.[1] Level I contains the submental and submandibular triangles bounded by the posterior belly of the digastric muscle, the hyoid bone inferiorly, and the body of the mandible superiorly. Level II contains the upper jugular lymph nodes and extends from the level of the hyoid bone inferiorly to the skull base superiorly. Level III contains the middle jugular lymph nodes from the hyoid bone superiorly to the cricothyroid membrane inferiorly. Level IV contains the lower jugular lymph nodes from the cricothyroid membrane superiorly to the clavicle inferiorly. The posterior limit of levels II, III, and IV is the posterior border of the sternocleidomastoid muscle (SCM). Level V contains the lymph nodes in the posterior triangle bounded by the anterior border of the trapezius posteriorly, the posterior border of the SCM anteriorly, and the clavicle inferiorly.

Metastasis to regional lymph nodes occurs in a predictable fashion through sequential spread. Regional lymph nodes at highest risk for metastases from primary SCCs of the oral cavity include those at levels I, II, and III, collectively known as the supraomohyoid triangle.[2,3] Tumors of each anatomic site have their own predictable first-echelon lymph nodes of involvement.[1] Nodal metastasis of SCC of the lip is uncommon but tends to involve initially adjacent submental and submandibular nodes and then jugular nodes. Tumors of the buccal mucosa tend to spread first to submental and submandibular nodes. Cancer of the upper and lower alveolar ridges involve buccinator, submandibular, jugular, and, occasionally, retropharyngeal nodes (Figure 61–3). Lymphatic drainage of the retromolar trigone involves upper jugular nodes, as well as subparotid and retropharyngeal nodes. Like tumors of the alveolar ridges, tumors of the hard palate do not commonly metastasize to regional lymph nodes, but

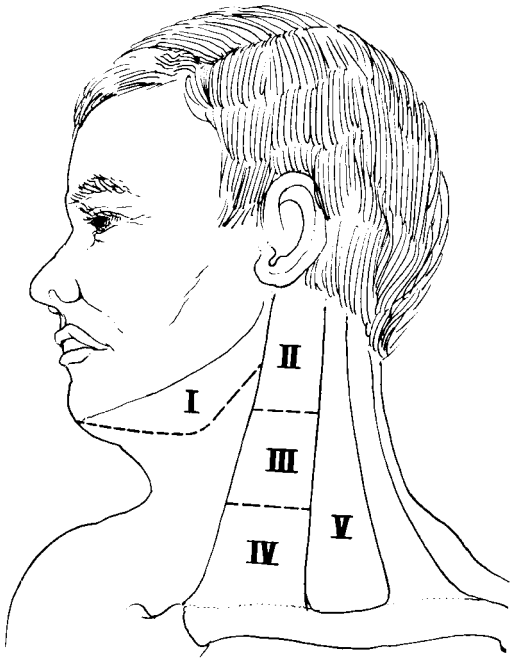


FIGURE 61–2. Cervical lymph node levels. Reproduced with permission from Alvi A, Myers EN, Johnson JT. Cancer of the oral cavity. In: Myers EN, Suen JY, editors. Cancer of the head and neck. 3rd ed. Philadelphia: WB Saunders; 1996. p. 322.

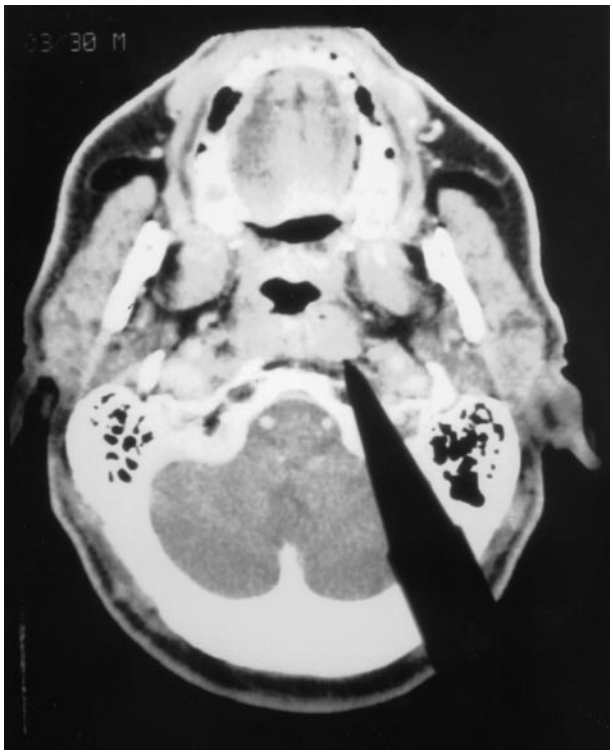


FIGURE 61–3. Axial computed tomographic scan of the neck demonstrating metastasis to a retropharyngeal lymph node (*pointer*).

lymphatic spread may involve buccinator, submandibular, jugular, and, occasionally, retropharyngeal nodes. Cancer of the floor of the mouth tends to spread first to submandibular and jugular nodes. Lymphatics of the tongue involve submental, submandibular, and jugular nodes. Lesions of the midline of the lip, floor of the mouth, or tongue commonly drain to lymphatic channels in both sides of the neck.

Oropharynx

The oropharynx is the midportion of the pharynx connecting the nasopharynx above with the hypopharynx below. The oropharynx extends from the plane of the inferior surface of the hard palate to the plane of the superior surface of the hyoid bone and opens anteriorly into the oral cavity. The oropharynx includes the base of the tongue, soft palate, tonsillar regions, and lateral and posterior pharyngeal walls. The anatomy of the oropharynx and divisions of the pharynx are illustrated in Figure 61–4.

Lymphatics The primary routes of lymphatic spread from primary tumors of the oropharynx, hypopharynx, and larynx involve deep jugular chain lymph nodes at levels II, III, and IV.[2,3] Midline structures such as the base of the tongue, soft palate, and posterior pharyngeal wall commonly drain to lymphatic channels in both sides of the neck. The soft palate, tonsillar region, and lateral and posterior pharyngeal walls drain to retropharyngeal lymph nodes.

EPIDEMIOLOGY

In the United States, the incidence of oral cavity and pharyngeal cancer is estimated to number approximately 30,200 new cases in 2000, representing 2.5% of all new cases of cancer.[4] Approximately 7,800 deaths from oral cavity and pharyngeal cancer are anticipated in 2000, accounting for 1.4% of the cancer mortality in the United States.

Patients with cancer of the oral cavity and pharynx in the United States remain predominantly male. The estimated incidence in men compared with women is approximately 2 to 1. This gender disparity, however, continues to narrow as tobacco use among women becomes relatively more prevalent.

The relative proportion of cancer of the oral cavity and oropharynx worldwide is significantly higher than that in the United States, representing greater than 30% of cancers in parts of India and Southeast Asia. This higher incidence worldwide is likely related to geographic differences in tobacco consumption, as described in the following section.

ETIOLOGY

The majority of patients with cancer of the oral cavity and oropharynx have a significant history of tobacco and alcohol consumption. Both tobacco and alcohol abuse independently contribute to the development of cancer of the oral cavity and oropharynx.[5,6] The cancer risk has been shown to be significantly elevated both for smokers who did not drink alcohol and for nonsmokers who abused alcohol.[6] When present together, however, the combined effects of tobacco and alcohol abuse in the development of oral and oropharyngeal cancer are multiplicative rather than simply additive.[7,8] These multiplicative effects of smoking and alcohol abuse are shown in Figure 61–5. The synergistic effect of

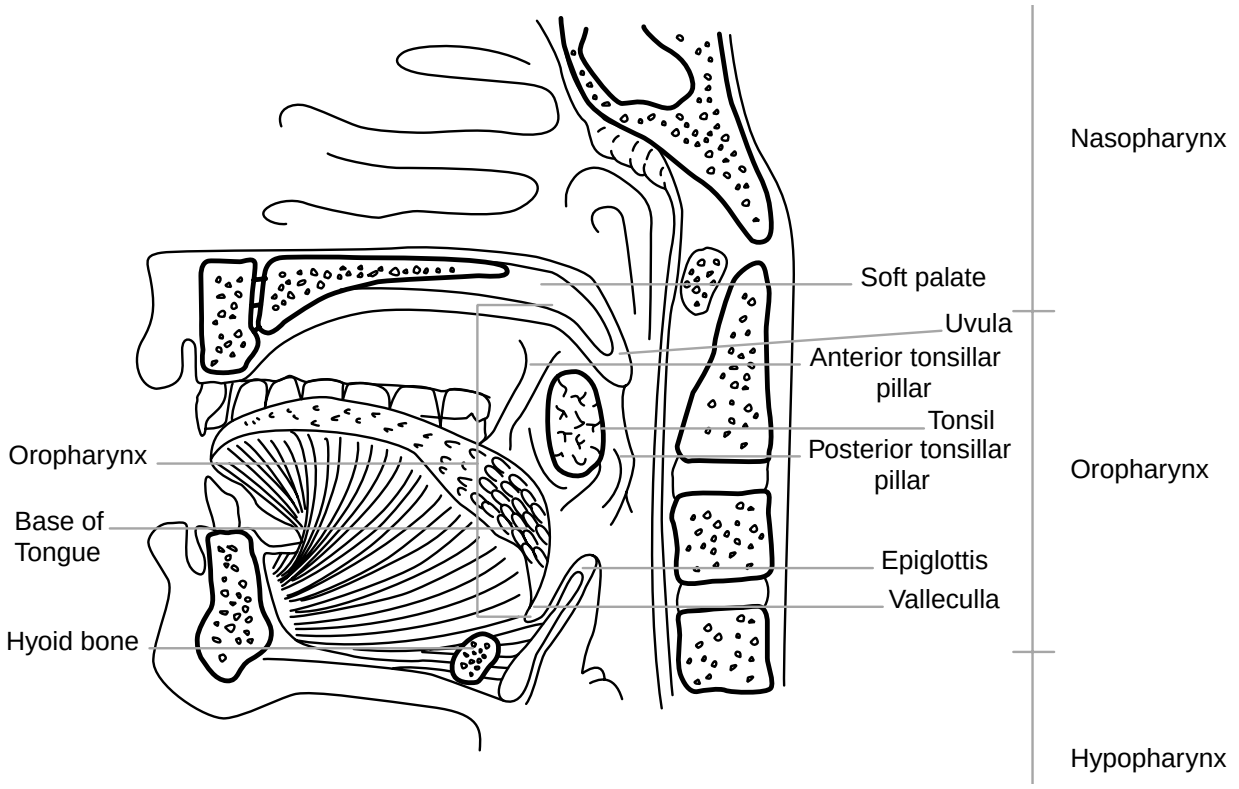


FIGURE 61–4. Anatomy of the oropharynx and divisions of the pharynx.

tobacco and alcohol is shown in the elevation of cancer risk more than 35-fold for men who smoke two or more packs of cigarettes and consume more than four alcoholic drinks per day.[9]

As previously mentioned, the geographic differences in tobacco consumption may explain the higher proportion of cancers arising from the oral cavity and oropharynx worldwide. In Southeast Asia, cultural practices such as reverse smoking, in which the lighted end of the cigarette is held within the mouth, have been shown to produce dysplastic changes at the hard palate.[10] Similarly, betel, a compound chewed regularly throughout Southeast Asia and the western Pacific basin, has been implicated in oral carcinogenesis. Composed of the nut of the areca palm (*Areca catechu*), the leaf of the betel pepper (*Piper betle*), and lime (calcium hydroxide), betel is chewed for its mild psychoactive and cholinergic effects.[11] Other practices common in Southeast Asia include bidi smoking, composed of tobacco rolled within a betel leaf, and consumption of paan, a quid of *Piper betle* leaf containing areca nut, lime, condiment, sweeteners, and sometimes tobacco.[12]

In nearly any form, tobacco consumption has been associated with oral cancer. For example, chronic snuff dipping, in which smokeless tobacco is held at the gingivobuccal sulcus, has been associated with a nearly 50-fold increase in cancers of the gum and buccal mucosa.[13] Pipe smoking and cigar smoking are other types of tobacco consumption linked to cancer of the oral cavity and oropharynx.

The mechanisms through which tobacco contributes to the development of carcinogenesis likely involve deoxyribonucleic acid (DNA) damage from tobacco-related nitrosamines and polycyclic aromatic hydrocarbons. The deleterious effects of these initiators and promoters of carcinogenesis may be partially reversible as cessation of smoking is associated with a sharply reduced risk of cancer, particularly for those who have quit for periods greater than 10 years.[9]

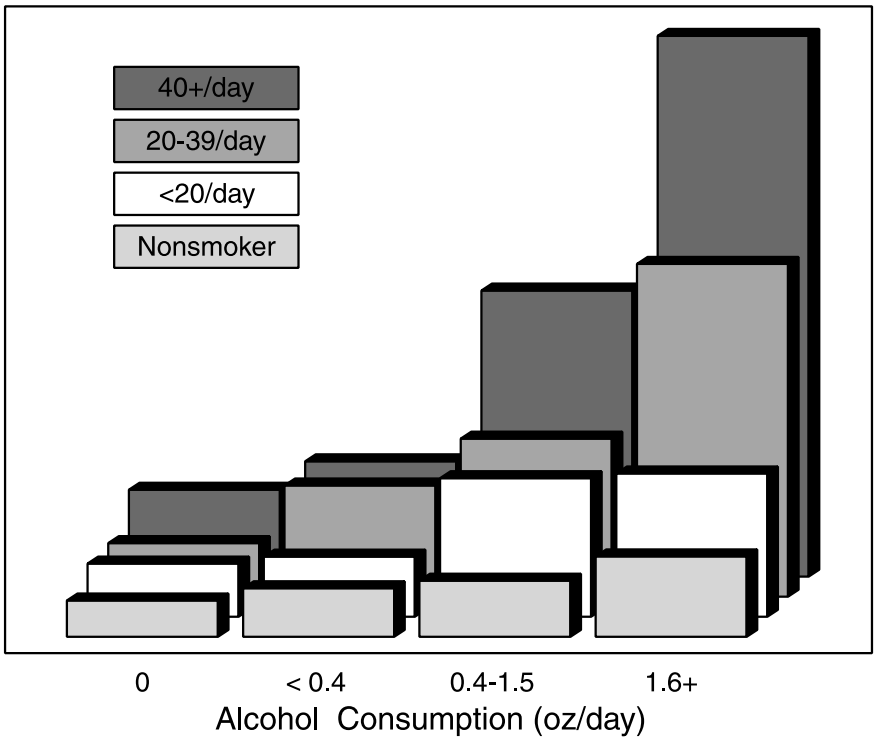
The mechanisms through which alcohol induces cancer development are less well characterized. The increased risk of cancer of the oral cavity and pharynx associated with frequent use of mouthwash high in alcohol content suggests that the etiology involves topical exposure.[14] The specific agent responsible has yet to be identified as experimental studies with laboratory animals have not demonstrated a causal effect between alcohol and carcino-

genesis.[15] Possible mechanisms proposed include alcohol possibly enhancing the metabolism of other carcinogens or enhancing nutritional deficiencies, thus promoting neoplastic change in oral and pharyngeal mucosa. The synergistic effects of combined tobacco and alcohol abuse multiplying the risk of cancer underscore the need for programs to encourage abstinence and cessation from these habits.

Viral infection has been correlated with the development of head and neck cancer. Human papillomavirus (HPV) is a well-established cause for benign recurrent oral and laryngeal papillomas. Genomic DNA sequences of HPV type 16 have been detected in SCCs of the tongue, tonsil, and pharynx.[16] The presence of HPV in the oral cavity has been associated with an increased risk of cancer of the oral cavity or oropharynx (odds ratio = 3.70; 95% confidence interval = 1.47 to 9.32), independent of exposure to tobacco and alcohol.[17] Similar to HPV, infection with human immunodeficiency virus (HIV) appears to confer increased risk for neoplasia. Kaposi's sarcoma is a common manifestation of acquired immune deficiency syndrome (AIDS). Squamous cell carcinoma in patients infected with HIV appears to assume an accelerated course. Squamous cell carcinoma of the upper aerodigestive tract in patients infected with HIV who also exhibited tobacco and/or alcohol abuse occurred at an earlier

Relative risk* of oral cancer according to level of exposure to smoking and alcohol

		Smoking (cigarette equivalents/day)			
		0	< 20	20-39	40+
Alcohol (oz/day)	0	1.00	1.52	1.43	2.43
	< 0.4	1.40	1.67	3.18	3.25
	0.4-1.5	1.60	4.36	4.46	8.21
	1.6+	2.33	4.13	9.59	15.5



*Risks are expressed relative to a risk of 1.00 for persons who neither smoked nor drank.

FIGURE 61–5. Multiplicative effects of smoking and alcohol abuse on the development of oral cancer. Table modified with permission from Rothman K and Keller A.[7] Graph modified with permission from Larson DL. Management of the mandible. In: Close LG, Larson DL, Shah JP, editors. Essentials of head and neck oncology. New York: Thieme Medical Publishers; 1998.

age and resulted in a significantly poorer disease-specific survival.[18]

Poor socioeconomic status, neglected oral hygiene, recurrent trauma from ill-fitting dentures, vocal abuse, and gastroesophageal reflux are other factors associated with an increased risk of carcinoma of the oral cavity and oropharynx. Outdoor occupations requiring prolonged exposure to ultraviolet B radiation in sunlight pose a greater risk of carcinoma of the lower lip.[19] Ionizing radiation has been implicated in the development of minor salivary gland malignancy of the oral cavity and pharynx. Certain dietary habits have been linked to an increased risk for oral cancer, including dietary deficiencies of vitamin A or riboflavin. Achlorhydric iron deficiency anemia associated with Plummer-Vinson syndrome has been shown to be a risk factor for cancer of the oropharynx in Scandinavian women.

PATHOLOGY

Slaughter et al proposed the concept of "field cancerization," a theory predicated on the notion that carcinogen-induced cytologic changes throughout the mucosal lining of the upper aerodigestive tract result in the persistent risk for multiple primary cancers.[20] This theory is supported by the rate of second primary tumors in patients with oral and pharyngeal cancer, approximately 4% annually.[21] The risk of second primary tumors for SCC of the head and neck is approximately 4% annually, up to 25% at 10 years. Both tobacco and alcohol consumption contribute independently to the risk of second primary tumors, with the effects of smoking more pronounced than those of alcohol.[22] The risk of developing subsequent primary tumors in patients who cease smoking after control of the first cancer is one-sixth the risk of those who continue to smoke.[23]

Molecular Events

There have been significant advances recently in our understanding of the cascade of molecular events responsible for carcinogenesis. The development of malignant tumors appears to be the result of multiple accumulated genetic alterations.[24] These genetic mutations must occur in the correct number and sequence for cancer to develop, a progression of genetic events first described in colorectal tumorigenesis.[25] A genetic progression model for head and neck cancer has been established demonstrating that field cancerization seems to involve the expansion and migration of clonally related preneoplastic cells.[26] Microsatellite-based genetic analysis of chromosomal alterations demonstrates that recurrent premalignant lesions arise from a common clonal progenitor with subsequent outgrowth of clonally divergent populations.[27] Phenotypically benign mucosa in the upper aerodigestive tract may harbor foci of clonal, preneoplastic cells that are genetically related to sites of cancer.[28] This observation may explain the incidence for local recurrence following surgical resection of cancers in the oral cavity and oropharynx with margins seemingly adequate by gross and microscopic examination.

Attempts to characterize the genetic events that occur during the progression to a malignant phenotype have focused much attention on determining potential candidates to serve as predictors of behavior or targets for therapy. Genetic alterations in the progression to carcinogenesis include activation of proto-oncogenes and the inactivation of tumor suppressor genes. *P53* is a tumor suppressor gene that plays an important role in arresting cell growth in the presence of genetic damage to permit DNA repair or lead to apoptosis. Mutations in and subsequent inactivation of the *P53* tumor suppressor gene may result in accumulation of DNA damage and uncontrolled cellular growth. It has been shown that the incidence of *P53* mutations increases throughout the progression from premalignant lesions to invasive carcinomas.[29] A history of tobacco and alcohol use is associated with a high frequency of *P53* mutations in patients with SCC of the head and neck, providing an important link between these etiologic factors and the molecular progression to carcinogenesis.[30] Consistent with the aforementioned clonal, preneoplastic cells described in seemingly benign oral and pharyngeal mucosa, *P53* may play a role in the concept of field cancerization. Previous reports have examined pathologic specimens for *P53* mutations in patients undergoing surgery for SCC of the head and neck.[31] Among tumor resection margins and regional lymph nodes initially judged to be negative for cancer on light microscopy, a substantial percentage was positive for *P53* mutations specific for the primary tumor of the head and neck. Patients with these margins positive for *P53* mutations appeared to demonstrate increased risk of local recurrence. In addition to an

association with local recurrence following surgery, mutations in *P53* have also been implicated in locoregional failure in patients with SCC treated with radiation therapy.[32] The precise relationship of *P53* and other specific genetic targets involved in malignant transformation has yet to be fully delineated; however, their study continues to further our understanding of the complex cascade of molecular events that result in carcinogenesis.

Precancerous Lesions

Akin to the progression of genetic events leading to phenotypic evidence of malignancy, various precancerous lesions affect the oral cavity and oropharynx with the potential for malignant degeneration. Leukoplakia is a clinical descriptive term for a white patch in the oral cavity or pharynx that does not rub off. The most common histologic change is epithelial hyperplasia in approximately 80% of lesions, but the prevalence of premalignant or malignant transformation is variable. A review of over 3,000 specimens clinically diagnosed as leukoplakia over a 13-year period revealed mild to moderate epithelial dysplasia in 12.2%, severe epithelial dysplasia or carcinoma in situ in 4.5%, and infiltrating SCC in 3.1%.[33]

Oral lichen planus is a white, lacy, striated lesion of unknown etiology most commonly occurring on the buccal mucosa. Malignant change in oral lichen planus is rare and likely prompted by carcinogenic cofactors.[34]

Erythroplasia appears as a red, slightly raised, granular lesion in the oral cavity and oropharynx. In contrast to the variable incidence of cancer in patients with leukoplakia, erythroplasia has a much higher correlation with concurrent or subsequent malignancy.[35] Minimal size of erythroplastic lesions does not preclude invasiveness, mandating biopsy at times of clinical suspicion.[36]

Oral submucous fibrosis is a precancerous lesion predominantly seen among people from India and Pakistan. It is a chronic disease of the oral mucosa characterized by inflammation and progressive fibrosis of the lamina propria and deeper connective tissues.[37] Bands of oral submucous fibrosis are commonly found in the posterior part of the oral cavity in mild cases and, as the disease increases in severity, are found anteriorly as well.[38] The malignant transformation frequently associated with oral submucous fibrosis is likely multifactorial in origin.[37] After adjusting for other risk factors in a study in Pakistan, patients with oral submucous fibrosis were 19 times more likely to develop oral cancer than those without it.[12]

Histopathology

Squamous cell carcinoma constitutes more than 90% of the malignant neoplasms of the oral cavity and oropharynx. The gross morphologic growth patterns of SCC include exophytic, ulcerative, and infiltrative. The ulcerative type is the most common form of SCC in the oral cavity and demonstrates proclivity for rapid invasion.

Histologically, SCC may be classified into the following categories: keratinizing, nonkeratinizing, spindle cell, adenoid squamous, and verrucous carcinoma. Keratinizing SCCs generally arise from ectodermally derived tissue, tend to be ulcerative or fungating, have less tendency for submucosal spread, and have infiltrating margins. Nonkeratinizing carcinomas typically arise from endodermally derived mucosa, spread submucosally, and have "pushing" margins. Spindle cell carcinoma is a rare variant of SCC demonstrating spindle-shaped mesenchymal cells resembling highly anaplastic sarcoma admixed with components of epidermoid carcinoma.[39] Adenoid squamous carcinoma is a rare variant of SCC in which adenoid transformation appears to arise within preexisting SCC.[40]

Described by Ackerman, verrucous carcinoma presents as a slowly growing exophytic or warty neoplasm in the oral cavity.[41] A well-differentiated variant of SCC, verrucous carcinoma has the histologic appearance of keratinized epithelium arranged in long, papillomatous folds with a "pushing" border. Verrucous carcinoma typically affects the buccal mucosa of elderly patients with a history of tobacco exposure or poor oral hygiene.[42] True verrucous carcinoma does not have metastatic potential.[43] The recommended treatment for verrucous carcinoma is wide surgical excision, although irradiation may be considered in selected patients.[44] Anaplastic transformation of verrucous carcinoma has been reported to occur following treatment with radiation therapy,[45] but this theory is controversial. Foci of less differentiated SCC have been observed to coexist within verrucous lesions, and recurrent lesions tend to be in the form of less differentiated

carcinomas.[44] There is a high rate of recurrence following single-modality treatment with either surgery or irradiation, so combined therapy may be considered for extensive lesions.[46] The histologic change in recurrent lesions to a less differentiated form may explain prior reports of malignant transformation following radiotherapy.

MECHANISMS OF CANCER SPREAD

Local Spread

Primary tumors of the oral cavity and oropharynx spread along tissue planes of least resistance, initially extending superficially along surface mucosa. As lesions enlarge, deeper structures may become involved, including the submucosa, underlying muscle, cartilage, or bone. Squamous cell carcinoma of the head and neck typically extends from its epicenter by irregular fingerlike projections, frequently with microscopic tumor as far as 1 cm beyond tumor margins apparently clear by inspection or palpation.[47] This insidious microscopic course of SCC may provide a mechanism for local recurrence in the presence of apparently negative surgical margins.

As previously mentioned, the risk of developing second primary cancers is approximately 4% per year, up to 25% at 10 years, mandating the need for regular, comprehensive follow-up. Both tobacco and alcohol have each been shown to contribute independently to the formation of multiple primaries. The concept of field cancerization demonstrates the potential of multifocal carcinogenesis through the accumulation of genetic alterations from chronic exposure to mutagenic agents. Multiple primary malignancies are described as simultaneous if diagnosed at the same time as the primary tumor, synchronous if diagnosed within 6 months of the primary tumor, or metachronous if diagnosed later than 6 months after the primary tumor.

Primary tumors may track along the course of nerves and spread outside the oral cavity and oropharynx. The negative prognostic implications of perineural spread are multifactorial. Tumor may spread to the base of the skull, making it less amenable to surgical resection and intensive radiotherapy. Tumors with perineural extension raise the question as to what constitutes an adequate margin of excision in the presence of possible discontinuous spread along the course of the nerve. Microvascular invasion has been demonstrated to have a negative impact on survival for oral and pharyngeal cancers, correlating with higher rates of locoregional recurrence and distant metastases.[48]

Regional Spread

Metastasis to regional lymphatics by carcinoma of the oral cavity and oropharynx is related to tumor size and location, depth of invasion, and the extent of lymphatic drainage from the primary site. Previous surgery, radiation, or inflammation may result in aberrant lymphatic drainage secondary to fibrosis of lymphatic channels. The metastatic potential of primary tumors of the lip, alveolar ridge, and hard palate is less than that of tumors arising elsewhere in the oral cavity. In general, lesions in the posterior part of the oral cavity have a higher predilection of regional lymph node metastases than those lesions situated more anteriorly in the oral cavity.[49] There is an increased risk for bilateral or contralateral metastases from primary tumors arising from midline structures, such as the midline lip, floor of the mouth, oral tongue, base of the tongue, soft palate, and posterior pharyngeal wall. The 5-year survival rate of patients with cervical metastases is approximately 50% that of patients without regional metastases.[50]

Extracapsular spread of carcinoma in cervical lymph nodes portends a poor prognosis. Extracapsular extension has been associated with increased rates of regional recurrence[51] as well as significantly decreased survival rates.[52] Other negative prognostic factors with regard to regional metastases include an increased number of involved nodes, as well as spread of tumor to lymph node levels more inferiorly in the neck.

Distant Spread

Distant metastases from cancers of the oral cavity and oropharynx generally do not occur until advanced stages of disease. Distant metastases typically involve first the lungs or bones.[53] Surveillance chest radiographs, therefore, are essential as part of the routine follow-up for a patient treated for head and neck SCC.

Previous studies have demonstrated that although up to 20% of patients who succumb to head and neck cancer develop metastases to distant

sites, only 4% had distant spread without antecedent failure in the head and neck.[54] Treatment efforts have therefore traditionally concentrated on locoregional control of disease above the clavicles.

CANCER OF THE ORAL CAVITY AND OROPHARYNX IN YOUNG ADULTS

Squamous cell carcinoma of the oral cavity and oropharynx tends to affect patients around the sixth decade of life. A topic of much debate concerns the biologic behavior of cancer of the oral cavity and oropharynx in the younger adult population. Although relatively infrequent, SCC of the oral cavity is observed in increasing numbers in patients younger than 40 years in age.[55] It has been previously reported that SCC of the tonsil arising in patients younger than 40 years resulted in poorer 5-year survival rates.[56] Molecular differences in tumor biology have not been demonstrated in the younger population, and the theory of a more aggressive behavior in the younger population is controversial. Three-year disease-free survival in patients with SCC of the tongue who were less than 40 years in age was not statistically different from that of patients older than 40 years with similar extent of disease.[57] It has been further demonstrated that although adults younger than 40 years with SCC of the tongue had higher rates of local recurrence than did older patients, this difference did not translate into poorer survival rates.[58] It should be noted that studies on both sides of this debate used small patient samples in light of the relatively infrequent occurrence of oral or pharyngeal cancer in the younger adult. However, the rising incidence of malignancy in this subset of the population warrants further efforts toward prevention and cessation of tobacco and alcohol abuse.

OTHER MALIGNANT TUMORS OF THE ORAL CAVITY AND OROPHARYNX

Squamous cell carcinoma accounts for greater than 90% of the malignancies of the oral cavity and oropharynx. Other malignant tumors affecting this region include Kaposi's sarcoma, extranodal non-Hodgkin's lymphoma, mucosal melanoma, and minor salivary gland cancer.

Kaposi's sarcoma is a common manifestation of AIDS and clinically presents as nonpruritic violaceous macules of the mucosa of the upper aerodigestive tract. These tumors grow slowly and may be treated with radiation, laser excision, cryotherapy, or intralesional injection of vinblastine.

Extranodal non-Hodgkin's lymphoma may present as a painless mass along the palate or gingiva or as an asymptomatic enlargement of one tonsil.

Mucosal melanoma of the head and neck is rare, representing only 3% of all melanomas.[59] The palate is the most common site of involvement in the oral cavity, with intralesional lymphatic and blood vessel invasion frequently seen. Surgical excision is the recommended method of treatment, although the ultimate prognosis is typically poor owing to distant dissemination.[60]

The most common minor salivary gland tumor of the oral cavity and oropharynx is adenoid cystic carcinoma, although mucoepidermoid carcinoma and adenocarcinoma may also be present in this region. The incidence of minor salivary gland cancer in this region is increased following previous therapeutic irradiation.

EVALUATION

History and Physical Examination

The first step when evaluating a patient with cancer of the oral cavity or oropharynx is a thorough history and comprehensive examination of the head and neck. The patient should be asked about symptoms of dysphagia, odynophagia, dysarthria, globus sensation, difficulty breathing, hemoptysis, otalgia (possibly referred), weight loss, or other constitutional symptoms. One component of a complete history is an adequate social history establishing risk factors for cancer of the head and neck. The patient should be asked about consumption of tobacco and alcohol, occupational exposures (including exposure to sunlight), and previous radiation exposure.

During the comprehensive examination of the head and neck, assess thoroughly the patient's oral cavity and oropharynx. Look for pigmentation changes or ulcerations of the oral mucosa. Assess the state of dentition for gingival disease and missing or loose teeth. Test the patient's maximal interincisor distance on mouth opening to rule out trismus (indicating possible involvement of the pterygoid musculature). A thorough examination includes visual inspection and palpation of all sites in the oral cavity and oropharynx for mucosal lesions or masses. Bimanual palpation is particularly useful to

evaluate the buccal surfaces and the floor of the mouth. Assess voluntary movement of the oral tongue. Do not neglect to palpate the base of the tongue. Indirect mirror laryngoscopy or fiber-optic laryngoscopy may assist inspection of the oropharynx. Test the patient's somatosensory function of the lower divisions of the trigeminal nerve (V2 and V3) for asymmetry. The patient's neck should be palpated for masses at all five levels. Palpable masses should be assessed for size, location, consistency, mobility, and tenderness.

The necessity for a systematic, comprehensive approach to the history and physical examination cannot be overemphasized. Tumor location is one factor influencing the time of presentation for patients complaining of symptoms. It has been shown that the delay in diagnosis for primary tumors of the oral cavity is greatest for floor of the mouth cancer and shortest for tongue cancer.[61] A number of other factors are involved in patient presentation, including tumor size; bleeding; pain; nodal metastases; interference with speech, breathing, or eating; and tendency to seek treatment. In light of the variable stage of disease at time of presentation, a complete examination of the head and neck is critical. The concept of field cancerization lends credence to the adage that finding one lesion in the oral cavity or pharynx does not preclude the presence of others.

Imaging Evaluation for Regional Metastases

When evaluating for nodal metastases, errors may occur in the clinical assessment of regional lymph nodes by palpation alone. It may be difficult to appreciate lymphadenopathy in patients with a short, stout neck, fibrotic changes secondary to previous radiotherapy, and metastatic lymph nodes at sites typically inaccessible by routine physical examination, such as the parapharyngeal space.[62] There is no substitute for a systematic, comprehensive examination of the neck, but imaging techniques provide valuable supplemental information regarding the status of regional lymph nodes. The sensitivity of CT and magnetic resonance imaging (MRI) for accurately detecting nodal metastases has been reported as 84% and 92%, respectively.[63] Although the presence of metastatic involvement of a lymph node is a pathologic and not a radiologic diagnosis, there are characteristic changes apparent on CT and MRI that suggest the presence of carcinoma, including rim enhancement, central necrosis, and a nodal size in excess of 1 cm in diameter. Positron emission tomography (PET) has also been shown to be useful in detecting regional metastases, with a sensitivity reported as 82%.[64] The utility of fluorine-18-fluorodeoxyglucose (FDG) PET in detecting primary and metastatic SCC of the oral cavity and oropharynx is demonstrated in Figure 61–6. Both physical examination and imaging are techniques prone to false-positive and false-negative errors in assessing for metastases to regional lymph nodes.[65] Pathologic confirmation by fine-needle aspiration (FNA) biopsy is therefore critical for any suspicious neck mass. Fine-needle aspiration is a highly sensitive procedure with minimal morbidity that allows cytologic examination of the aspirate. Repeat FNA should be considered following an initial nondiagnostic FNA. Open neck biopsy is a procedure with extremely limited indications and should be considered only when both the surgeon and the patient are prepared to proceed immediately with a neck dissection if indicated.[62]

Imaging Evaluation of the Mandible

Computed tomography and MRI provide useful adjuvant information regarding soft tissue and bony invasion of advanced oral and oropharyngeal tumors, particularly with extension toward the nasal cavity, maxillary sinus, parapharyngeal space, and larynx. An area of controversy, however, has surrounded the optimal evaluation of mandibular involvement by cancer.

Patterns of spread of SCC within the mandible include spread along the inferior alveolar nerve as well as direct extension through the cancellous bony trabeculae.[66] Clinical evaluation by physical examination has been studied in comparison to panoramic radiographs (panorex) and to CT scans for assessing mandibular invasion by tumor.[67] Clinical evaluation was found to be the most accurate modality to determine the extent of bony invasion and to decide the optimal method of mandibular resection. Panoramic films were found to be helpful in evaluating gross extent of mandibular invasion but not helpful for assessing minimal bony invasion or cortical invasion. The utility of CT scans was reported to be limited owing to the presence of irregular dental sockets and artifacts. Computed

tomographic scans demonstrating mandibular invasion from SCC of the oral cavity are shown in Figure 61–7. DentaScan software provides imaging of the mandible and maxilla in three planes of reference: axial, panoramic, and oblique sagittal. DentaScan imaging minimizes dental restoration artifact, and the multiple planes of visualization improve evaluation of cortical bony margins.[68,69]

ENDOSCOPY

Endoscopic examination of the upper aerodigestive tract under anesthesia provides both thorough inspection of the primary tumor and evaluation for second primary tumors, with the ability to biopsy suspicious sites. The oropharynx, hypopharynx, larynx, and esophagus should be examined in a systematic fashion. The routine use of bronchoscopy is controversial owing to its low yield of synchronous carcinoma of the lung, in the presence of a normal chest radiograph and in the absence of clinical symptoms.[70]

STAGING

Staging systems for cancer provide a standardized measurement for assessing the extent of disease. Staging allows estimation of prognosis, assistance in

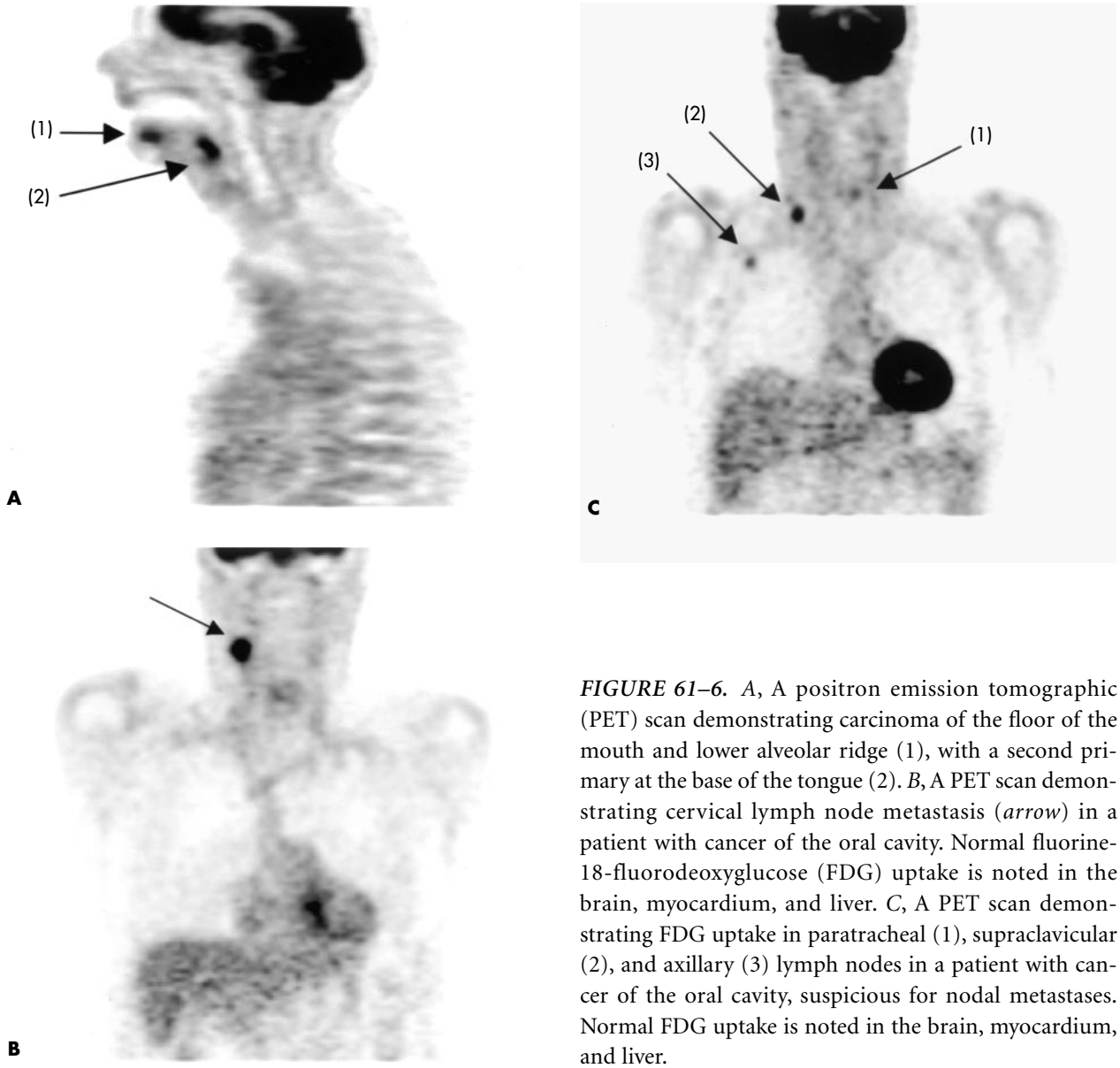


FIGURE 61–6. A, A positron emission tomographic (PET) scan demonstrating carcinoma of the floor of the mouth and lower alveolar ridge (1), with a second primary at the base of the tongue (2). *B*, A PET scan demonstrating cervical lymph node metastasis (*arrow*) in a patient with cancer of the oral cavity. Normal fluorine-18-fluorodeoxyglucose (FDG) uptake is noted in the brain, myocardium, and liver. *C*, A PET scan demonstrating FDG uptake in paratracheal (1), supraclavicular (2), and axillary (3) lymph nodes in a patient with cancer of the oral cavity, suspicious for nodal metastases. Normal FDG uptake is noted in the brain, myocardium, and liver.

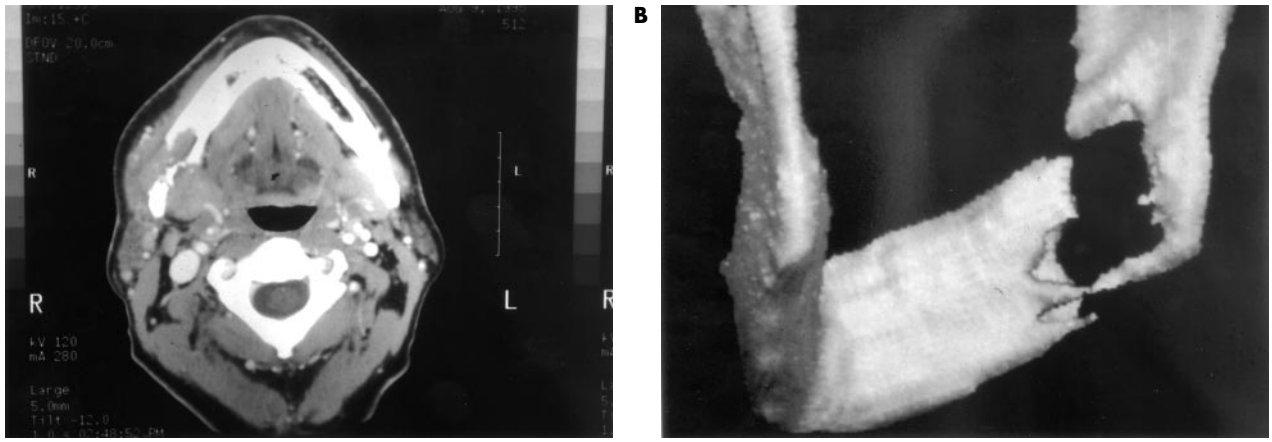


FIGURE 61–7. *A*, Axial computed tomographic (CT) scan of the neck demonstrating mandibular invasion from oral cavity carcinoma. *B*, Three-dimensional reformatted CT scan of the mandible demonstrating bony invasion in the same patient.

planning treatment options, evaluation of the effectiveness of treatment modalities, and promotion of a uniform comparison of treatment results. The staging system of the American Joint Committee on Cancer is based on the TNM staging system in which T category refers to the primary tumor, N category refers to regional cervical lymph node metastases, and M category refers to distant metastases (Tables 61–1 and 61–2).[1] Patient factors such as comorbidity[71,72] and molecular characteristics of tumors are not incorporated in the current staging system, but such variables may provide useful information regarding prognosis and outcomes analysis.

TREATMENT

The primary objective in treating patients with SCC of the oral cavity or oropharynx should focus on rendering the patient free of disease. In general, early-stage lesions may be treated by surgery or radiation with comparable results, and more advanced cancers are best approached with combined therapy. The proclivity of these tumors, particularly late-stage disease, for locoregional failure, however, mandates recognizing the specific prognostic factors associated with each case. Tailoring the most appropriate therapy involves considering patient characteristics, including comorbidity and motivation; physician resources, including experience and facilities; and tumor presentation, including location, grade, and extent of disease. Further, important principles that must be reconciled with tumor eradication include optimizing esthetic form and preserving proper function, including respiration, swallowing, and speech.

SURGERY

In light of the comparable results for early-stage lesions treated with either surgery or radiation therapy, factors in favor of surgical resection include the pathologic staging of local invasion and avoiding complications associated with radiotherapy, including xerostomia and osteoradionecrosis. The primary limitation of surgery involves functional impairment, particularly when resection involving the mandible or tongue is necessary.

Surgical Approaches Different surgical approaches to the oral cavity and oropharynx exist, depending on tumor location and extent. Peroral excision is a suitable approach for small lesions of the oral cavity, including those arising from the floor of the mouth or oral tongue. A cheek flap, either upper or lower, may be constructed for larger lesions in the oral cavity. A visor flap may be used for anterior lesions in the oral cavity to avoid external facial incisions but at the expense of bilateral mental nerve sacrifice.

Tumors arising at or extending to the posterior part of the oral cavity or oropharynx may require mandibulotomy for adequate exposure. A paramedian mandibulotomy approach with paralingual extension (mandibular swing), illustrated in Figure

TABLE 61–1. American Joint Committee on Cancer Staging of Lip and Oral Cavity Cancer

Primary tumor (T)			
TX	Primary tumor cannot be assessed		
T0	No evidence of primary tumor		
Tis	Carcinoma in situ		
T1	Tumor 2 cm or less in greatest dimension		
T2	Tumor more than 2 cm but not more than 4 cm in greatest dimension		
T3	Tumor more than 4 cm in greatest dimension		
T4 (lip)	Tumor invades adjacent structures (eg, through cortical bone, inferior alveolar nerve, floor of mouth, skin of face)		
T4 (oral cavity)	Tumor invades adjacent structures (eg, through cortical bone, into deep [extrinsic] muscle of tongue, maxillary sinus, skin; superficial erosion alone of bone/tooth socket by gingival primary is not sufficient to classify as T4)		
Regional lymph nodes (N)			
NX	Regional lymph nodes cannot be assessed		
N0	No regional lymph node metastases		
N1	Metastasis in a single ipsilateral lymph node, 3 cm or less in greatest dimension		
N2	Metastasis in a single ipsilateral lymph node, more than 3 cm but not more than 6 cm in greatest dimension; or in multiple ipsilateral lymph nodes, none more than 6 cm in greatest dimension; or in bilateral or contralateral lymph nodes, none more than 6 cm in greatest dimension		
N2a	Metastasis in single ipsilateral lymph node more than 3 cm but not more than 6 cm in greatest dimension		
N2b	Metastasis in multiple ipsilateral lymph nodes, none more than 6 cm in greatest dimension		
N2c	Metastasis in bilateral or contralateral lymph nodes, none more than 6 cm in greatest dimension		
N3	Metastasis in a lymph node more than 6 cm in greatest dimension		
Distant metastases (M)			
MX	Distant metastasis cannot be assessed		
M0	No distant metastasis		
M1	Distant metastasis		
Stage grouping			
Stage 0	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1	N1	M0
	T2	N1	M0
	T3	N1	M0
Stage IVA	T4	N0	M0
	T4	N1	M0
	Any T	N2	M0
Stage IVB	Any T	N3	M0
Stage IVC	Any T	Any N	M1

Adapted from American Joint Committee on Cancer.[1]

TABLE 61–2. American Joint Committee on Cancer Staging of Oropharynx Cancer

Primary tumor (T)			
TX	Primary tumor cannot be assessed		
T0	No evidence of primary tumor		
Tis	Carcinoma in situ		
T1	Tumor 2 cm or less in greatest dimension		
T2	Tumor more than 2 cm but not more than 4 cm in greatest dimension		
T3	Tumor more than 4 cm in greatest dimension		
T4	Tumor invades adjacent structures (eg, pterygoid muscle[s], mandible, hard palate, deep muscle of tongue, larynx)		
Regional lymph nodes (N)			
NX	Regional lymph nodes cannot be assessed		
N0	No regional lymph node metastases		
N1	Metastasis in a single ipsilateral lymph node, 3 cm or less in greatest dimension		
N2	Metastasis in a single ipsilateral lymph node, more than 3 cm but not more than 6 cm in greatest dimension; or in multiple ipsilateral lymph nodes, none more than 6 cm in greatest dimension; or in bilateral or contralateral lymph nodes, none more than 6 cm in greatest dimension		
N2a	Metastasis in single ipsilateral lymph node more than 3 cm but not more than 6 cm in greatest dimension		
N2b	Metastasis in multiple ipsilateral lymph nodes, none more than 6 cm in greatest dimension		
N2c	Metastasis in bilateral or contralateral lymph nodes, none more than 6 cm in greatest dimension		
N3	Metastasis in a lymph node more than 6 cm in greatest dimension		
Distant metastases (M)			
MX	Distant metastasis cannot be assessed		
M0	No distant metastasis		
M1	Distant metastasis		
Stage grouping			
Stage 0	Tis	N0	M0
Stage I	T1	N0	M0
Stage II	T2	N0	M0
Stage III	T3	N0	M0
	T1	N1	M0
	T2	N1	M0
	T3	N1	M0
Stage IVA	T4	N0	M0
	T4	N1	M0
	Any T	N2	M0
Stage IVB	Any T	N3	M0
Stage IVC	Any T	Any N	M1

Adapted from American Joint Committee on Cancer.[1]

61–8, offers more versatility and less morbidity than mandibulotomy through a lateral approach.[73] The mandibular swing procedure provides the option for surgical resection with adjuvant radiotherapy as an alternative to single-modality high-dose irradiation to treat oropharyngeal cancer.[74] Placing the mandibulotomy anterior to the inferior alveolar (mental) foramen is ideal to preserve the integrity of the inferior alveolar nerve and to exclude the osteotomy from the portals employed for adjuvant external beam radiotherapy.[75]

One evolution in the surgical approach to cancers of the oral cavity and oropharynx is the attempt to spare the mandible in the absence of bony invasion.[76] Mandibular preservation stems from several principles. Oncologically, it was recognized to be no longer necessary to remove a portion of the mandible simply to keep contents from a neck dissection in continuity with resection of a primary oral tumor.[77] Functionally, resection of the anterior arch of the mandible results in disabilities, including drooling and interference with eating, directly related to the amount of bone removed. Esthetically, this "Andy Gump" deformity results in an inferior cosmetic appearance. Segmental mandibulectomy is illustrated in Figure 61–9.

Instances of bony invasion by tumor mandate resection of that portion of the mandible. The amount of bone to be resected is predicated on the location and extent of bony involvement. Marginal mandibulectomy, shown in Figure 61–10, has been recommended for lower gingival cancers if erosive bony defects did not extend beyond the inferior alveolar (mandibular) canal or if invasive bony defects were confined to a superficial area of alveolar bone.[78] Rim mandibulectomy may be effective, provided that there is preservation of at least 1 cm thickness of bone inferiorly to prevent mandibular fracture.[79]

Surgical Margins An important concept in the surgical management of cancer of the oral cavity and oropharynx concerns the importance for adequate margins of resection. Local recurrence in patients with positive surgical margins following resection of SCC from the oral cavity was twice as high as those with negative margins.[80] Furthermore, local recurrence rates and 5-year survival rates of patients with "close" surgical margins (less than

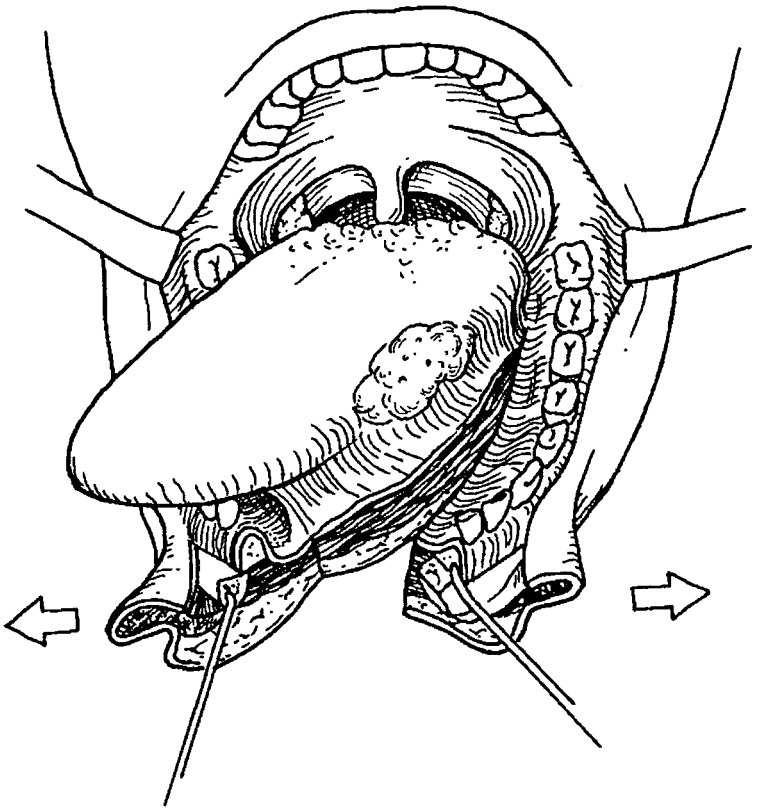


FIGURE 61–8. Paramedian mandibulotomy for resection of carcinoma of the posterior part of the tongue. Reproduced with permission from Larson DL. Management of the mandible. In: Close LG, Larson DL, Shah JP, editors. Essentials of head and neck oncology. New York: Thieme Medical Publishers; 1998. p. 196.

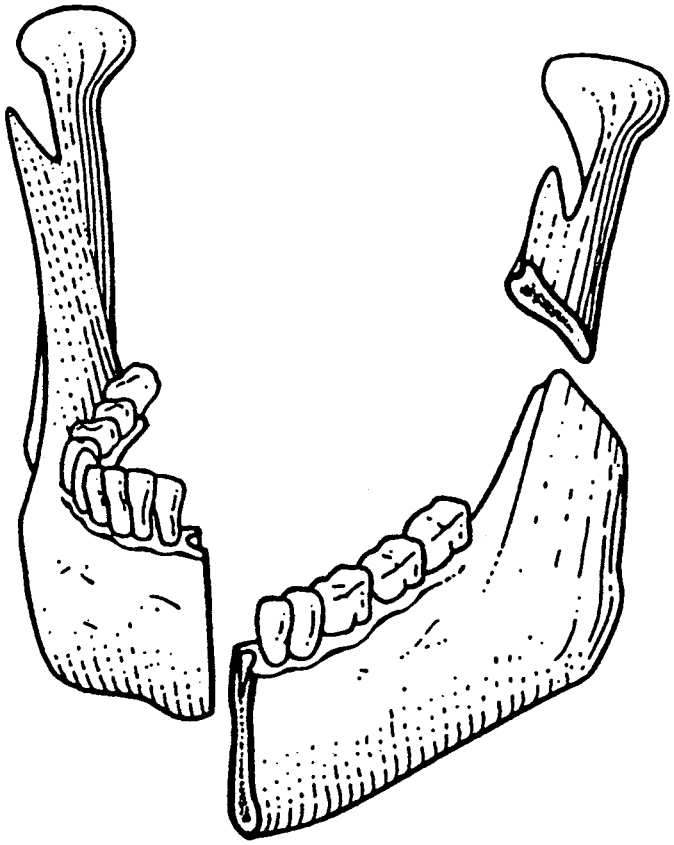


FIGURE 61–9. Segmental mandibulectomy. Reproduced with permission from Alvi A, Myers EN, Johnson JT. Cancer of the oral cavity. In: Myers EN, Suen JY, editors. Cancer of the head and neck. 3rd ed. Philadelphia: WB Saunders; 1996.

5 mm) were comparable to those with microscopically positive margins.[81]

In light of the impact of histologically "close" or positive margins on local recurrence and survival, it is critical at the time of surgery to ensure resection margins of adequate size. Another factor to consider at the time of surgery is the reported change in tissue margins before and after pathologic fixation. In an animal model, oral cavity margins shrank from approximately 30 to 50% from the in situ measurement of the surgeon at the time of resection to final pathologic evaluation on the microscopic slide after formalin fixation.[82]

Radiation Therapy

General principles of radiotherapy summarized by Wang include the generally favorable radioresponsiveness of most early-stage head and neck SCCs; a well-oxygenated environment provides an ideal setting for therapeutic radiation; invasion of bone or deep muscle portends a poorer response to radiotherapy; and neck dissection, with or without adjuvant irradiation, is the preferred treatment for large cervical metastases.[83]

Modalities for primary irradiation include external beam radiation therapy, interstitial implantation, or a combination of the two. Definitive radiotherapy shares comparable results with surgical treatment for early-stage oral or oropharyngeal cancers. Radiation therapy is typically favorable for tonsillar carcinoma, which is particularly responsive to radiation therapy, and is preferred for early-stage lesions in the oral cavity or oropharynx in patients who pose an anesthetic risk.

Radiation dose depends on the extent of disease and the tolerance of the surrounding normal tissues. Definitive radiotherapy may be delivered with doses of 60 to 65 Gy over 6 weeks for smaller tumors. Larger tumors may require 65 to 70 Gy delivered over 7 weeks. Adjuvant radiotherapy following surgery may be delivered with doses of 60 to 65 Gy over 6 weeks for microscopic residual disease. Greater doses may be required for gross residual disease.

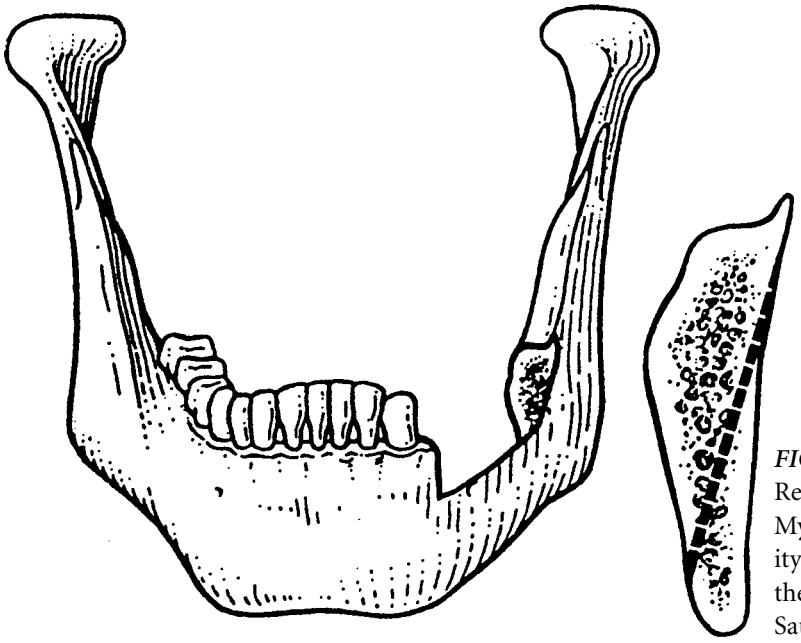


FIGURE 61–10. Marginal mandibulectomy. Reproduced with permission from Alvi A, Myers EN, Johnson JT. Cancer of the oral cavity. In: Myers EN, Suen JY, editors. Cancer of the head and neck. 3rd ed. Philadelphia: WB Saunders; 1996.

Improvement continues in the mode of delivery of radiotherapy. Improved 5-year local control rates and disease-specific survival rates have been reported for T1 to T3 base of tongue and faucial tonsillar lesions using an accelerated hyperfractionated radiotherapy twice-daily program compared with conventional once-daily radiation therapy.[84] A phase III randomized trial 9003 of the Radiation Therapy Oncology Group (RTOG) has reported improved locoregional control of locally advanced head and neck cancer using hyperfractionation or accelerated fractionation with concomitant boost compared with conventional once-daily fractionation.[85] Interstitial brachytherapy has been demonstrated to be effective for carcinoma of the oral cavity and oropharynx.[86]

In addition to side effects of mucositis and xerostomia, major complications of definitive radiotherapy for SCC of the oral cavity and oropharynx have been reported, including osteoradionecrosis, pathologic fracture, or ulceration of mucous membranes.[87] The risk of osteoradionecrosis necessitates comprehensive dental care prior to the initiation of radiotherapy.[88]

Combined Therapy

A combination of surgery and radiation therapy is recommended for advanced cancers of the oral cavity and oropharynx in light of the high propensity for locoregional failure. The timing of irradiation in relation to surgery has been a matter of debate. The advantages of preoperative radiotherapy include reduction of tumor volume in a better-oxygenated environment and possible sterilization of the surgical field. The disadvantages of preoperative radiotherapy include possible surgical wound complications and possible obscuring of tumor margins at the time of resection. The advantages of postoperative (adjuvant) radiotherapy include the safer administration of higher radiation doses, improved patient compliance, and irradiation of potential microscopic disease rather than gross, bulky tumor. The disadvantages of postoperative radiotherapy include the possibility of surgical complications delaying the initiation of radiation treatments. The final report of study 73-03 of the RTOG did not detect statistically significant differences in locoregional control or overall survival between preoperative radiotherapy and postoperative radiotherapy for carcinoma of the oral cavity and oropharynx.[89]

At the present time, combined treatment for advanced carcinoma of the oral cavity typically involves surgical resection and postoperative radiation therapy. Indications for adjuvant radiotherapy include multiple positive nodes, extracapsular lymph node extension, and positive margins of

resection. Adjuvant radiation therapy has been shown to improve locoregional control in the presence of extracapsular spread or positive resection margins following surgery.[90] Postoperative radiotherapy, however, should not be construed as a panacea for an inadequate operation. Adjuvant radiotherapy does not appear to decrease the risk of local recurrence in patients with positive margins to a level similar to that in comparable patients with negative margins not treated with radiotherapy.[80]

The ideal timing of postoperative radiotherapy has been recommended as starting treatments within 6 weeks following surgery. For stage III and IV SCC of the head and neck, locoregional recurrence was 5.5% when adjuvant radiotherapy was initiated within 6 weeks following surgery. Locoregional recurrence was 31.5% when postoperative radiation therapy was delayed longer than 6 weeks.[91] This increase in nodal recurrence rate when adjuvant radiotherapy was delayed longer than 6 weeks was observed whether nodal metastases were confined to a single nodal level or whether multiple levels were involved. Furthermore, this higher recurrence rate with delayed radiotherapy was not abrogated by using higher doses of radiation.[92]

Neck dissection is the preferred modality for salvage. Irradiation for surgical failure in the neck is associated with worse survival rates compared to salvage neck dissection following failed radiotherapy.[93]

Chemotherapy

The effect of using chemotherapy to treat head and neck SCC has been studied in a variety of settings. Chemotherapy administered prior to definitive local therapy is referred to as neoadjuvant (induction) chemotherapy. Adjuvant chemotherapy describes chemotherapy delivered after definitive local therapy. Concomitant chemotherapy refers to chemotherapy administered simultaneously with irradiation.

A multi-institutional prospective randomized trial of the Southwest Oncology Group demonstrated no survival advantage with the addition of neoadjuvant chemotherapy, administered prior to surgery and postoperative radiotherapy, for advanced resectable head and neck cancer.[94] Jacobs et al found no demonstrable difference in 5-year survival rates between neoadjuvant and adjuvant chemotherapy using cisplatin and 5-fluorouracil (5-FU) for advanced, resectable head and neck cancer.[95] Another study using adjuvant cisplatin and 5-FU following surgery with postoperative radiotherapy demonstrated a decreased incidence of distant metastases but no difference in 4-year locoregional control, overall survival, or disease-free survival.[96] In light of the effective treatment of concomitant chemotherapy and irradiation for advanced laryngeal cancer,[96,97] a randomized trial of concomitant carboplatin and 5-FU combined with radiotherapy was performed for stage III and IV oropharyngeal carcinoma.[96,98] These results demonstrated greater 3-year overall actuarial survival and disease-free survival rates than conventional radiotherapy alone. Locoregional control was improved in concomitant chemotherapy and radiation therapy compared with radiation therapy alone. However, mucositis and hematologic toxicity, measured by neutrophil count and hemoglobin level, were higher for the concomitant treatment group.

Further investigation into the benefit of chemotherapy for carcinoma of the oral cavity and oropharynx is warranted as it currently plays a supplemental role to the established treatment modalities, surgery and radiation. Chemotherapy, however, is frequently employed in the palliation of advanced head and neck cancer not amenable to the primary methods of treatment or in the setting of recurrent disease.

Chemoprevention

Recent efforts to expand our armamentarium to treat cancer of the oral cavity and oropharynx have included investigation of chemopreventive agents. As the genetic events leading to carcinogenesis are further elucidated, attempts have been made to identify potential agents that can prevent DNA damage or enhance DNA repair. A prospective randomized trial has demonstrated effective treatment of oral leukoplakia with a 6-month trial of isotretinoin (13-*cis*-retinoic acid), with significant reduction in lesion size and reversal of dysplasia.[99] Relapse, however, occurred in 56% 2 to 3 months following cessation of treatment. A follow-up prospective randomized study used a 12-month trial of isotretinoin following primary treatment for SCC of the oral cavity, pharynx, or larynx.[100] After a median period of 32 months, no significant differences were found in local, regional, or distant recurrences of the

primary cancers. However, patients treated with isotretinoin had significantly fewer second primary tumors. The response to chemopreventive agents may include genetic alterations not readily apparent on phenotypic presentation. This discordance suggests the need to consider not only phenotypic changes, such as lesion size and differentiation, but also genotypic response, using critical genetic markers, when evaluating the effect of chemopreventive agents.[101]

Management of the Neck

An issue of ongoing controversy concerns the management of the neck in patients without clinical evidence of regional nodal metastases (the N0 neck). Lymphatic metastases from primary tumors of the oral cavity and oropharynx occur in a predictable fashion through sequential spread. Although previous surgery or irradiation may result in aberrant lymphatic drainage, there are well-established first-echelon lymph nodes at highest risk for metastases depending on the primary site of cancer. Pathologic evidence of micrometastases has been observed on gross section in up to 33% of elective neck dissection specimens. More sensitive detection methods such as serial sectioning or polymerase chain reaction may result in an even higher prevalence. In light of the propensity for various sites of the oral cavity and oropharynx to spread to regional lymphatics and the high risk of nodal recurrence, what constitutes the ideal approach to the N0 neck? Is the best approach observation or treatment? If treatment is chosen, is elective neck dissection or elective neck irradiation the optimal modality?

Elective Neck Dissection It has been previously reported that there is no survival advantage offered by elective neck dissection when compared to observation.[102] However, when observation is used for the neck at risk for metastasis, patients tend to fail with advanced nodal disease, even with close follow-up.[103] Elective neck dissection has been shown to improve locoregional control and may therefore positively impact on the quality of patient's survival.[104]

One criticism of elective neck dissection has focused on potential complications from the procedure. Initially described by Crile in 1906[105] and popularized by Martin et al,[106] the so-called radical neck dissection or comprehensive neck dissection, with sacrifice of the spinal accessory nerve (XI), SCM, and internal jugular vein, is associated with significant morbidity. Significant compromise of shoulder function results from sacrifice of nerve XI, and cosmetic deformity with platysmal banding occurs from resection of the SCM.[107] Removal of the internal jugular vein risks significant edema of the face and occasional neurologic deficit, particularly following bilateral neck dissection or when radiotherapy is employed.

In light of the morbidity associated with radical neck dissection, there has been a trend toward selective, rather than comprehensive, neck dissection, based on the predictable pattern of cervical lymph node metastases. Selective neck dissection has been demonstrated to be an oncologically sound procedure, providing effective treatment for the N0 neck.[108] Supraomohyoid neck dissection (SOHND), selective lymphadenectomy clearing cervical nodal levels I, II, and III, has been recommended for N0 patients with primary SCCs of the oral cavity.[3,109] Other reports have extended the indication for SOHND for primary tumors in the oral cavity to include not only N0 necks but also N1 necks without evidence of extracapsular spread.[110,111] A proposed limitation of the SOHND for primary tumors of the oral tongue concerns the potential for "skip metastases," or metastases to inferior cervical nodes at level III or IV in the absence of demonstrable involvement at levels I and II.[112] Although clearing level IV during a SOHND adds little in terms of morbidity and operative time, the risk of isolated level IV involvement in the absence of other cervical metastases may be sufficiently low as to obviate the need for routine addition of level IV to the SOHND for the N0 neck with an oral cavity primary tumor.[113]

Based on the predictable spread of cervical nodal metastases, SOHND may not be sufficient for the N0 neck with a primary arising in the oropharynx. The risk for level IV spread is higher from primary tumors of the oropharynx compared with those arising from the oral cavity. Thus, an anterolateral neck dissection encompassing levels II, III, and IV of the deep jugular chain has been advocated for N0 necks with an oropharyngeal primary.[3]

For primary tumors of the oral cavity or oropharynx, the risk of level V involvement has been shown to be minimal when only one level is involved, unless level IV is the involved level or multiple levels are involved.[114] Level V, therefore, is not

routinely included in the selective lymphadenectomy of N0 necks for primaries of the oral cavity or oropharynx.

In this era of more selective neck dissections, the utility of lymphoscintigraphy with nodal sampling, clinically useful for treatment of cutaneous melanoma, has been investigated for primaries of the oral cavity. The fundamental principle behind determining the sentinel lymph node of a primary cancer is based on the well-established patterns of cervical lymph nodes at highest risk for metastasis. Given the potential savings of time, cost, and morbidity, the detection of the sentinel node by lymphoscintigraphy has been shown to be a feasible procedure for selected patients with oral cancer and clinically negative necks. However, the ability to detect metastases reliably has not been established.[115] Previous neck dissection has been demonstrated to result in modified lymphatic drainage in patients with tongue cancer, thereby possibly further complicating accurate detection by lymphoscintigraphy.[116] The applicability of lymphoscintigraphy-assisted sentinel lymph node sampling may be limited in oral cavity cancer, secondary to the technical challenges of accurate detection, as well as the risk of associated morbidity that may not be significantly different from that of a SOHND.

Elective Neck Irradiation The theoretical advantages of treating the N0 neck with surgical lymphadenectomy include pathologic staging information, preparation of donor and recipient vessels for microvascular reconstruction, and avoidance of the morbidity associated with radiotherapy. However, elective neck irradiation has been shown to provide lymph node failure rates of less than 5% when the primary lesion has been adequately controlled.[117] In fact, no significant differences in regional recurrence have been demonstrated following treatment with elective neck irradiation compared to elective neck dissection.[118]

The modality chosen to address the clinically negative neck may be largely influenced by philosophy of the treating physician. Nevertheless, as a general rule, the treatment modality for the primary tumor of the oral cavity and oropharynx may be used concurrently to address occult metastases in the clinically negative neck. Recognizing that some sites of the oral cavity have less propensity to spread to regional lymphatics (hard palate, alveolar ridge, and nonmidline aspects of the lip), the risk of regional failure with primaries of the oral cavity and oropharynx is sufficiently high to warrant elective treatment of the neck in most instances.

Therapeutic Neck Dissection For the clinically positive neck, the traditional surgical procedure of choice has been comprehensive neck dissection with preservation of nerve XI when technically feasible. The continuing evolution of a more selective approach to the neck, however, has included the clinically positive neck as well. Comprehensive neck dissection, with preservation of nerve XI, the SCM, and the internal jugular vein, may be performed for N1, N2a, or N2b disease when technically feasible. Supraomohyoid neck dissection may be acceptable for N1 disease without extracapsular extension arising from primaries of the oral cavity, particularly when the involved node is at level I.

SPECIFIC SITES IN THE ORAL CAVITY

Lip

Squamous cell carcinoma of the lip, by virtue of its location, tends to present at an early stage. The lower lip is affected more commonly, presumably secondary to sunlight exposure, with origin at the commissure occurring in less than 1% of cases.[119]

Comparable cure rates have been reported for small tumors using either surgery or radiation therapy, but surgical excision is the treatment of choice in most instances owing to its lower morbidity and better cosmetic result.[120] Advantages of surgery over irradiation include the ability to assess tumor margins, avoidance of the complications associated with radiotherapy, and rapid rehabilitation.[121] Lesions up to one-third the lip length may be resected with primary closure, and a number of options exist for flap reconstruction for larger lesions.

An important consideration for advanced lesions is possible involvement of the mental branch of the mandibular nerve (V3), which provides sensory innervation to the lower lip. Surgical salvage for local recurrence is successful in greater than 70% of cases.[120]

Metastases from carcinoma of the lip are uncommon except in advanced lesions, recurrent lesions, or lesions arising at the oral commissure. When lymphatic spread arises from midline lip

lesions, bilateral nodal metastases are more prevalent. Depth of invasion is an important prognostic factor for regional lymph node metastasis, particularly for lesions greater than 6 mm thick.[122]

Overall survival rates reported for stages I to IV are approximately 90% when using surgery for stages I and II and combination treatment of surgery plus adjuvant radiotherapy for stages III and IV.[120]

Buccal Mucosa

Cancers of the buccal mucosa are characteristically locally aggressive, spreading to adjacent gingiva in approximately 38% of cases and involving the retromolar trigone or pterygoid region in approximately 13%.[123] The intraoperative photograph in Figure 61–11 depicts an advanced-stage SCC of the right buccal mucosa extending through the external skin.

Surgical resection is recommended for stage I and II tumors, with combination therapy using surgery and postoperative radiotherapy for stage III and IV tumors.

Using this treatment approach, the following 5-year survival rates have been reported: stage I, 77%; stage II, 65%; stage III, 27%; and stage IV, 18%.[123]

Upper and Lower Alveolar Ridges

Tumors of the alveolar ridge typically present with soreness or gum pain, ulceration, intraoral bleeding, loosening of teeth, or ill-fitting dentures. Patients tend to present with symptoms within 3 months duration.[124]

Cancer of the upper alveolar ridge may extend superiorly to involve the floor of the nasal cavity or the maxillary sinus. Carcinoma of the alveolar ridge more commonly arises from the lower alveolus, and cancers of the lower gum may spread along the course of the inferior alveolar nerve toward the skull base.

Based on 2-year survival rates, marginal mandibulectomy has been demonstrated to be a sound oncologic procedure for stage I and II disease of the lower alveolus.[125] Marginal mandibulectomy is an effective procedure for lower alveolar carcinoma but is contraindicated if the mandible is enveloped by tumor, if there is tumor invasion through the cortical plate, or if there has been recent tooth extraction in proximity to the tumor.[124] The use of adjuvant radiotherapy has been advocated for inadequate margins of surgical resection, perineural invasion, or extensive nodal metastases,[126] although regional metastases to cervical lymphatics are uncommon in carcinoma of the alveolar ridge.

Using surgery as the primary treatment modality, with the addition of postoperative radiotherapy for advanced lesions, the following 5-year survival rates have been demonstrated: stage I, 77%; stage II, 70%; stage III, 42%; and stage IV, 24%.[124]

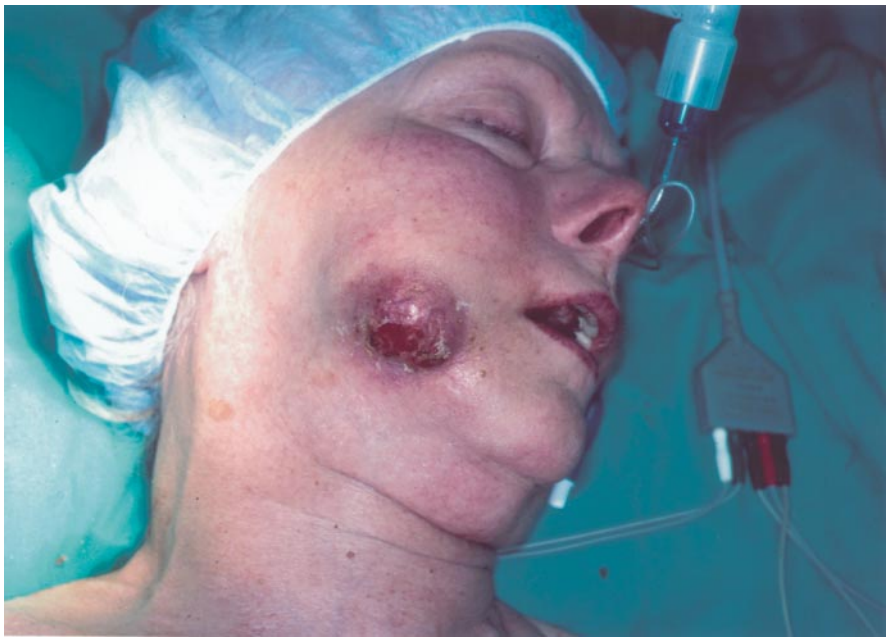


FIGURE 61–11. Advanced-stage squamous cell carcinoma of the right buccal mucosa penetrating through the skin.

Retromolar Trigone

Branches from the glossopharyngeal (IX) nerve provide sensory innervation to the retromolar trigone, so patients with carcinoma of the retromolar trigone may present complaining of pain referred to the ipsilateral ear. Numbness in the distribution of the inferior alveolar nerve may be observed if tumor invades the mandible. Trismus may result from invasion of the pterygoid musculature. Extension into the pterygopalatine fossa may lead to disease at the skull base. The proximity of the structures of the oropharynx, including the base of the tongue, soft palate, and tonsil, places them at risk for tumor involvement by direct extension. Proximity to the ramus of the mandible often necessitates resection of the ascending portion of the mandible.

Hard Palate

Although relatively rare in the United States, carcinoma of the hard palate is more prevalent in places of the world where reverse smoking is practiced. The heat generated from holding the lit end of the cigarette near the hard palate has been demonstrated to result in malignant transformation. Presenting symptoms of cancer of the hard palate include pain, bleeding, improper denture fit, or altered speech. Biopsy may be necessary to distinguish carcinoma of the hard palate from other disease entities, such as necrotizing sialometaplasia. One confounding factor in establishing a diagnosis of hard palate cancer is the considerable delay before patients present complaining of symptoms. One study showed that 58% of patients had had symptoms for at least 12 weeks before seeking treatment, and 38% had a delay of more than 6 months.[127] Advanced tumors may extend into the nasal cavity through the incisive foramen or may approach the skull base via the greater palatine foramina.

For early-stage lesions, no differences have been shown between single-modality treatment using either surgery or irradiation, so selection of therapy should be based on the anatomic location and extent of disease, the presence of second primaries, and associated patient comorbidities.[127] Surgery followed by adjuvant radiotherapy is recommended for advanced-stage disease.

Regional lymphatic drainage from the hard palate is sparse, so metastases to cervical lymph nodes are uncommon.

Floor of the Mouth

By virtue of their location, cancers of the floor of the mouth may remain undetected until they progress to advanced disease. Large bulky tumors may affect normal speech and deglutition. Patients may complain of pain referred to the ipsilateral ear from tumor involvement of the lingual nerve extending to the main trunk of the mandibular nerve (V3). Advanced tumors may invade the tongue or mandible by direct extension. Second primary tumors are particularly common with floor of the mouth cancer.[128]

Elective treatment of the neck is warranted in carcinoma of the floor of the mouth owing to the significant incidence of occult nodal metastases, even with T1 lesions.[129] Patients with high-stage primary cancers are at high risk for advanced nodal metastases as approximately one-third of patients with T4 tumors present with N3 disease.[128] The risk of contralateral nodal metastases is significant with floor of the mouth cancer,[130] and multiple levels are often involved.[128]

The following 5-year survival rates, using surgery as primary therapy with postoperative radiotherapy for stage III and IV disease, have been reported: stage I, 88%; stage II, 80%; stage III, 66%; and stage IV, 32%.[128]

Figure 61–12 illustrates the surgical management of a patient with biopsy-proven SCC of the floor of the mouth invading the ventral surface of the oral tongue. The patient underwent resection of the lesion at the floor of the mouth and ventral surface of the tongue, with marginal mandibulectomy, bilateral SOHND, and tracheostomy. Repair of the surgical defect was accomplished through radial forearm free flap reconstruction.

Oral Tongue

Common presenting symptoms and signs for carcinoma of the oral tongue include localized pain and the presence of an ulcer, frequently at the middle third of the tongue. Patients may report dysarthria or pain with eating. As with the floor of the mouth, cancers of the oral tongue may present with referred pain in the ipsilateral ear owing to involvement of the mandibular nerve (V3).

Second primary tumors are common in cancers of the tongue, occurring in approximately 24% of patients.[131]

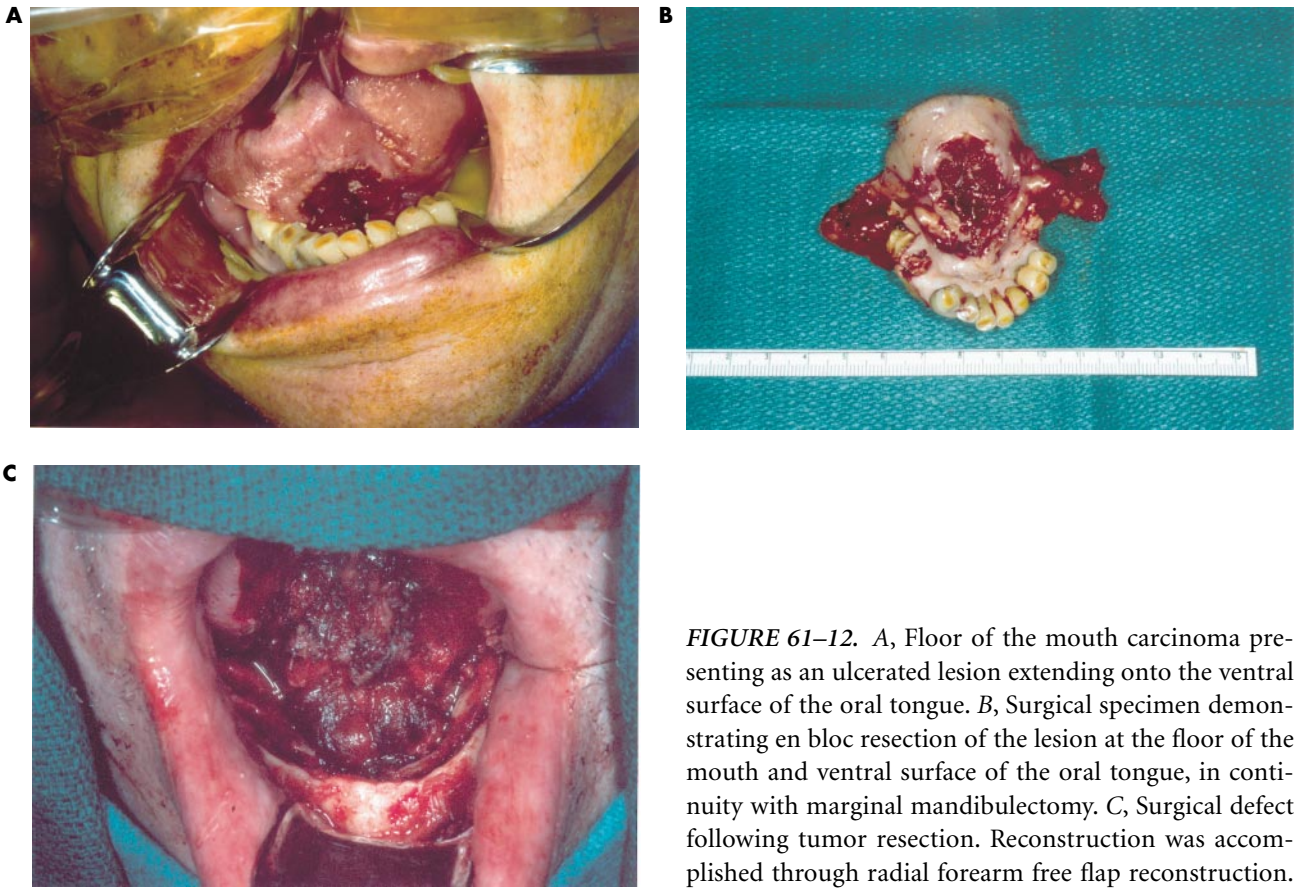


FIGURE 61–12. *A*, Floor of the mouth carcinoma presenting as an ulcerated lesion extending onto the ventral surface of the oral tongue. *B*, Surgical specimen demonstrating en bloc resection of the lesion at the floor of the mouth and ventral surface of the oral tongue, in continuity with marginal mandibulectomy. *C*, Surgical defect following tumor resection. Reconstruction was accomplished through radial forearm free flap reconstruction.

There have been no demonstrable differences in rates of locoregional control or survival between surgery and irradiation for T1 and T2 lesions.[132] Surgical resection is typically employed for stage I and II lesions, whereas combined therapy using surgery and postoperative radiotherapy is indicated for stage III and IV disease.[131]

Occult metastases in regional lymphatics are common with carcinoma of the tongue, particularly when the depth of the primary tumor is greater than 2 mm thick.[133] The risk of regional metastases approaches 40% in this group of patients, necessitating elective treatment of the N0 neck with carcinoma of the oral tongue. The rates of both locoregional control and survival are improved with elective neck dissection, even for T1 and T2 lesions.[134] The N0 neck may be treated with selective neck dissection as no differences in survival or locoregional control have been demonstrated between comprehensive and selective neck dissection for elective lymphadenectomy[131,135]

Improvements in observed rates of overall 5-year survival have been attributed to a more aggressive approach to the neck in patients with early-stage (stages I and II) tumors and the addition of postoperative radiotherapy in patients with advanced-stage (stages III and IV) disease. Reported 5-year survival rates for cancer of the tongue using surgery has primary treatment with adjuvant radiotherapy: stage I, 90%; stage II, 72%; stage III, 54%; and stage IV, 34%.[131] The increasing use of mandible-sparing procedures and selective neck dissections may result in improved quality of survival as well.

RECONSTRUCTION OF THE ORAL CAVITY

Following resection of cancers of the oral cavity, defects may be approached in a variety of ways. Select small intraoral defects may be closed primarily, covered with a local flap or split-thickness skin graft, or left to heal by secondary intention. Larger defects may require more complex flap reconstruction. A number of factors affect the type of reconstruction, if any, used, including tumor extent and location; possible dysfunction of breathing, swal-

lowing, or speech; patient comorbidity; and the surgeon's experience and preference.

A comprehensive description of the variety of reconstructive techniques for soft tissue and bony defects following ablation of oral cavity cancers is beyond the scope of this chapter (see Chapters 40 and 41). Nevertheless, two key developments in the evolution of head and neck reconstruction that deserve special mention include pedicled flaps and microvascular free tissue transfer.

The pectoralis major myocutaneous flap provides well-vascularized soft tissue for reconstructing oral cavity defects in a single stage.[136] It is one example of a pedicled flap, receiving its blood supply from the thoracoacromial artery, and is useful for reconstructing intraoral as well as external defects. Its reliable vascularity, proximity to the head and neck region, and availability for harvest with the patient in a supine position led to its wide acceptance for reconstruction of head and neck defects. The pectoralis major myocutaneous flap, however, does have several disadvantages in intraoral reconstruction, including hair-bearing skin, soft tissue bulk inappropriate for smaller defects, and limitation in the arc of rotation.

Microvascular free tissue transfer revolutionized head and neck reconstruction, providing flaps and their nutrient vessels to repair complex defects with uniformly perfused tissue. The radial forearm free flap and rectus abdominis free flap are two types of free tissue transfer useful for intraoral soft tissue reconstruction.

Mandibular defects may be successfully reconstructed with composite bone flaps, including the fibular osteocutaneous free flap or iliac crest osteocutaneous free flap. Such vascularized composite bone flaps provide reliable bone stock for optimal esthetic contour and masticatory function. Figure 61–13 illustrates the surgical management of a patient with biopsy-proven SCC of the floor of the mouth involving the anterior part of the mandible and anterior part of the tongue. The patient underwent resection of the floor of the mouth, subtotal glossectomy, segmental mandibulectomy, tracheostomy, and bilateral modified radical neck dissection incorporating cervical nodal levels I through IV. Mandibular reconstruction was accomplished using fibular free flap, with soft tissue coverage provided by radial forearm free flap reconstruction.

SPECIFIC SITES IN THE OROPHARYNX

Base of the Tongue

Typical of cancers of the oropharynx, base of the tongue cancers often present at an advanced stage of disease. Overall, 62% of patients present with nodal metastases,[137] and contralateral or bilateral nodal involvement is frequently seen.

A variety of treatment options for base of the tongue cancer have been proposed, including surgery with or without postoperative radiotherapy, primary external beam radiotherapy, external beam radiation therapy with brachytherapy implantation with or without neck dissection, and induction chemotherapy with external beam radiation therapy with or without brachytherapy. Given the proximity of the larynx to lesions of the base of the tongue, an important consideration in planning therapy is optimizing functions of speech and swallowing while eradicating disease.

Table 61–3 illustrates 5-year overall survival rates for patients with base of the tongue carcinoma treated with different treatment modalities.

Using surgery as definitive treatment with postoperative radiotherapy for advanced disease, 5-year survival rates for patients with stage I and II lesions were found to be 77%, whereas that for stage III lesions was 64% and for stage IV disease 59%.[137] In this study, mandible continuity was maintained in 86%, and the larynx was preserved in 80%.

Combination therapy using external beam radiotherapy and implantation of iridium 192 has been employed for T1 to T4 base of the tongue cancers. Using irradiation alone for N0 necks and treating N+ necks with irradiation and neck dissection, overall 5- and 10-year survival rates have been reported at 86% and 52%, respectively.[138]

Five-year overall and disease-free survival rates for patients treated with continuous-course external beam irradiation (with or without neck dissection) who did not receive interstitial implantation were reported to be 43% and 58%, respectively.[139] This study found improved local control for T4 tumors treated with twice-a-day fractionation compared with once-a-day fractionation.

Finally, in light of the favorable results using induction chemotherapy followed by irradiation in the treatment of advanced laryngeal cancer, similar regimens have been studied for carcinoma of the

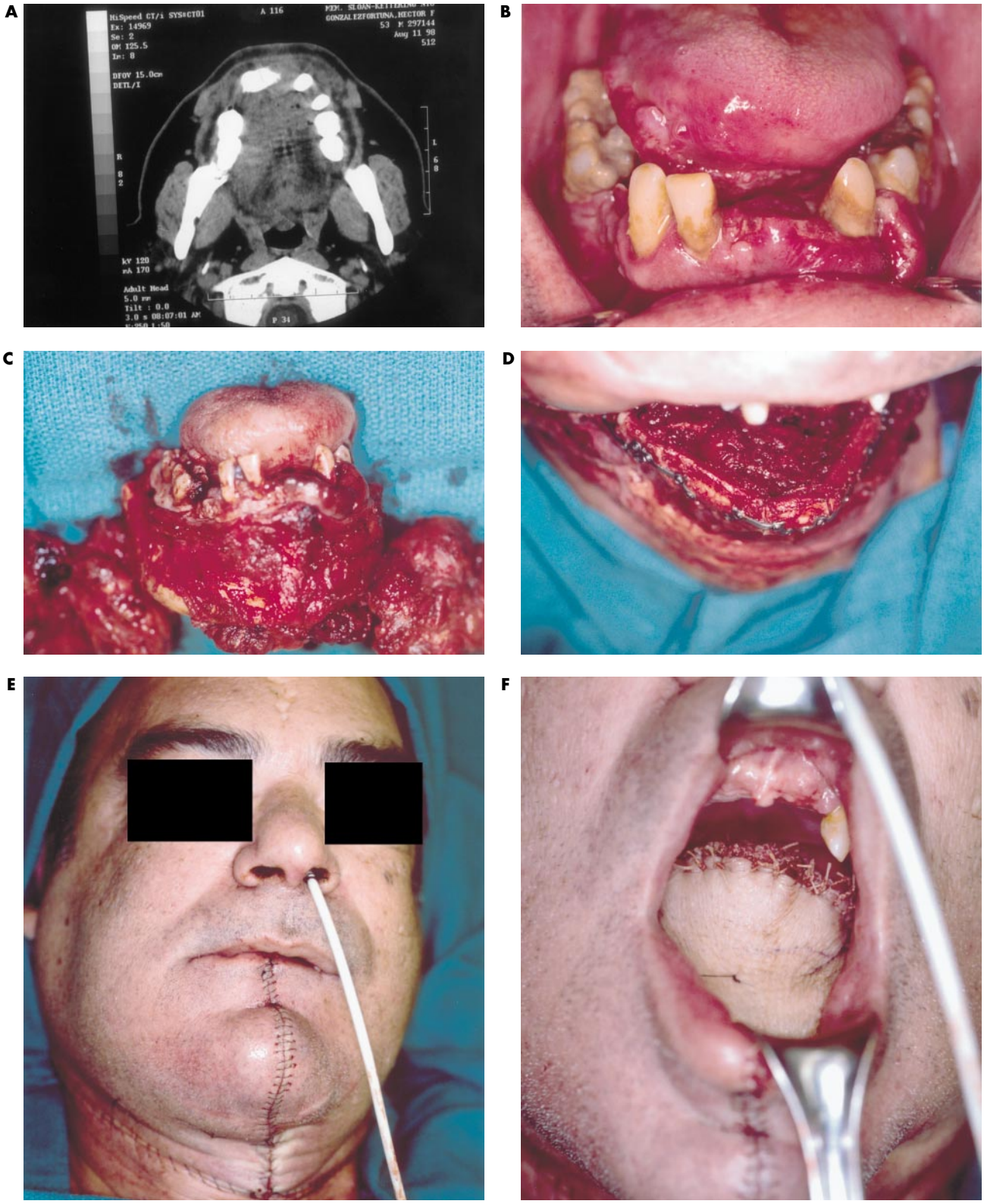


FIGURE 61–13. *A*, Axial computed tomographic scan of the neck demonstrating extensive squamous cell carcinoma of the floor of the mouth. *B*, Tumor extended through the dental alveoli of the anterior part of the mandible, as well as into the anterior portion of the oral tongue. *C*, Surgical specimen demonstrating en bloc segmental mandibular resection, in continuity with the oral tongue, floor of the mouth, and contents of bilateral modified radical neck dissections. *D*, Reconstruction of the anterior part of the mandible using contoured fibular free flap. *E*, Closure of the midline lip split incision. *F*, Soft tissue reconstruction using radial forearm free flap.

TABLE 61–3. 5-Year Overall Survival for Base of the Tongue Squamous Cell Carcinoma

Surgery with postoperative radiotherapy (137 patients): 55%
External beam radiotherapy (138 patients): 43%
External beam radiotherapy with brachytherapy (139 patients): 86%
Chemotherapy with radiotherapy (140 patients): 41%

base of the tongue. Patients in one study with advanced SCC of the oropharynx were given one to three cycles of cisplatin, followed by definitive radiotherapy, for a complete or partial response at the primary site. External beam radiotherapy was administered with or without interstitial implant and with or without accompanying neck dissection. Five-year overall and failure-free survival rates were reported as 41% and 42%, respectively.[140]

Comparing the results of different treatment modalities is difficult owing to the inherent diversity of patient samples and extent of disease included in reported studies. Primary radiotherapy is associated with favorable rates of local control and overall survival, but large, deeply invasive tumors may not be as radiosensitive as exophytic lesions. Larger, well-defined tumors can often be resected with preservation of the mandible and larynx, with satisfactory function of speech and swallowing.[137] Further investigation is necessary to compare functional results following surgery versus those following radiotherapy. Regardless of the modality chosen for the primary tumor, the propensity for base of the tongue carcinoma to spread to regional lymphatics mandates an aggressive approach to the neck. Neck dissection should be considered for clinically evident nodal metastases, recognizing the potential for bilateral involvement.

Soft Palate

Surgery or radiotherapy provides good local control and survival rates in early-stage (stages I and II) tumors, but combination therapy, using surgical resection followed by adjuvant radiation therapy, is recommended for stage III and IV tumors. Resection of large cancers of the soft palate may be complicated by postoperative velopharyngeal insufficiency, best treated with a palatal obturator.

Tonsil

Of all of the tumors in the oral cavity and oropharynx, cancers of the tonsil are the most radiosensitive, particularly exophytic lesions. Diagnosis at an early stage (stages I and II) is uncommon in light of the disease being present for some time before the patient presents complaining of symptoms. Stage I and II tumors may be best approached with primary radiotherapy, recognizing the potential for occult nodal metastases. Treatment planning for stage III and IV tumors is slightly more complicated as the method of choice may differ depending on the particular situation. A small tonsil primary with extensive nodal metastases may be treated with primary radiotherapy with consideration for neck dissection. In contrast, a large primary tumor at the tonsillar region without regional lymph node metastases could be approached with combination surgery and postoperative radiotherapy or with combination chemotherapy and irradiation. The following 3-year survival rates, using radiotherapy for stage I and II tumors and combined therapy with surgery and postoperative radiation therapy for stage III and IV disease, have been reported: stage I, 89%; stage II, 83%; stage III, 58%; and stage IV, 49%.[141]

Posterior Pharyngeal Wall

Cancers of the posterior pharyngeal wall commonly present with dysphagia and odynophagia. Alternatively, the patient may complain of a globus sensation or change in voice.

Second primary tumors are common, affecting 37%, most commonly arising in the oral cavity or larynx.[142] Spread to retropharyngeal lymph nodes may occur in 30 to 40% of cases.

Both surgery and radiation have been singly employed, often with disappointing results. Overall 5-year survival rates using surgery were reported at 32%, ranging from 44% for favorable lesions to 15% for extensive tumors.[142] Posterior pharyngeal wall resection with laryngeal preservation and radial forearm free flap reconstruction may be appropriate for selected lesions. This approach, however, is not recommended for patients with significant comor-

bidities owing to the potential for postoperative complications.[143] Previously reported single-modality radiation survival rates have been dismal, with approximately 3% surviving to 5 years.[144] Although the data are sparse, it appears that combination therapy of surgery and postoperative radiotherapy or chemoradiation may be indicated, particularly for advanced disease.

RECONSTRUCTION OF THE OROPHARYNX

As in the oral cavity, surgical defects in the oropharynx may be approached in a variety of ways. Certain lesions may not require complex reconstructive efforts, being amenable to primary closure or healing by secondary intention. Larger defects may require split-thickness skin graft coverage, local or regional flap reconstruction, or microvascular free tissue transfer. As in the oral cavity, the need for reconstruction is based on a variety of factors, including tumor extent and location; possible dysfunction of breathing, swallowing, or speech; patient comorbidity; and the surgeon's experience and preference.

Kraus has previously reported on reconstruction following resection of cancers of the oropharynx.[145] The majority of cases involving resection of a base of the tongue primary may be closed primarily. Early-stage (stage I and II) cancers of the tonsil may heal by secondary intention following peroral resection. More advanced lesions of the tonsil may require flap reconstruction, depending on the size of the mucosal defect and the need to resect portions of the mandible. Tension-free primary closure is often possible following segmental mandibulectomy. Advanced lesions necessitating a mandibulotomy approach with mandibular preservation may require flap reconstruction to resurface the inner portion of the mandible. The thin, pliable nature of the radial forearm free flap makes it a preferable choice over the pectoralis major myocutaneous flap for reconstruction in this region. Similarly, the radial forearm free flap may be a suitable choice for reconstruction of extensive soft palate defects. For posterior pharyngeal wall defects, laryngeal preservation with radial forearm free flap reconstruction should be reserved for patients with high performance status in the hope of minimizing postoperative morbidity.

REFERENCES

1. American Joint Committee on Cancer. AJCC cancer staging manual. 5th ed. New York: Lippincott-Raven; 1997.
2. Lindberg R. Distribution of cervical lymph node metastases from squamous cell carcinoma of the upper respiratory and digestive tracts. Cancer 1972; 29:1446–9.
3. Shah JP. Patterns of cervical lymph node metastasis from squamous carcinomas of the upper aerodigestive tract. Am J Surg 1990;160:405–9.
4. Greenlee RT, Murray T, Bolden S, Wingo PA. Cancer statistics, 2000. CA Cancer J Clin 2000;50:7–33.
5. Deleyiannis FW, Thomas DB, Vaughan TL, Davis S. Alcoholism: independent predictor of survival in patients with head and neck cancer. J Natl Cancer Inst 1996;88:542–9.
6. Talamini R, La Vecchia C, Levi F, et al. Cancer of the oral cavity and pharynx in nonsmokers who drink alcohol and in nondrinkers who smoke tobacco. J Natl Cancer Inst 1998;90:1901–3.
7. Rothman K, Keller A. The effect of joint exposure to alcohol and tobacco on risk of cancer of the mouth and pharynx. J Chron Dis 1972;25:711–6.
8. Mashberg A, Boffetta P, Winkelman R, Garfinkel L. Tobacco smoking, alcohol drinking, and cancer of the oral cavity and oropharynx among U.S. veterans. Cancer 1993;72:1369–75.
9. Blot WJ, McLaughlin JK, Winn DM, et al. Smoking and drinking in relation to oral and pharyngeal cancer. Cancer Res 1988;48:3282–7.
10. Reddy CR, Raju MV, Ramulu C, Reddy PG. Changes in the ducts of the glands of the hard palate in reverse smokers. Cancer 1972;30:231–8.
11. Norton SA. Betel: consumption and consequences. J Am Acad Dermatol 1998;38:81–8.
12. Merchant A, Husain SS, Hosain M, et al. Paan without tobacco: an independent risk factor for oral cancer. Int J Cancer 2000;86:128–31.
13. Winn DM, Blot WJ, Shy CM, et al. Snuff dipping and oral cancer among women in the southern United States. N Engl J Med 1981;304:745–9.
14. Winn DM, Blot WJ, McLaughlin JK, et al. Mouthwash use and oral conditions in the risk of oral and pharyngeal cancer. Cancer Res 1991;51:3044–7.
15. Blot WJ. Alcohol and cancer. Cancer Res 1992;52(7 Suppl):2119s–23s.
16. Brandsma JL, Abramson AL. Association of papillomavirus with cancers of the head and neck. Arch Otolaryngol Head Neck Surg 1989;115:621–5.

17. Smith EM, Hoffman HT, Summersgill KS, et al. Human papillomavirus and risk of oral cancer. Laryngoscope 1998;108:1098–103.
18. Singh B, Balwally AN, Shaha AR, et al. Upper aerodigestive tract squamous cell carcinoma. The human immunodeficiency virus connection [published erratum appears in Arch Otolaryngol Head Neck Surg 1996;122:944]. Arch Otolaryngol Head Neck Surg 1996;122:639–43.
19. Baker SR. Risk factors in multiple carcinomas of the lip. Otolaryngol Head Neck Surg 1980;88:248–51.
20. Slaughter DP, Southwick HW, Smejkal W. "Field cancerization" in oral stratified squamous epithelium: clinical implications of multicentric origin. Cancer (Phila) 1953;6:963–8.
21. Day GL, Blot WJ. Second primary tumors in patients with oral cancer. Cancer 1992;70:14–9.
22. Day GL, Blot WJ, Shore RE, et al. Second cancers following oral and pharyngeal cancers: role of tobacco and alcohol. J Natl Cancer Inst 1994;86:131–7.
23. Moore C. Cigarette smoking and cancer of the mouth, pharynx, and larynx. A continuing study. JAMA 1971;218:553–8.
24. Knudson AG. Hereditary cancer, oncogenes, and antioncogenes. Cancer Res 1985;45:1437–43.
25. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759–67.
26. Califano J, van der Riet P, Westra W, et al. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res 1996;56: 2488–92.
27. Califano J, Westra WH, Meininger G, et al. Genetic progression and clonal relationship of recurrent premalignant head and neck lesions. Clin Cancer Res 2000;6:347–52.
28. Califano J, Westra WH, Koch W, et al. Unknown primary head and neck squamous cell carcinoma: molecular identification of the site of origin. J Natl Cancer Inst 1999;91:599–604.
29. Boyle JO, Hakim J, Koch W, et al. The incidence of P53 mutations increases with progression of head and neck cancer. Cancer Res 1993;53:4477–80.
30. Brennan JA, Boyle JO, Koch WM, et al. Association between cigarette smoking and mutation of the p53 gene in squamous-cell carcinoma of the head and neck. N Engl J Med 1995;332:712–7.
31. Brennan JA, Mao L, Hruban RH, et al. Molecular assessment of histopathological staging in squamous-cell carcinoma of the head and neck. N Engl J Med 1995;332:429–35.
32. Koch WM, Brennan JA, Zahurak M, et al. p53 mutation and locoregional treatment failure in head and neck squamous cell carcinoma. J Natl Cancer Inst 1996;88:1580–6.
33. Waldron CA, Shafer WG. Leukoplakia revisited. A clinicopathologic study of 3256 oral leukoplakias. Cancer 1975;36:1386–92.
34. Batsakis JG, Cleary KR, Cho KJ. Lichen planus and lichenoid lesions of the oral cavity. Ann Otol Rhinol Laryngol 1994;103:495–7.
35. Shafer WG, Waldron CA. Erythroplakia of the oral cavity. Cancer 1975;36:1021–8.
36. Mashberg A, Meyers H. Anatomical site and size of 222 early asymptomatic oral squamous cell carcinomas: a continuing prospective study of oral cancer. II. Cancer 1976;37:2149–57.
37. Pillai R, Balaram P, Reddiar KS. Pathogenesis of oral submucous fibrosis. Relationship to risk factors associated with oral cancer. Cancer 1992;69: 2011–20.
38. Haider SM, Merchant AT, Fikree FF, Rahbar MH. Clinical and functional staging of oral submucous fibrosis. Br J Oral Maxillofac Surg 2000;38:12–5.
39. Lichtiger B, Mackay B, Tessmer CF. Spindle-cell variant of squamous carcinoma. A light and electron microscopic study of 13 cases. Cancer 1970;26: 1311–20.
40. Takagi M, Sakota Y, Takayama S, Ishikawa G. Adenoid squamous cell carcinoma of the oral mucosa: report of two autopsy cases. Cancer 1977;40:2250–5.
41. Ackerman LV. Verrucous carcinoma of the oral cavity. Surgery 1948;23:678–90.
42. Kraus FT, Perezmesa C. Verrucous carcinoma. Clinical and pathologic study of 105 cases involving oral cavity, larynx and genitalia. Cancer 1966;19: 26–38.
43. Batsakis JG, Hybels R, Crissman JD, Rice DH. The pathology of head and neck tumors: verrucous carcinoma, part 15. Head Neck Surg 1982;5:29–38.
44. Medina JE, Dichtel W, Luna MA. Verrucous-squamous carcinomas of the oral cavity. A clinicopathologic study of 104 cases. Arch Otolaryngol 1984;110: 437–40.
45. Proffitt SD, Spooner TR, Kosek JC. Origin of undifferentiated neoplasm from verrucous epidermal carcinoma of oral cavity following irradiation. Cancer 1970;26:389–93.
46. McClure DL, Gullane PJ, Slinger RP, Wysocki GP. Verrucous carcinoma—changing concepts in management. J Otolaryngol 1984;13:7–12.

47. Davidson TM, Haghighi P, Astarita R, et al. MOHS for head and neck mucosal cancer: report on 111 patients. Laryngoscope 1988;98:1078–83.
48. Close LG, Brown PM, Vuitch MF, et al. Microvascular invasion and survival in cancer of the oral cavity and oropharynx. Arch Otolaryngol Head Neck Surg 1989;115:1304–9.
49. Farr HW, Goldfarb PM, Farr CM. Epidermoid carcinoma of the mouth and pharynx at Memorial Sloan-Kettering Cancer Center, 1965 to 1969. Am J Surg 1980;140:563–7.
50. Shah JP. Cancer of the upper aerodigestive tract. In: Alfonso AE, Gardner B, editors. The practice of cancer surgery. New York: Appleton-Century-Crofts; 1982.
51. Snow GB, Annyas AA, van Slooten EA, et al. Prognostic factors of neck node metastasis. Clin Otolaryngol 1982;7:185–92.
52. Johnson JT, Barnes EL, Myers EN, et al. The extracapsular spread of tumors in cervical node metastasis. Arch Otolaryngol 1981;107:725–9.
53. Merino OR, Lindberg RD, Fletcher GH. An analysis of distant metastases from squamous cell carcinoma of the upper respiratory and digestive tracts. Cancer 1977;40:145–51.
54. Farr HW, Arthur K. Epidermoid carcinoma of the mouth and pharynx 1960-1964. J Laryngol Otol 1972;86:243–53.
55. Myers JN, Elkins T, Roberts D, Byers RM. Squamous cell carcinoma of the tongue in young adults: increasing incidence and factors that predict treatment outcomes. Otolaryngol Head Neck Surg 2000;122:44–51.
56. Johnston WD, Byers RM. Squamous cell carcinoma of the tonsil in young adults. Cancer 1977;39:632–6.
57. Pitman KT, Johnson JT, Wagner RL, Myers EN. Cancer of the tongue in patients less than forty. Head Neck 2000;22:297–302.
58. Friedlander PL, Schantz SP, Shaha AR, et al. Squamous cell carcinoma of the tongue in young patients: a matched-pair analysis. Head Neck 1998; 20:363–8.
59. Shah JP, Huvos AG, Strong EW. Mucosal melanomas of the head and neck. Am J Surg 1977;134:531–5.
60. Snow GB, van der Waal I. Mucosal melanomas of the head and neck. Otolaryngol Clin North Am 1986;19:537–47.
61. Rubright WC, Hoffman HT, Lynch CF, et al. Risk factors for advanced-stage oral cavity cancer. Arch Otolaryngol Head Neck Surg 1996;122:621–6.
62. Shah JP. Cervical lymph node metastases—diagnostic, therapeutic, and prognostic implications. Oncology (Huntingt) 1990;4:61–9.
63. Hillsamer PJ, Schuller DE, McGhee RB, et al. Improving diagnostic accuracy of cervical metastases with computed tomography and magnetic resonance imaging. Arch Otolaryngol Head Neck Surg 1990;116:1297–301.
64. McGuirt WF, Williams DW, Keyes JW, et al. A comparative diagnostic study of head and neck nodal metastases using positron emission tomography. Laryngoscope 1995;105(4 Pt 1):373–5.
65. Ali S, Tiwari RM, Snow GB. False-positive and false-negative neck nodes. Head Neck Surg 1985;8:78–82.
66. McGregor AD, MacDonald DG. Patterns of spread of squamous cell carcinoma within the mandible. Head Neck 1989;11:457–61.
67. Shaha AR. Preoperative evaluation of the mandible in patients with carcinoma of the floor of mouth. Head Neck 1991;13:398–402.
68. King JM, Caldarelli DD, Petasnick JP. DentaScan: a new diagnostic method for evaluating mandibular and maxillary pathology. Laryngoscope 1992;102:379–87.
69. Yanagisawa K, Friedman CD, Vining EM, Abrahams JJ. DentaScan imaging of the mandible and maxilla. Head Neck 1993;15:1–7.
70. Maisel RH, Vermeersch H. Panendoscopy for second primaries in head and neck cancer. Ann Otol Rhinol Laryngol 1981;90(5 Pt 1):460–4.
71. Pugliano FA, Piccirillo JF, Zequeira MR, et al. Clinical-severity staging system for oral cavity cancer: five-year survival rates. Otolaryngol Head Neck Surg 1999;120:38–45.
72. Pugliano FA, Piccirillo JF, Zequeira MR, et al. Clinical-severity staging system for oropharyngeal cancer: five-year survival rates. Arch Otolaryngol Head Neck Surg 1997;123:1118–24.
73. Spiro RH, Gerold FP, Strong EW. Mandibular "swing" approach for oral and oropharyngeal tumors. Head Neck Surg 1981;3:371–8.
74. Spiro RH, Gerold FP, Shah JP, et al. Mandibulotomy approach to oropharyngeal tumors. Am J Surg 1985;150:466–9.
75. Dubner S, Spiro RH. Median mandibulotomy: a critical assessment. Head Neck 1991;13:389–93.
76. Callery CD, Spiro RH, Strong EW. Changing trends in the management of squamous carcinoma of the tongue. Am J Surg 1984;148:449–54.
77. Spiro RH, Strong EW. Discontinuous partial glossectomy and radical neck dissection in selected

patients with epidermoid carcinoma of the mobile tongue. Am J Surg 1973;126:544–6.
78. Totsuka Y, Usui Y, Tei K, et al. Results of surgical treatment for squamous carcinoma of the lower alveolus: segmental vs. marginal resection. Head Neck 1991;13:114–20.
79. Barttelbort SW, Ariyan S. Mandible preservation with oral cavity carcinoma: rim mandibulectomy versus sagittal mandibulectomy. Am J Surg 1993; 166:411–5.
80. Loree TR, Strong EW. Significance of positive margins in oral cavity squamous carcinoma. Am J Surg 1990;160:410–4.
81. Looser KG, Shah JP, Strong EW. The significance of "positive" margins in surgically resected epidermoid carcinomas. Head Neck Surg 1978;1: 107–11.
82. Johnson RE, Sigman JD, Funk GF, et al. Quantification of surgical margin shrinkage in the oral cavity. Head Neck 1997;19:281–6.
83. Wang CC. Radiation therapy in the management of oral malignant disease. Otolaryngol Clin North Am 1979;12:73–80.
84. Wang CC, Montgomery W, Efird J. Local control of oropharyngeal carcinoma by irradiation alone. Laryngoscope 1995;105(5 Pt 1):529–33.
85. Fu KK, Pajak TF, Trotti A, et al. A Radiation Therapy Oncology Group (RTOG) phase III randomized study to compare hyperfractionation and two variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: first report of RTOG 9003. Int J Radiat Oncol Biol Phys 2000;48:7–16.
86. Pernot M, Hoffstetter S, Peiffert D, et al. Role of interstitial brachytherapy in oral and oropharyngeal carcinoma: reflection of a series of 1344 patients treated at the time of initial presentation. Otolaryngol Head Neck Surg 1996;115:519–26.
87. Larson DL, Lindberg RD, Lane E, Goepfert H. Major complications of radiotherapy in cancer of the oral cavity and oropharynx. A 10 year retrospective study. Am J Surg 1983;146:531–6.
88. Marx RE, Johnson RP. Studies in the radiobiology of osteoradionecrosis and their clinical significance. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1987;64:379–90.
89. Kramer S, Gelber RD, Snow JB, et al. Combined radiation therapy and surgery in the management of advanced head and neck cancer: final report of study 73-03 of the Radiation Therapy Oncology Group. Head Neck Surg 1987;10:19–30.
90. Huang DT, Johnson CR, Schmidt-Ullrich R, Grimes M. Postoperative radiotherapy in head and neck carcinoma with extracapsular lymph node extension and/or positive resection margins: a comparative study. Int J Radiat Oncol Biol Phys 1992;23:737–42.
91. Vikram B, Strong EW, Shah J, Spiro RH. Elective postoperative radiation therapy in stages III and IV epidermoid carcinoma of the head and neck. Am J Surg 1980;140:580–4.
92. Vikram B, Strong EW, Shah JP, Spiro R. Failure in the neck following multimodality treatment for advanced head and neck cancer. Head Neck Surg 1984;6:724–9.
93. Khafif RA, Rafla S, Tepper P, et al. Effectiveness of radiotherapy with radical neck dissection in cancers of the head and neck. Arch Otolaryngol Head Neck Surg 1991;117:196–9.
94. Schuller DE, Metch B, Stein DW, et al. Preoperative chemotherapy in advanced resectable head and neck cancer: final report of the Southwest Oncology Group. Laryngoscope 1988;98:1205–11.
95. Jacobs JR, Fu KK, Lowry LD, et al. 5-year results of cisplatin and fluorouracil infusion in head and neck cancer. Arch Otolaryngol Head Neck Surg 1991;117:288–91.
96. Laramore GE, Scott CB, Al Sarraf M, et al. Adjuvant chemotherapy for resectable squamous cell carcinomas of the head and neck: report on Intergroup Study 0034. Int J Radiat Oncol Biol Phys 1992;23:705–13.
97. Department of Veterans Affairs Laryngeal Cancer Study Group. Induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. The Department of Veterans Affairs Laryngeal Cancer Study Group. N Engl J Med 1991;324:1685–90.
98. Calais G, Alfonsi M, Bardet E, et al. Randomized trial of radiation therapy versus concomitant chemotherapy and radiation therapy for advanced-stage oropharynx carcinoma. J Natl Cancer Inst 1999;91:2081–6.
99. Hong WK, Endicott J, Itri LM, et al. 13-*cis*-retinoic acid in the treatment of oral leukoplakia. N Engl J Med 1986;315:1501–5.
100. Hong WK, Lippman SM, Itri LM, et al. Prevention of second primary tumors with isotretinoin in squamous-cell carcinoma of the head and neck. N Engl J Med 1990;323:795–801.
101. Mao L, El-Naggar AK, Papadimitrakopoulou V, et al. Phenotype and genotype of advanced prema-

lignant head and neck lesions after chemopreventive therapy. J Natl Cancer Inst 1998;90:1545–51.
102. Vandenbrouck C, Sancho-Garnier H, Chassagne D, et al. Elective versus therapeutic radical neck dissection in epidermoid carcinoma of the oral cavity: results of a randomized clinical trial. Cancer 1980; 46:386–90.
103. Andersen PE, Cambronero E, Shaha AR, Shah JP. The extent of neck disease after regional failure during observation of the N0 neck. Am J Surg 1996;172:689–91.
104. Hughes CJ, Gallo O, Spiro RH, Shah JP. Management of occult neck metastases in oral cavity squamous carcinoma. Am J Surg 1993;166:380–3.
105. Crile GW. Excision of cancer of the head and neck with special reference to the plan of dissection based on one hundred and thirty-two operations. JAMA 1906;47:1780–6.
106. Martin H, DelValle B, Ehrlich H, Cahan WG. Neck dissection. Cancer 1951;4:441–99.
107. Schuller DE, Reiches NA, Hamaker RC, et al. Analysis of disability resulting from treatment including radical neck dissection or modified neck dissection. Head Neck Surg 1983;6:551–8.
108. Shah JP, Candela FC, Poddar AK. The patterns of cervical lymph node metastases from squamous carcinoma of the oral cavity. Cancer 1990;66: 109–13.
109. Spiro JD, Spiro RH, Shah JP, et al. Critical assessment of supraomohyoid neck dissection. Am J Surg 1988;156:286–9.
110. Byers RM. Modified neck dissection. A study of 967 cases from 1970 to 1980. Am J Surg 1985;150: 414–21.
111. Medina JE, Byers RM. Supraomohyoid neck dissection: rationale, indications, and surgical technique. Head Neck 1989;11:111–22.
112. Byers RM, Weber RS, Andrews T, et al. Frequency and therapeutic implications of "skip metastases" in the neck from squamous carcinoma of the oral tongue. Head Neck 1997;19:14–9.
113. O'Brien CJ, Traynor SJ, McNeil E, et al. The use of clinical criteria alone in the management of the clinically negative neck among patients with squamous cell carcinoma of the oral cavity and oropharynx. Arch Otolaryngol Head Neck Surg 2000;126:360–5.
114. Davidson BJ, Kulkarny V, Delacure MD, Shah JP. Posterior triangle metastases of squamous cell carcinoma of the upper aerodigestive tract. Am J Surg 1993;166:395–8.
115. Koch WM, Choti MA, Civelek AC, et al. Gamma probe-directed biopsy of the sentinel node in oral squamous cell carcinoma. Arch Otolaryngol Head Neck Surg 1998;124:455–9.
116. Chiesa F, Mauri S, Grana C, et al. Is there a role for sentinel node biopsy in early N0 tongue tumors? Surgery 2000;128:16–21.
117. Rabuzzi DD, Chung CT, Sagerman RH. Prophylactic neck irradiation. Arch Otolaryngol 1980;106: 454–5.
118. Weissler MC, Weigel MT, Rosenman JG, Silver JR. Treatment of the clinically negative neck in advanced cancer of the head and neck. Arch Otolaryngol Head Neck Surg 1989;115:691–4.
119. Wilson JS, Kemble JV. Cancer of the lip at risk. Br J Oral Surg 1972;9:186–94.
120. Heller KS, Shah JP. Carcinoma of the lip. Am J Surg 1979;138:600–3.
121. Baker SR, Krause CJ. Carcinoma of the lip. Laryngoscope 1980;90:19–27.
122. Frierson HF, Cooper PH. Prognostic factors in squamous cell carcinoma of the lower lip. Hum Pathol 1986;17:346–54.
123. Bloom ND, Spiro RH. Carcinoma of the cheek mucosa. A retrospective analysis. Am J Surg 1980; 140:556–9.
124. Soo KC, Spiro RH, King W, et al. Squamous carcinoma of the gums. Am J Surg 1988;156:281–5.
125. Wald RM, Calcaterra TC. Lower alveolar carcinoma. Segmental v marginal resection. Arch Otolaryngol 1983;109:578–82.
126. Byers RM, Newman R, Russell N, Yue A. Results of treatment for squamous carcinoma of the lower gum. Cancer 1981;47:2236–8.
127. Evans JF, Shah JP. Epidermoid carcinoma of the palate. Am J Surg 1981;142:451–5.
128. Shaha AR, Spiro RH, Shah JP, Strong EW. Squamous carcinoma of the floor of the mouth. Am J Surg 1984;148:455–9.
129. Hicks WL, Loree TR, Garcia RI, et al. Squamous cell carcinoma of the floor of mouth: a 20-year review. Head Neck 1997;19:400–5.
130. Kowalski LP, Bagietto R, Lara JR, et al. Factors influencing contralateral lymph node metastasis from oral carcinoma. Head Neck 1999;21:104–10.
131. Franceschi D, Gupta R, Spiro RH, Shah JP. Improved survival in the treatment of squamous carcinoma of the oral tongue. Am J Surg 1993; 166:360–5.
132. Fein DA, Mendenhall WM, Parsons JT, et al. Carcinoma of the oral tongue: a comparison of results

and complications of treatment with radiotherapy and/or surgery. Head Neck 1994;16:358–65.
133. Spiro RH, Huvos AG, Wong GY, et al. Predictive value of tumor thickness in squamous carcinoma confined to the tongue and floor of the mouth. Am J Surg 1986;152:345–50.
134. Lydiatt DD, Robbins KT, Byers RM, Wolf PF. Treatment of stage I and II oral tongue cancer. Head Neck 1993;15:308–12.
135. Yuen AP, Wei WI, Wong YM, Tang KC. Elective neck dissection versus observation in the treatment of early oral tongue carcinoma. Head Neck 1997; 19:583–8.
136. Ariyan S. The pectoralis major myocutaneous flap. A versatile flap for reconstruction in the head and neck. Plast Reconstr Surg 1979;63:73–81.
137. Kraus DH, Vastola AP, Huvos AG, Spiro RH. Surgical management of squamous cell carcinoma of the base of the tongue. Am J Surg 1993;166:384–8.
138. Harrison LB, Lee HJ, Pfister DG, et al. Long term results of primary radiotherapy with/without neck dissection for squamous cell cancer of the base of tongue. Head Neck 1998;20:668–73.
139. Foote RL, Parsons JT, Mendenhall WM, et al. Is interstitial implantation essential for successful radiotherapeutic treatment of base of tongue carcinoma? Int J Radiat Oncol Biol Phys 1990;18: 1293–8.
140. Pfister DG, Harrison LB, Strong EW, et al. Organ-function preservation in advanced oropharynx cancer: results with induction chemotherapy and radiation. J Clin Oncol 1995;13:671–80.
141. Spiro JD, Spiro RH. Carcinoma of the tonsillar fossa. An update. Arch Otolaryngol Head Neck Surg 1989; 115:1186–9.
142. Spiro RH, Kelly J, Vega AL, et al. Squamous carcinoma of the posterior pharyngeal wall. Am J Surg 1990;160:420–3.
143. Lydiatt WM, Kraus DH, Cordeiro PG, et al. Posterior pharyngeal carcinoma resection with larynx preservation and radial forearm free flap reconstruction: a preliminary report. Head Neck 1996; 18:501–5.
144. Pene F, Avedian V, Eschwege F, et al. A retrospective study of 131 cases of carcinoma of the posterior pharyngeal wall. Cancer 1978;42:2490–3.
145. Kraus DH. Reconstruction of oropharyngeal defects. In: Proceedings of the Fourth International Conference on Head and Neck Cancer, Toronto. 28 July–1 August 1996, Arlington, VA. Pittsburgh (PA): The Society of Head and Neck Surgeons and the American Society of Head and Neck Surgery; 1996.

CHAPTER 62

Diseases of the Salivary Glands

William R. Carroll, MD, C. Elliott Morgan, DMD, MD

The spectrum of diseases involving the salivary glands includes congenital anomalies, infections, localized and systemic inflammatory conditions, and benign and malignant neoplasms.

ANATOMY

Parotid Gland

The parotid gland is located in the space between the ramus of the mandible and the external auditory canal and mastoid tip. The anterior portion of the gland overlies the masseter muscle, and the posterior portion overlies the sternocleidomastoid muscle. The deep, medial portion of the gland is adjacent to the parapharyngeal space. The styloid process is adjacent and deep to the gland. The carotid artery and internal jugular vein are medial to the styloid process.

The parotid gland is covered by the parotid fascia. This layer is an extension of the superficial layer of the deep cervical fascia that splits to envelop the gland. On its inferomedial surface, the condensation of the fascial envelope suspended between the styloid process and the mandible is termed the stylomandibular ligament. This ligament separates the inferior portion of the parotid gland from the posterior portion of the submandibular gland. The parotid duct (Stensen's duct) exits the superficial part of the gland anteriorly and passes across the masseter muscle at a location halfway between the zygoma and the oral commissure. As the duct reaches the anterior portion of the masseter muscle, it turns medially to pierce the buccinator muscle and enters the buccal mucosa of the oral cavity opposite the second upper molar tooth. Lymph nodes may be found within the capsule of the parotid gland. These lymph nodes drain the scalp, forehead, eye, cheek, and oral commissure.

Surgeons often speak of superficial and deep lobes of the parotid gland and suggest that the facial nerve lies in a true cleavage plane between these two lobes. Most anatomists conclude that there is no true anatomic division between the superficial and deep lobes. These terms simply refer to the presence of tissue both lateral and medial to the facial nerve.

The position of the facial nerve is the dominant consideration in surgery of the parotid gland. As the nerve exits the stylomastoid foramen, it enters the posterior and medial portion of the gland, usually as a single trunk. Its primary bifurcation occurs at the pes anserinus with subsequent secondary division into five facial and cervical branches. The nerve lies within the substance of the gland and is not simply covered by folds of the gland. Numerous variations in the branching pattern of the facial nerve occur, and the surgeon must be observant during approach and dissection of the nerve.[1]

Parotid gland secretions are stimulated by parasympathetic fibers originating in the inferior salivatory nucleus of the brainstem. These fibers join the glossopharyngeal nerve and then the lesser petrosal nerve before joining the otic ganglion just outside the foramen ovale. The auriculotemporal nerve picks up these parasympathetic fibers and distributes them to the parotid gland.

Submandibular Gland

The submandibular gland is located within the submandibular triangle of the neck. Like the parotid gland, it is covered with a fascial capsule, which originates from the superficial layer of the deep cervical fascia. At the posterior border of the gland, the stylomandibular ligament separates the submandibular gland from the parotid gland. The anterior

border of the submandibular gland is folded over the posterior border of the mylohyoid muscle, creating portions of the gland both superficial and deep to the mylohyoid muscle. The submandibular gland duct (Wharton's duct) arises from the medial aspect of the gland and proceeds submucosally in the floor of the mouth medial to the sublingual gland to reach its papilla at the side of the frenulum of the tongue. The vascular supply to the submandibular gland is from the facial artery, which enters the gland just superior to the posterior belly of the digastric muscle. Small branches from the facial artery enter the gland directly. The lingual nerve is located just medial and superior to the gland within the sublingual space. The hypoglossal nerve courses between the medial portion of the submandibular gland and the hyoglossus muscle deep to the digastric muscle. The marginal mandibular branch of the facial nerve often overlies the lateral aspect of the submandibular gland. It is positioned in a plane deep to the platysma muscle but superficial to the submandibular gland fascia. Protection of these nerves during surgery of the submandibular gland will be discussed in more detail.

Parasympathetic stimulation of the submandibular gland causes formation and release of saliva. The responsible parasympathetic fibers originate in the superior salivatory nucleus in the brainstem. These fibers travel with the seventh cranial nerve and enter the chorda tympani nerve as it branches from the facial nerve within the mastoid. After traversing the middle ear and entering the infratemporal fossa, the chorda tympani joins the lingual nerve just below the foramen ovale. The parasympathetic fibers run with the lingual nerve to the submandibular ganglion, which is suspended from the lingual nerve just superior to the gland.

Sublingual Glands

The sublingual glands are located submucosally in the floor of the mouth. They are paired structures and nearly meet in the anterior part of the floor of the mouth. Typically, there is no single duct draining the submandibular glands. Rather, there are 10 to 12 smaller ducts that pass directly into the mucosa of the floor of the mouth. The sublingual glands are innervated by parasympathetic fibers of the chorda tympani nerve.

Minor Salivary Glands

There are hundreds of small, unnamed minor salivary glands distributed throughout the upper aerodigestive tract. These glands are both mucus and serous producing, and each has its own small duct draining directly to the mucous membrane. The minor salivary glands are most prominent in the oral cavity and are located in the hard and soft palate, lips, buccal mucosa, floor of the mouth, and tongue. As is discussed, many of the diseases involving the major salivary glands may also be found within the minor salivary glands. The neoplasms arising in the minor salivary glands are very often malignant.

HISTOLOGY

Microscopically, the salivary glands are composed of acini, secretory tubules, and collecting ducts (Figure 62–1). The acinar cells are responsible for producing the secretions, which may be either serous or mucinous. The acinar cells within the parotid gland are primarily serous. The acinar cells in the sublingual glands are mostly mucinous. The submandibular gland and minor salivary glands contain a mixture of serous- and mucus-producing cells. As will be discussed, neoplasms of the salivary glands arise from the different histologic components of the secretory apparatus (see Figure 62–1).

PHYSIOLOGY

Saliva has several important functions, including initiation of digestion of food, lubrication of food, protection of dental structures, control of oral cavity bacterial counts, and immune system function. When stimulated, the acinar cells of the salivary glands produce a transudate of fluids; proteins and other organic substances are actively excreted. The flow of saliva results from release of stored saliva as well as production and secretion of newly formed saliva. Chemically, normal saliva typically includes 99% water, electrolytes, and organic compounds including proteins, urea, lipids, and amino acids.[2] Included in the protein portion of saliva are amylase, albumin, immunoglobulin A (IgA), and lysozyme. The submandibular glands account for approximately 70% of the total volume of saliva. The parotid glands produce approximately 25%, and the

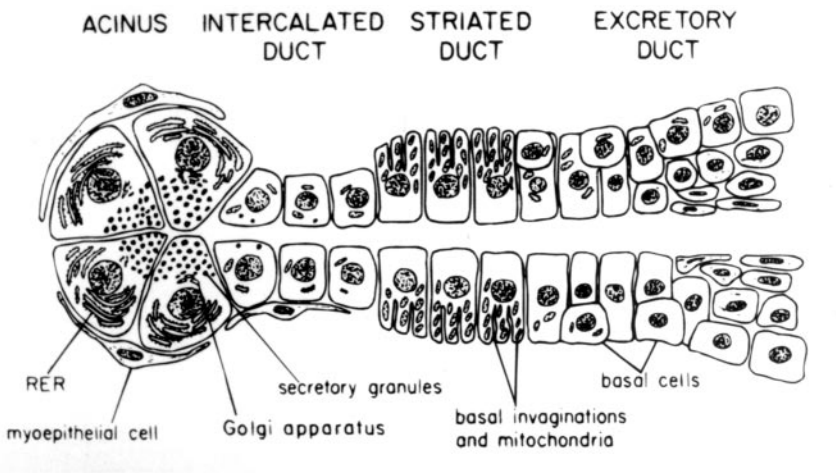


FIGURE 62–1. The salivary gland acinus and duct system. RER = rough endoplasmic reticulum.

sublingual and minor salivary glands are responsible for the balance. There is a baseline secretion rate of saliva that occurs to lubricate and protect the oral cavity and pharynx.

INFLAMMATORY DISEASES OF THE SALIVARY GLANDS

Viral Infections (Mumps)

Mumps (epidemic parotitis) is a disease of viral origin most commonly occurring in the pediatric age group. Viral parotitis is usually caused by a paramyxovirus (specifically the *Rubulavirus*), but many viral pathogens may cause acute infections within salivary glands.

In contrast to bacterial sialadenitis, viral infections of the salivary glands are systemic from the outset. There is a 2- to 3-week incubation period after exposure when the virus multiplies in the upper respiratory tract or parotid gland. A 3- to 5-day period of viremia follows, and the virus localizes in the parotid gland, germinal tissues, and central nervous system. Typical symptoms include fever, malaise, and headaches, followed by tenderness and enlargement of the parotid glands. As the parotid glands enlarge, trismus and dysphagia may develop. Typically, the disease involves the parotid glands primarily, but the submandibular and sublingual glands may be involved on occasion. The disease is usually self-limited and uncomplicated. The swelling of the glands characteristically subsides within 2 weeks.

Complications of viral parotitis may occur owing to the systemic nature of the infection. The most common complication from paramyxovirus parotitis is orchitis, occurring in 20 to 30% of males. Oophoritis occurs in 5% of females. Involvement of the germinal tissues does not usually cause sterility. Aseptic meningitis is a complication in about 10% of patients with viral parotitis. Pancreatitis occurs in 5%.

Sensorineural hearing loss occurs in 0.5 to 4% of patients with mumps syndrome. The hearing loss progresses rapidly and occurs in the later stages of parotitis. Tinnitus, vertigo, and aural fullness develop with the onset of hearing loss. The hearing loss is usually profound, permanent, and unilateral in the majority of cases.

Diagnosis of viral parotitis can be confirmed with viral serology. Complement-fixing antibodies appear following exposure to paramyxovirus. Complete blood count usually shows leukocytopenia with relative lymphocytosis. Amylase levels are increased in the early stages of infection.

Treatment of viral salivary gland infection is supportive. Live attenuated mumps vaccine as part of mumps-measles-rubella immunization is given to children after 12 months of age. Antibody titers are elevated in 90% of recipients. There have been no deaths and very few side effects from the vaccine.[3,4]

Sjögren's Syndrome

Sjögren's syndrome is a chronic autoimmune disorder of the exocrine glands, which affects predominantly, but not exclusively, the salivary glands. It is the second most common autoimmune disease, trailing only rheumatoid arthritis. The signs and symptoms of Sjögren's syndrome were first described by Haddu in 1883. Sjögren, a Swedish ophthalmologist, published a monograph in 1933

describing Sjögren's syndrome as a distinct clinical syndrome with systemic manifestations.

Sjögren's syndrome is now classified in two forms: primary (confined to exocrine glands) or secondary (systemic involvement in extraglandular sites). In 80% of primary and 35% of secondary cases, unilateral or bilateral salivary gland swelling occurs. The pathogenesis of Sjögren's syndrome remains unclear; viral infection is a possible trigger. Sjögren's syndrome has been seen in various ethnic groups. Women with the onset in the fourth to fifth decades of life constitute 90% of cases.

The salivary gland enlargement associated with Sjögren's syndrome is often bilateral and diffuse. The parotid or submandibular glands may be involved, and the swelling in the glands may be episodic and fluctuating. The salivary gland enlargement is typically painless, with firmness throughout the gland.

The most common symptom of Sjögren's syndrome is xerostomia (dry mouth). The decreased salivation causes difficulty with swallowing, altered taste, and speech difficulties. Long-term xerostomia causes an increase in dental caries and tongue atrophy. Ocular involvement in Sjögren's syndrome is characterized as keratoconjunctivitis sicca. Symptoms include pain, burning, and photophobia. Long-term xerophthalmia may cause corneal abrasions, scarring, and vision loss.

The secondary form of Sjögren's syndrome may progress to multisystem involvement. The systemic symptoms are linked to connective tissue disease. Rheumatoid arthritis occurs in 50% of patients with Sjögren's syndrome. Other diseases associated with the secondary form include lupus, vasculitis, scleroderma, dermatomyositis, Crohn's disease, pulmonary interstitial fibrosis, interstitial nephritis, pancreatitis, thyroid disease, and peripheral neuropathy. A high prevalence of sensorineural hearing loss occurs in secondary Sjögren's syndrome.

Asymmetric enlargement of the parotid that persists or is rapid in onset should raise suspicion of lymphoma. Patients have 44 times greater risk of developing lymphoma after developing Sjögren's syndrome. Parotid enlargement in the presence of splenomegaly, lymphadenopathy, immunosuppression, or previous radiation should raise suspicion for lymphoma.

The most appropriate diagnostic test for Sjögren's syndrome is controversial. Schirmer's test is used to document decreased lacrimation. Slit lamp examination can identify corneal ulcerations and keratosis. Sialometry is not of practical use to document xerostomia owing to variable results.

Laboratory tests for Sjögren's syndrome are of limited benefit. SS-A (anti-Ro) and SS-B (anti-La) autoantibodies are found in 38 to 60% and 25 to 40% of patients with primary Sjögren's syndrome, respectively. Enzyme-linked immunosorbent assay techniques may reveal SS-A and SS-B antibodies in 95% and 87% of patients, respectively, but also in 10 to 15% of healthy individuals. Unfortunately, SS-A and SS-B antibodies may be found in other autoimmune diseases. The occurrence of SS-A and SS-B antibodies correlates with earlier onset and longer duration of disease.

Salivary tissue biopsy is the most commonly employed method for diagnosing Sjögren's syndrome. The minor salivary glands of the lower lip are most commonly biopsied owing to ease of access and minimal complications. At least four minor salivary glands should be present in the specimen for a representative biopsy. Histologic diagnosis is based on the presence of more than one cluster of greater than 50 lymphocytes per 4 mm^2. Lip biopsy has the highest specificity (95%), sensitivity (58 to 100%), positive predictive value, and negative predictive value of all tests for Sjögren's syndrome.[5–12]

Human Immunodeficiency Virus

Patients infected with the human immunodeficiency virus (HIV) may develop a spectrum of salivary gland disorders, including diffuse infiltrative lymphocytosis, benign lymphoepithelial lesions, lymphoepithelial cysts, and malignant salivary gland tumors, including lymphoma, Kaposi's sarcoma, and adenoid cystic carcinoma.

Diffuse infiltrative lymphocytosis is very much like Sjögren's syndrome with bilateral parotid enlargement, xerostomia, xerophthalmia, and arthralgia. SS-A, SS-B, and rheumatoid factor serology are negative, but salivary gland biopsy has characteristics of Sjögren's syndrome.

Lymphoepithelial cysts may occur unilaterally or bilaterally in the parotid glands. Multiple lymphoepithelial cysts are so characteristic of HIV infection as to be considered almost pathognomonic. The pathogenesis of the cysts is unknown, and they can occur at any stage of the disease. Needle biopsy of the cysts should be done along with HIV testing.

Patients with HIV with parotid swelling should be evaluated for parotid tumors like uninfected patients. Patients with HIV have a higher incidence of lymphoma, general lymphadenopathy, and Kaposi's sarcoma associated with parotid enlargement than the general population. Fine-needle biopsy is often useful in patients with HIV with parotid enlargement.[13]

Sarcoidosis

The salivary glands are involved in 6 to 10% of patients with sarcoidosis. The parotid gland is involved most commonly. Heerfordt's syndrome (uveoparotid fever) results from sarcoidosis and includes uveitis, lacrimal and salivary gland inflammation, and facial paralysis. The diagnosis is one of exclusion because the noncaseating granulomas are not pathognomonic for the disease. Anemia, thrombocytopenia, leukopenia, eosinophilia, decreased albumin, and hyperglobulinemia support the diagnosis. Elevated angiotensin-converting enzyme levels, sedimentation rates, and calcium levels may also be present. Hilar adenopathy may be present on a chest radiograph.

The most common neurologic manifestation of sarcoidosis is facial paralysis, occurring in 50% of patients with nervous system involvement. The paralysis is bilateral in one-third of patients and usually resolves spontaneously but may recur. Corticosteroids are used to treat patients with more advanced forms of sarcoidosis.[14]

Wegener's Granulomatosis

Wegener's granulomatosis is a systemic disorder often involving the respiratory tract from the nose to the lungs, as well as the kidneys. Wegener's granulomatosis affects both sexes equally, occurs in all ages, and is usually seen in Caucasians. Histologically, Wegener's granulomatosis is characterized by vasculitis of medium and small vessels.

More than 70% of the features associated with Wegener's granulomatosis are related to the ears, nose, and throat. Symptoms develop insidiously, with sinusitis being the most frequent presentation. Salivary gland enlargement may accompany the nasal and sinus symptoms, with the parotid or submandibular glands, or both, being affected. Subglottic stenosis commonly occurs. In the lower airway, pulmonary infiltrates or cavitary nodules are also noted. Renal involvement indicates systemic Wegener's granulomatosis and is the most frequent cause of death.

Laboratory evaluation of Wegener's granulomatosis includes complete blood count, sedimentation rate, autoimmune serology, creatinine level, chest radiograph, and sinus computed tomographic (CT) scan. Elevated levels of antineutrophil cytoplasmic antibody (C-ANCA) are indicative of Wegener's granulomatosis. The sensitivity of C-ANCA is 40 to 90%, depending on the activity of the disease, with a specificity of 90%.

Intranasal biopsies should be taken with a generous section of viable nasal mucosa. These can be taken from the turbinate, septum, or lateral nasal wall. The specimen should be examined for fungal microorganisms as well. Wegener's granulomatosis must be differentiated from other granulomatous diseases including polymorphic reticulosis, Churg-Strauss syndrome, lymphoma, sarcoidosis, scleroma, tuberculosis, relapsing polychondritis, and fungal disease.

The "gold standard" for treatment of Wegener's granulomatosis is cyclophosphamide and glucocorticoids. Methotrexate has been used with corticosteroids in milder forms of Wegener's granulomatosis. Upper airway involvement responds well to trimethoprim and sulfamethoxazole used in combination with cyclophosphamide and glycocorticoids.[15–17]

Acute and Chronic Bacterial Infections of the Salivary Glands

Acute bacterial sialadenitis may involve any salivary gland, although the parotid gland is affected most frequently. Suppurative infection of the salivary glands may be an isolated, acute event or chronic with recurrent acute exacerbations.

Acute bacterial infection of the salivary glands may occur either by retrograde transmission of bacteria from the oral cavity or by stasis of salivary flow. Saliva contains lysozyme and IgA, which protect the salivary glands from infection. Stone formation (sialolithiasis) may cause mechanical obstruction of the salivary duct, causing stasis of flow with resultant bacterial infection. Elderly patients are at high risk for salivary gland infection owing to medications that decrease salivary flow. These medications

include diuretics, antidepressants, beta blockers, anticholinergics, and antihistamines. Patients with chronic, debilitating conditions, patients with compromised immune function, HIV-positive patients, and anorexic and bulimic patients, as well as depressed patients, are at an increased risk for acute bacterial salivary gland infection. Xerostomia of any cause increases the risk for bacterial parotitis.

Staphylococcus aureus is the most common microorganism causing acute bacterial parotitis. Other microorganisms include beta-hemolytic streptococcus, *Haemophilus influenzae*, pneumococcus, and, less frequently, gram-negative microorganisms.

Symptoms of acute bacterial sialadenitis include rapid onset of pain, swelling, induration, and fever. The symptoms may worsen during eating. Examination reveals induration, erythema, edema, and tenderness over the gland and purulence at the ductal orifice.

Bacterial sialadenitis may progress to abscess formation. Pitting edema of the skin overlying the involved gland is said to be indicative of abscess formation. Needle aspiration of the gland may confirm abscess formation. Cultures are done on all aspirated material. Failure to improve after 48 to 72 hours of antibiotic therapy may also indicate abscess formation. Radiographic scans followed by incision and drainage are indicated.

Patients who do not respond to standard therapy for bacterial infection should be evaluated for other disease processes that may mimic bacterial sialadenitis. Lymphoma, cat-scratch disease, Sjögren's syndrome, and Wegener's granulomatosis may mimic acute bacterial sialadenitis. Appropriate imaging and laboratory studies should be done to rule in or out these diseases.

Treatment of acute sialadenitis is directed at improving salivary flow, reversing any underlying conditions responsible for the onset of the infection, and appropriate antibiotic therapy. Antibiotic therapy should be directed toward the gram-positive and anaerobic microorganisms commonly involved. Oral antibiotic therapy can be given if no other systemic illness is present. Incision and drainage are reserved for localized abscesses.

Facial nerve paresis may rarely occur in acute sialadenitis. Because of their location, submandibular and parotid infections may spread into the deep neck spaces. Hematogenous spread of infection can also occur, causing sepsis or even shock and multisystem organ failure in debilitated patients.

Chronic Sialadenitis

Patients with chronic sialadenitis experience recurrent, low-grade inflammation and edema of the gland, minor pain, and sialorrhea that may be slightly purulent. *Streptococcus viridans* is the usual infecting microorganism. Measures to increase salivary flow should be instituted and appropriate antibiotics given. Benign lymphoepithelial lesions may develop from chronic parotitis. Attempts to identify stones or duct strictures should be made. Sialography may reveal a stone or ductal stenosis. The inflammatory process may wax and wane, with symptom-free periods. Conservative therapy usually manages the recurrent inflammation.

Surgical incision is reserved for refractory recurrent infections or in cases with a suspicion of malignancy.[18–25]

IMAGING

Plain radiography, ultrasonography, sialography, radionuclide scanning, CT scanning, and magnetic resonance imaging (MRI) have all been employed to image the salivary glands.

Plain Radiography

Salivary calculi can be detected within the ducts of the major salivary glands on occasion using plain radiographs. Stones occur much more commonly in the submandibular duct and are opaque 80 to 94.7% of the time. Parotid stones are more likely radiolucent, but by using high-quality intraoral plain films, up to 70% of parotid stones may be visualized.

Ultrasonography

Ultrasonography is an inexpensive, noninvasive way to evaluate for submandibular or parotid stones. Up to 90% of stones greater than 2 mm in size can be detected. Ultrasonography also shows the location of the stone and can be used in acute sialadenitis when sialography is contraindicated.

Contrast Sialography

Contrast sialography has been the mainstay for detecting sialolithiasis. Sialography is also used to detect chronic inflammatory changes in the ductal systems. Among the changes seen are dilation of terminal ducts and acini and strictures. Sialography has been replaced by CT and MRI for evaluation of neoplastic parotid masses.

SALIVARY GLAND NEOPLASMS

Salivary gland neoplasms can occur at any site where salivary tissue is found. The most common benign and malignant neoplasms occurring in salivary tissue are discussed below. The predilection of these tumors to occur in specific subsites is highlighted for each tumor type. The common surgical techniques to deal with neoplasms in the parotid, submandibular, and sublingual glands are discussed in detail.

Histogenesis

There is no universal agreement on the cell of origin for salivary gland neoplasms. The two most common theories of origin are the "stem cell" theory and the "multicellular" theory.[26,27] The stem cell theory proposes that the intercalated duct cells and myoepithelial cells of the salivary glands arise from one type of stem cell, and, in turn, these cells give rise to the group of tumors including pleomorphic adenomas, oncocytomas, acinic cell carcinoma, and adenoid cystic carcinoma. The other type of stem cell gives rise to the excretory ducts, and from this lineage, the epidermoid tumors such as mucoepidermoid carcinoma arise.

The multicellular theory attempts to pair each type of parotid neoplasm with a particular cell of origin within the salivary apparatus. Hence, squamous cell and mucoepidermoid carcinomas arise from excretory duct cells, pleomorphic adenomas arise from intercalated duct cells, and Warthin's and oncocytic tumors arise from striated ductal cells. There is no consensus currently whether the stem cell theory or the multicellular theory best explains the histogenesis of salivary gland neoplasms. Recent evidence seems to be swinging more toward the unicellular, stem cell theory.[26]

Benign Salivary Gland Neoplasms

Pleomorphic Adenoma Pleomorphic adenoma or benign mixed tumor is the most common salivary tumor, accounting for up to two-thirds of all salivary gland neoplasms.[28] Approximately 85% of all pleomorphic adenomas are located in the parotid glands, 10% in the minor salivary glands, and 5% in the submandibular glands.[26] Pleomorphic adenomas contain both mesenchymal and epithelial cells. Grossly, the tumors appear encapsulated but, on close inspection, have pseudopod extensions into the surrounding tissues. This growth pattern is thought to be responsible for the high rate of local recurrence (approximately 30%) when these tumors are enucleated.

Within the parotid gland, the majority of pleomorphic adenomas arise in the superficial lobe. Adequate surgical therapy involves nerve identification and protection with removal of the tumor and an adequate cuff of surrounding parotid gland parenchyma. A complete superficial parotidectomy has been recognized for years as the standard procedure in dealing with parotid gland neoplasms. However, most head and neck surgeons now recognize that an "adequate" parotidectomy can be tailored to the size and location for a given parotid tumor.

Genetic alterations associated with the formation of pleomorphic adenoma have been identified. Tumor deoxyribonucleic acid (DNA) has been shown to have chromosome abnormalities involving chromosome 8q12. This region is the site of the pleomorphic adenoma gene *PLAG-1*.[29] In a series of pleomorphic adenomas, loss of heterozygosity at the loci 8q and 12q were detected in 47% and 27%.[30] Carcinoma ex-pleomorphic adenoma has been noted to have some of the same genetic alterations. Investigators have noted a higher expression of p53 in patients with carcinoma ex-pleomorphic adenomas than in benign pleomorphic adenomas.[31]

Warthin's Tumor Warthin's tumor, or papillary cystadenoma lymphomatosum, is the second most common benign neoplasm of the salivary glands. The tumor is named for the University of Michigan pathologist who first reported two cases in the United States. Interestingly, Warthin's tumor occurs almost exclusively in the parotid glands. It typically involves the lower pole of the parotid gland and may be bilateral in up to 10% of cases. The most popular etiologic theory suggests that Warthin's tumor arises in salivary ducts that are trapped within intraparotid

lymph nodes. Because the parotid gland is encapsulated relatively late during development, it is the only salivary gland with lymphoid tissue located within the substance of the gland.

The recommended treatment for Warthin's tumor is complete surgical excision similar to that described for pleomorphic adenomas. In elderly or debilitated patients, a fine-needle aspirate, which is diagnostic of Warthin's tumor, coupled with a consistent clinical picture, may be grounds for expectant, nonsurgical management. No definite etiologic factors or genetic alterations have yet been identified in Warthin's tumor. It is felt that the lymphoid component of Warthin's tumor is benign and derived from a preexisting lymph node.[26,27] Recent DNA studies suggest that no clonal expansion is seen in the cells of Warthin's tumor. In other words, the tumor may eventually be proven to be a reactive process.

Monomorphic Adenomas Monomorphic adenomas include basal cell adenoma, clear cell adenoma, and glycogen-rich adenoma among other less common tumors. The most common monomorphic adenoma is the basal cell adenoma, which comprises 1 to 3% of salivary gland neoplasms.[32] The monomorphic adenomas may be separated into basaloid and nonbasaloid types. The basaloid types occur more commonly in the major salivary glands, particularly the parotid gland.[33] Treatment of monomorphic adenomas includes wide surgical excision with an adequate cuff of normal surrounding tissue.

Oncocytomas Oncocytomas comprise less than 1% of all salivary gland neoplasms. These rare tumors are typically found in the superficial lobe of the parotid gland. Oncocytomas are grossly encapsulated, single lesions. Histologically, the tumors are composed of large cells with round nuclei. The cytoplasm is filled with mitochondria.[26,27] Treatment of oncocytoma, much like the other benign salivary tumors discussed above, involves wide local excision with a cuff of surrounding gland parenchyma. Because they most commonly occur in the parotid gland, a superficial parotidectomy with facial nerve preservation is usually required. Rarely, malignant oncocytomas are detected. Cytologic differentiation from benign oncocytoma can be difficult, and malignancy is usually defined by invasive clinical and histologic features.[26]

Hemangiomas Hemangiomas are the most common tumor arising in the salivary gland from the connective tissue elements. They are the most common salivary gland tumor of any type in children and are often detected within the first year of life. Hemangiomas often occur over the angle of the mandible, and the overlying skin may contain a bluish discoloration. Engorgement of the lesion with crying or straining is often also seen.[28] Most hemangiomas in the salivary glands undergo spontaneous resolution. Nonsurgical therapy may be indicated for massive lesions that are interfering with the aerodigestive tract or causing hematologic complications. Surgical treatment is only rarely recommended.[28]

Malignant Tumors of the Salivary Glands

Salivary gland malignancies make up a relatively small percentage of all cancer occurring in the head and neck region. Benign salivary gland tumors outnumber malignant salivary tumors by approximately 3 to 1 overall.[34]

As a general rule, residents and medical students are taught that 20% of parotid tumors are malignant, 50% of submandibular gland tumors are malignant, and 80% of minor salivary gland and sublingual gland tumors are malignant. Spiro and Spiro analyzed over 7,000 reported salivary gland neoplasms. Seventy-eight percent of the parotid neoplasms were benign, 54% of submandibular neoplasms were benign, and 35% of minor salivary gland neoplasms were benign.[35]

Malignant salivary gland tumors often present in a similar fashion to their benign counterparts. Pain is present in 10 to 29% of patients with cancer in the parotid gland, and facial paralysis is detected in 10 to 15% of parotid gland malignancies.

Because malignant salivary gland tumors are reasonably rare, characteristic genetic alterations, oncogene markers, and predisposing factors are still being determined. A brief synopsis of these factors will be presented with each tumor type. The etiology of malignant salivary gland tumors is not well defined. The clear association with smoking and alcohol use found in many other head and neck carcinomas has not been documented. Radiation exposure is one known risk factor, and there seems to be a dose-response relationship for the development of malignant tumors.[36,37] Of 2,807 salivary gland neo-

plasms presented by Spiro, 35% of the malignant neoplasms were mucoepidermoid carcinomas, 23% were adenoid cystic carcinoma, 18% were adenocarcinoma, 13% were malignant mixed tumor, and 7% were acinic cell carcinoma.[38] Most authors comment on the difficulty encountered in classifying salivary gland tumors, and, indeed, the classification schemes have changed from time to time.

Mucoepidermoid Carcinoma Mucoepidermoid carcinoma is the most common malignant salivary gland tumor. Approximately one half of all mucoepidermoid carcinomas occur in the parotid gland, with the majority of the remainder occurring in minor salivary glands. As the name implies, mucoepidermoid carcinomas are composed of both mucus and epidermoid cells. Pathologists also describe the presence of intermediate and clear cells. Histologic grading of mucoepidermoid carcinoma has correlated with prognosis. Most surgeons recognize high- and low-grade subtypes of mucoepidermoid carcinoma. Some authors include an intermediate grade as well. Clearly, the prognosis with high-grade tumors is worse with respect to local recurrence and distant metastases, and overall survival is diminished. For high-grade lesions, the local recurrence rates following surgery approach 60%. Nodal metastases may develop in 40 to 70% of patients, and 30% develop distant metastases. Five-year survival rates are typically 30 to 50% for high-grade lesions.[38] In contrast, low-grade mucoepidermoid carcinomas have 5-year survival rates in the 80 to 95% range. They are much less likely to develop nodal or distant metastasis. Most patients with mucoepidermoid carcinoma present with an asymptomatic mass at the primary site. Pain, facial paralysis, and neck masses are also seen and are usually associated with high-grade and highly malignant mucoepidermoid carcinomas.

Surgical treatment for mucoepidermoid carcinoma is influenced by histologic grading. For low-grade mucoepidermoid carcinoma, treatment usually involves wide surgical excision. Neck dissection or adjuvant irradiation is used only when clinically evident metastases are detected or when there is evidence of bone, nerve, or extraglandular invasion. In contrast, high-grade mucoepidermoid carcinomas are treated more like squamous cell carcinomas. Wide surgical excision combined with regional lymph node dissection and adjuvant radiation therapy is commonplace. The issue of facial nerve preservation in the presence of high-grade mucoepidermoid carcinoma is controversial. Some surgeons prefer a wide block excision of the parotid gland (radical parotidectomy) with no attempt at facial nerve preservation. Others feel that if the tumor can be easily separable from the nerve and there is no histologic evidence of perineural invasion, the nerve can be preserved. No one recommends preserving the facial nerve when it is clearly invaded by tumor.

Five-year survival rates for low-grade mucoepidermoid carcinoma are 80 to 95% overall. Five-year survival rates for high-grade tumors are 30 to 50%.[38]

Adenoid Cystic Carcinoma Adenoid cystic carcinoma is the second most common malignancy of the salivary glands. It is the most common malignancy in the submandibular glands and minor salivary glands. The most common clinical presentation for adenoid cystic carcinoma is a painless, slowly enlarging mass. However, paresthesias and paralysis are more common with adenoid cystic carcinoma than other salivary malignancies. Indeed, perineural invasion and spread is a hallmark of adenoid cystic carcinoma.[36] There are three common histologic patterns for adenoid cystic carcinoma: tubular, cribriform, and solid. Low-grade adenoid cystic carcinomas typically contain a mixed tubular and cribriform pattern. The solid pattern is more common in high-grade malignancies.[39] The prognosis of adenoid cystic carcinoma may be related to histologic grade. Those patients with low-grade tumors have slow disease progression and infrequent distant metastasis. Patients with high-grade malignancies experience much more rapid growth, with higher frequency of distant metastasis and decreased survival.[39] In the Sloan Kettering series, however, the early trend toward better survival with low-grade lesions disappeared as the patient follow-up exceeded 10 years. In other words, disease progression was slower and less fulminant but relentless and equally deadly. Fifteen- to 20-year survival rates are in the 30% range regardless of histologic grade.[38]

Treatment of adenoid cystic carcinoma includes wide local excision of primary disease and therapeutic lymph node dissection. Elective node dissections are not typically recommended. Full-course radiation therapy is usually recommended

postoperatively.[36] For adenoid cystic carcinoma of the parotid gland, some surgeons have recommended wide-field radical parotidectomy with transection of the facial nerve within the mastoid to best control perineural spread. Fast neutron radiotherapy has proven effective for recurrent or unresectable adenocystic carcinoma. In the University of Washington series, when patients were able to undergo surgical resection (even if microscopic margins were not clear) and there was no direct skull base involvement, 5-year regional control rates of 80% were achieved. The authors concluded that fast neutron radiotherapy is effective for locally advanced adenoid cystic carcinoma but that overall survival is unlikely to be improved without improvement in control of distant metastatic disease.[40]

Polymorphous Low-Grade Adenocarcinoma Polymorphous low-grade adenocarcinoma (PLGA) was first described in 1983 by Batsakis et al and was initially termed "terminal duct adenocarcinoma."[41] In 1984, the term polymorphous low-grade adenocarcinoma was proposed.[26,42]

Polymorphous low-grade adenocarcinoma typically arises from the minor salivary glands and is frequently seen in the oral cavity. The palate is the most common location for PLGA. The lesions typically present as a circumscribed, slow-growing, painless mass. These minor salivary gland malignancies may be confused with pleomorphic adenoma on the benign side and with adenoid cystic carcinoma on the more malignant side.[43] Proper treatment of PLGA is surgical. The tumors are considered low grade, and in a review of over 200 cases, a 17% recurrence rate and 9% regional metastatic rate were noted. Elective node dissection is not recommended.[26,44]

Acinic Cell Carcinoma Acinic cell carcinoma is reasonably rare and comprises roughly 1 to 3% of all salivary gland tumors. In the past, acinic cell carcinoma was known as acinic cell tumor. Most investigators now believe that all of these tumors are malignant. Most acinic cell carcinomas are found in the parotid gland, where they comprise 12 to 17% of parotid malignancies. Bilateral parotid gland involvement has been reported in up to 3% of cases. The biologic behavior may range from slow, indolent, local growth to a more aggressive form with rapid growth and the potential for distant metastasis.

Treatment of acinic cell carcinoma is wide surgical excision. Determinate survival rates were reported by Spiro et al to be 76%, 63%, and 55% at 5, 10, and 15 years, respectively.[26,45] Radiation is not typically used for primary treatment. Adjuvant radiation is often recommended in the face of lymphatic or perineural invasion. The role of elective neck dissection has not been established. As with most parotid malignancies, the authors often perform a selective neck dissection of levels 2 and 3 when resecting acinic cell carcinoma.

Malignant Mixed Tumors This somewhat confusing group of tumors contains at least three distinct tumor types: (1) carcinoma ex-pleomorphic adenoma, (2) malignant mixed tumor (carcinosarcoma), and (3) benign metastasizing pleomorphic adenoma. Carcinoma ex-pleomorphic adenoma is the more common of these three. This tumor represents malignant transformation of the epithelial component of a pleomorphic adenoma. The risk of malignant transformation increases from 1.5 to approximately 9.5% over 5 to 15 years duration of the pleomorphic adenoma.[36,46] The biologic behavior of carcinoma ex-pleomorphic adenoma is generally more aggressive than other salivary gland malignancies. Treatment is wide local excision with consideration of lymph node dissection and postoperative radiation therapy. Five- and 10-year survival rates are 40 and 24%, respectively.[38]

Malignant mixed tumor is less common and contains malignant epithelial and mesenchymal components. This is a true carcinosarcoma, and both components of the malignancy are evident in metastatic sites. The sarcomatous component is often differentiated as a chondrosarcoma, but other types of sarcoma are also described.[43] Treatment is wide surgical excision and radiation. Distant metastases are common, and the prognosis is poor.

Metastasizing pleomorphic adenoma represents an unusual lesion in which benign-appearing pleomorphic adenoma appears in regional lymph nodes. A 22% mortality rate associated with this finding has been reported. There is a suggestion that the risk of developing metastasizing pleomorphic adenoma increases with the longevity of the original tumor and with local recurrence.

Squamous Cell Carcinoma Primary squamous cell carcinoma of the salivary glands is rare and probably occurs only in the parotid and submandibular glands. Squamous cell carcinoma is much more commonly metastatic to the parotid gland parenchyma or intraparotid lymph nodes. The diagnosis of squamous cell carcinoma within the salivary gland should always trigger an intensive search for a primary lesion.

Primary squamous cell carcinoma of the parotid typically presents as a firm, fixed mass. Treatment is aggressive wide surgical excision with consideration of regional lymph node dissection. Adjuvant radiation therapy would usually be recommended. Five-year survival is estimated at 45%.[36]

SURGERY OF THE MAJOR SALIVARY GLANDS

Surgery of the Submandibular Glands

The incisions for approach to the submandibular gland vary somewhat depending on the pathology. For benign neoplasms and inflammatory processes, a reasonably short, transverse incision in a natural skin crease two fingerbreadths below the angle of the mandible is used. If there is malignancy within the submandibular gland and a neck dissection is also anticipated, the preferred incision for neck dissection is used.

The skin and subcutaneous tissues are elevated in the subplatysmal layer. At the inferior border or the gland, flap elevation using cold dissection with scissors or scalpel is preferred to avoid damage to the underlying marginal mandibular nerve. The marginal mandibular nerve lies between the platysma/superficial musculoaponeurotic system (SMAS) layer and the superficial layer of the deep cervical fascia (capsule of the submandibular gland). The nerve lies superficial to the facial vein at the posterior border of the gland as well. If the skin flaps have been elevated in a proper plane, the nerve is often easily seen. It is typically small (1 to 2 mm in diameter) and often is accompanied by a small vasa nervorum, which runs along the superficial surface. If desired, the nerve can be isolated along its inferior border and raised out of the field at this point. Another method for preserving the nerve involves incising the capsule of the submandibular gland inferiorly. The facial vein is also divided, and these two structures are lifted to the inferior border of the mandible, elevating the nerve out of the field.

With the gland exposed and the marginal mandibular nerve protected, removal of the gland begins. The facial artery and vein are identified at the superior edge of the gland, where it lies adjacent to the inferior border of the mandible. There is rarely any consequence to complete ligation of the facial artery and vein. Nonetheless, the authors typically preserve these vessels for benign disease, choosing to ligate the three or four tributaries into the gland instead. The vessels are followed posteriorly and inferiorly to the digastric tendon. Attention is then turned to the anterior aspect of the gland. As the anterior aspect of the gland is elevated, staying close to the gland capsule, the mylohyoid muscle comes into view. The mylohyoid muscle is then mobilized along its posterior border from hyoid bone to the junction with the mandible. A retractor is placed under the mylohyoid muscle, exposing the tissues of the deep floor of the mouth. Any gland attachments to the inferior border of the mandible can be quickly taken down to this same plane without any risk of damage to neural structures.

Once the plane of the floor of the mouth is reached, the lingual nerve is identified as it arcs downward into the field. The submandibular ganglion carrying parasympathetic fibers into the gland is located at the apex of the arc and is divided. The submandibular duct is then identified at the anterosuperior aspect of the gland. It is isolated, teased downward from the floor of the mouth, and ligated and divided.

At this point, the gland is free from all important vascular and neural attachments. The gland can simply be rotated from superior to inferior, lifted off the digastric tendon, and removed. Care should be taken to keep the plane of dissection in the areolar tissue adjacent to the medial capsule of the gland. In doing so, the surgeon will stay superficial to the mylohyoid and hyoglossus muscles and avoid damage to the hyoglossal nerve. If any doubt regarding the position of the hypoglossal nerve exists, it can usually be identified in the deeper tissues just superior to the tendon of the digastric muscle.

In cases of malignancy, or when the submandibular gland is removed as part of a neck dissection, the perifacial lymph nodes along the superior and posterior aspects of the gland should

be removed as well. These are the cases in which damage to the marginal mandibular nerve is most commonly seen. The procedure can be modified by identifying the marginal mandibular nerve, dissecting down onto the inferior border of the mandible, and dividing the facial artery and vein as they cross the mandible. Rather than hugging the superior border of the gland, as in benign disease, the entire gland and the perifacial nodal tissue are removed by this modification.

Surgery of the Parotid Gland

Surgery of the parotid gland is challenging. Even with normal anatomy and normal surrounding parotid tissue, dissection of the branches of the facial nerve requires patience and great attention to detail. When the parenchyma of the gland is altered by an inflammatory process or the course of a nerve is distorted and attenuated by tumor, preservation of the facial nerve fibers can be very difficult. Schedule plenty of time to perform parotidectomies. Be patient, insist on careful hemostasis, and use loupe magnification if available.

Surgical preparation and draping of the patient are important in parotidectomy. The authors prefer to be able to see the entire side of the face during the operation to monitor for facial stimulation. An Io-Ban drape is used on the anterior half of the face for this purpose. For unusual cases, intraoperative facial nerve monitoring can be very helpful.

The operation begins with proper exposure. A modified Blair incision or a modified facelift incision is acceptable for parotidectomy. Skin flaps are elevated in the subplatysmal and sub-SMAS plane. The parotid gland, like the submandibular gland, is encapsulated by the superficial layer of the deep cervical fascia. The plane of flap elevation should be between the platysma/SMAS layer and the gland capsule. The tail of the gland is sharply separated from the sternocleidomastoid muscle. Often electrocautery is used for this portion of the dissection. The external jugular vein is divided. The anterior branch of the greater auricular nerve is divided while preserving the posterior branch if at all possible. The digastric muscle is identified and exposed from the submandibular triangle back to the mastoid tip. The posterior part of the parotid gland is then separated sharply from the ear structures from the zygoma superiorly down to the mastoid tip. Dissection can proceed rapidly down to the level of the lateral aspect of the mastoid tip and to the cartilaginous portion of the ear canal.

At this point in the operation, the main trunk of the facial nerve is identified. There are four or five common ways of identifying the facial nerve. In our opinion, the most reliable is by identification of the tympanomastoid suture. Looking at a skull, the tympanomastoid suture extends medially along the skull base until it ends at the stylomastoid foramen. The stylomastoid foramen is almost always within the suture line and, with the exception of congenital otic deformities, provides a safe and consistent guide to the facial nerve. The second common method uses the position of the tragal pointer cartilage. However, as elsewhere in the body, the shape and size of the cartilage vary. Previous experience has determined that the adage "one centimeter deep and one centimeter inferior to the tragal pointer" is much less reliable than the tympanomastoid suture.

Once the main trunk of the facial nerve is identified, only limited dissection of the parotid gland is completed until the pes anserinus is seen. The majority of neoplasms of the parotid gland lie within the superficial lobe. The direction of dissection is determined by the location of the tumor. For tumors in the lower portion of the gland, the upper division or perhaps one of the buccal branches can be followed anteriorly, maintaining an adequate cuff of parotid tissue superior to the tumor. Once the superior border has been established, the lower branches are subsequently exposed while rotating the tumor from superior to inferior. When palpating parotid glands, preoperatively, one is often tempted to simply perform wide local excision without facial nerve identification. Experience shows that most of these tumors lie directly adjacent to at least one branch of the facial nerve, and rarely, if ever, will the surgeon feel that the nerve has been dissected needlessly.

For deep lobe tumors, malignancies of the parotid gland, and chronic inflammation involving the entire gland, a more complete parotidectomy is necessary. For total parotidectomy, exposure of all of the nerve branches is accomplished. Sharp dissection frees the nerve from the main trunk to the periphery, exposing the underlying deep lobe. The deep lobe is then separated from the masseter muscle anteriorly, the zygoma and mandibular condyle superiorly, and the parapharyngeal structures medially. The superficial temporal artery at its ori-

gin from the internal maxillary artery is adjacent to the deep lobe of the parotid gland. This should be approached and controlled carefully. Impressive arterial bleeding from the internal maxillary artery can occur, and rapid application of clamps to control the bleeding can result in damage to the overlying facial nerve network.

A complete parotidectomy leaves a significant depression in the posterior lateral aspect of the face. There are differing opinions whether this should be camouflaged. Some surgeons feel that filling in the defect makes detection of recurrent disease more difficult. Others feel that in this day of advanced head and neck imaging, recurrence can be picked up easily by imaging techniques. The authors infrequently surgically fill these defects. In patients who are very concerned about appearance, however, camouflage of the defect is used. Options include fat grafts, sternocleidomastoid muscle rotation flaps, acellular dermal grafts, or free vascularized fascial grafts.

REFERENCES

1. Hollingshead WH. Anatomy for surgeons. Vol I. The head and neck. 3rd ed. Philadelphia: Harper & Row; 1982.
2. Mandel ID. Sialochemistry in diseases and clinical situations affecting the salivary glands. Crit Rev Clin Lab Sci 1980;12:321–66.
3. McAnnally T. Parotitis: clinical presentations and management. Post Grad Med 1982;71:87–99.
4. Hilleman MR, Buynak EB, Weibel RE, Stokes J. Live attenuated mumps-virus vaccine. N Engl J Med 1968;278:227–32.
5. Cell J, Rubbies PM, Prat J, et al. Sjögren's syndrome: a stepwise approach to the use of diagnostic tests. Ann Rheum Dis 1992;51:607–10.
6. Rothschild B. Sjögren's syndrome. Compr Ther 1996;22:39–43.
7. Talal N. Sjögren's syndrome New York: Academic Press; 1989.
8. Moutsopoulos HM, Chused TM, Mann DL, et al. Sjögren's syndrome. J Intern Med 1991;230:187.
9. Meilof JF, Bantjes I, De Jong J, et al. The detection of anti-Ro/SS-A and anti-La/SS-B antibodies. A comparison of counterimmunoelectrophoresis with immunoblot, ELISA and RNA-precipitation essays. J Immunol Methods 1990;133:215–26.
10. Chisholm DM, Mason DK. Labial salivary gland biopsy in Sjögren's disease. J Clin Pathol 1968;21:656–60.
11. Voulgarelis M, Dafni UG, Isenburg DA, Moutsopoulos HM. Malignant lymphoma in primary Sjögren's syndrome: a multicenter, retrospective clinical study by the European Concerted Action on Sjögren's Syndrome. Arthritis Rheum 1999;42:1765–72.
12. Papiris SA, Maniati M, Constantopoulos SHS, et al. Lung involvement in primary Sjögren's syndrome is mainly related to the small airway disease. Ann Rheum Dis 1999;58:61–4.
13. Schiodt M, Dodd CL, Greenspan D, et al. Natural history of HIV-associated salivary gland disease. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1992;74:326–31.
14. Hammer JE, Goefield HH. Cervical lymphadenopathy and parotid gland swelling in sarcoidosis: a study of 31 cases. J Am Dent Assoc 1967;74:1224.
15. Duna GF, Galperin C, Hoffman GS. Wegener's granulomatosis. Rheum Dis Clin North Am 1995;21:949–86.
16. Kallenberg CG. Treatment of Wegener's granulomatosis: new horizons? Clin Exp Rheum 1996;14:9–16.
17. Sneller MC, Hoffman GS, Tatar-Williams C, et al. An analysis of forty-two Wegener's granulomatosis patients treated with methotrexate and prednisone. Arthritis Rheum 1995;38:608–13.
18. Brook I. Diagnosis and management of parotitis. Arch Otolaryngol Head Neck Surg 1992;118:469–71.
19. Brown S, Bonifazi DZ. An overview of anorexia and bulimia nervosa, and the impact of eating disorders on the oral cavity. Compendium 1993;14:1596–608.
20. DeVries EJ, Kapadia SB, Johnson JT, et al. Salivary gland lymphoproliferative disease in acquired immune disease. Otolaryngol Head Neck Surg 1988;99:59–62.
21. McAnnally T. Parotitis: clinical presentations and management. Postgrad Med 1982;71:87–99.
22. Pang YT, Raine CH. Acute suppurative parotitis and facial paralysis. J Laryngol Otol 1996;110:91–2.
23. Spratt JS. The etiology and therapy of acute pyogenic parotitis. Surg Gynecol Obstet 1961;112:391–405.
24. Rosseau P. Acute suppurative parotitis. J Am Geriatr Soc 1990;38:897–8.
25. Loughran DH, Smith LG. Review: infectious disorders of the parotid gland. N J Med 1988;85:311–4.
26. Luna MA. Salivary glands. In: Pilch BZ, editor. Head and neck surgical pathology. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 284–349.
27. Johns M, Goldsmith M. Incidence, diagnosis and classification of salivary gland tumors. J Oncol 1989;3:47–56.

28. Califano J, Eisele DW. Benign salivary gland neoplasms. Otolaryngol Clin North Am 1999;32:861–73.
29. Roijer E, Kas K, Behrendt M, et al. Fluorescence in situ hybridization mapping of breakpoints in pleomorphic adenomas with 8q12-13 abnormalities identifies a subgroup of tumors without PLAG1 involvement. Genes Chromosomes Cancer 1999;24: 78–82.
30. Gillenwater A, Hurr K, Wolf P, et al. Microsatellite alterations at chromosome 8q loci in pleomorphic adenoma. Otolaryngol Head Neck Surg 1997;117: 448–52.
31. Li X, Tsuji T, Wen S, et al. Detection of numeric abnormalities of chromosome 17 and p53 deletions by fluorescence in situ hybridization in pleomorphic adenomas and carcinomas in pleomorphic adenoma. Correlation with p53 expression. Cancer 1997;79:2314–9.
32. Seifert G, Michlke A, Hanubrich J, et al. Diseases of the salivary glands: pathology, diagnosis, treatment, and facial nerve surgery. Stuttgart: Georg Thieme; 1986.
33. Batsakis JG, Luna MA. Undifferentiated carcinomas of the salivary glands. Ann Otol Rhinol Laryngol 1991;100:687–90.
34. Auclair PL, Ellis GL, Gnepp DR, et al. Salivary gland neoplasms: general considerations. In: Ellis GL, Auclair PL, Gnepp DR, editors. Surgical pathology of the salivary glands. Philadelphia: WB Saunders; 1991. p. 135.
35. Spiro RH, Spiro JD. Cancer of the salivary glands. In: Myers EN, Suen JY, editors. Cancer of the head and neck. 2nd ed. New York: Churchill Livingstone; 1989. p. 646.
36. Rice DH. Malignant salivary gland neoplasms. Otolaryngol Clin North Am 1999;32:875–86.
37. Land CE, Saku T, Hayashi Y, et al. Incidence of salivary gland tumors among atomic bomb survivors, 1950–1987. Evaluation of radiation-related risks. Radiat Res 1996;146:28–36.
38. Spiro RH. Salivary neoplasms: an overview of a 35-year experience with 2807 patients. Head Neck 1986; 8:177–84.
39. Alvi A, Myers EN, Carru RL. Malignant tumors of the salivary glands. In: Myers EN, Suen JY, editors. Cancer of the head and neck. 3rd ed. Philadelphia: WB Saunders; 1996. p. 525.
40. Douglas JG, Laramore GE, Austin-Seymour M, et al. Treatment of locally advanced adenocystic carcinoma of the head and neck with neutron radiotherapy. Int J Radiat Oncol Biol Phys 2000;46:551–7.
41. Batsakis JG, Pinkston GR, Luna MA, et al. Adenocarcinoma of the oral cavity: a clinico-pathologic study of terminal duct carcinoma. J Laryngol Otol 1983;97:825–35.
42. Evans HL, Batsakis JG. Polymorphous low-grade adenocarcinoma of the minor salivary glands. Cancer 1984;53:935–42.
43. Westra WH. The surgical pathology of salivary gland neoplasms. Otolaryngol Clin North Am 1999;32: 914–43.
44. Vincent SD, Hammond HL, Finkelstein MW. Clinical and therapeutic features of polymorphous low-grade adenocarcinoma. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1994;77:41–7.
45. Spiro RH, Huvos AF, Strong EW. Acinic cell carcinoma of salivary origin: a clinical pathologic study of 67 cases. Cancer 1978;41:924–35.
46. Seifert G, Sobin LH. The World Health Organization's Histological Classification of Salivary Gland Tumors. A commentary on the second edition. Cancer 1992;70:379–85.

CHAPTER 63

Diseases of the Thyroid and Parathyroid Glands

Robert A. Hendrix, MD

THYROID GLAND

The thyroid gland (from Greek *thyroeides*, meaning shield-like) is a brownish-red, highly vascular, ductless gland situated anteriorly in the visceral compartment of the neck at the level of the fifth, sixth, and seventh cervical and first thoracic vertebrae. The normal adult gland weighing 20 to 25 g is slightly larger in women. It enlarges further during puberty, menstruation, and pregnancy. It consists of right and left cone-shaped lobes 5 cm in length, connected by a narrow region of gland termed the isthmus. The base of each lobe lies on a level with the fourth or fifth tracheal ring, with the apex oriented laterally and superiorly to the level of the oblique line of the thyroid cartilage. The pyramidal lobe is a variable tail of thyroid tissue trailing superiorly from a paramedian attachment to the isthmus. A true capsule of fibrous connective tissue contains the parenchyma and sends fine septae between lobules of the thyroid gland.

PHYSIOLOGY

The thyroid gland contains two types of functioning endocrine cells of different origins. The follicular cells secrete L-thyroxine (T_4) and 3,5,3′-triiodo-L-thyronine (T_3), which influence a wide range of metabolic processes. The parafollicular cells or C cells influence calcium metabolism by secretion of calcitonin.

The lumina of follicles of the thyroid contain thyroglobulin as a colloid. This glycoprotein is produced only by follicular cells. Production of active thyroid hormone begins with energy-dependent uptake and transport of extracellular iodide across the follicular cell membrane. Following oxidation of iodide and repeated iodinations of tyrosine residues on newly synthesized thyroglobulin at the cell–colloid interface, T_3 and T_4 are produced and stored in the follicular lumen bound to thyroglobulin.

Liberation of active thyroid hormone requires uptake of colloid droplets from the lumen by pinocytosis and transport through the follicular cell with hydrolysis of thyroglobulin and release of free T_3 and T_4. Thus, normal secretion of thyroid hormone requires not only an adequate rate of hormone synthesis but also the capacity to hydrolyze thyroglobulin to liberate T_4.

The follicles secrete only 20% of serum T_3; dehalogenation of circulating T_4 produces the remainder. The concentration of unbound T_3 is about 10 times greater than unbound T_4 because of different affinities for thyroid-binding globulin and other plasma proteins. Nonetheless, plasma proteins reversibly bind almost all of the serum T_3 and T_4, leaving only 0.3% of T_3 and 0.03% of T_4 (or one-tenth the level of T_3) free to act on receptor sites.

The metabolic effects and the regulation of thyroid hormone depend solely on the concentration of free or unbound T_4 and T_3 in plasma. Approximately one-third of the T_4 is deiodinated to T_3, and *T_3 is three times more potent than T_4*. Thus, T_4 exerts most of its metabolic effects through its conversion to T_3. The weaker binding of T_3 to plasma protein carriers may also contribute to the more rapid onset and shorter duration of its action.

Thyrotropin (thyroid-stimulating hormone [TSH]) is a glycoprotein secreted by basophilic (thyrotropin) cells of the anterior pituitary gland. Thyroid-stimulating hormone mediates suprathyroid regulation of thyroid hormone secretion. It acts on the follicular cell receptor to increase all aspects of follicular cell metabolism, including synthesis of hormone and thyroglobulin and hormone secretion. The thyrotropin receptor belongs to the large and diverse family of G protein–coupled receptors,

which all share a common molecular structure: an extracellular amino terminal, seven transmembrane loops, and an intracytoplasmic carboxyl terminal. Of the subfamily of glycoprotein hormone receptors, binding specificity arises from an especially long amino-terminal extracellular domain. The gene for the thyrotropin receptor is encoded by 10 exons spread over 58 kilobases on chromosome 14. The first 9 exons encode the large extracellular domain, and the tenth exon encodes the carboxyl terminus. The thyrotropin receptor is coupled to a subunit of guanine-nucleotide-binding protein that activates adenylate cyclase and increases intracellular levels of cyclic adenosine monophosphate (cAMP). Follicular cell growth and function are both stimulated by the "second messenger" cAMP. Further, expression of thyroglobulin and thyroid peroxidase genes is indirectly regulated by cAMP. Were a single thyroid epithelial cell to contain a somatic mutation that caused chronic stimulation of the cAMP pathway, that cell would be stimulated to grow and clonal proliferation would result, giving rise to an autonomously functioning thyroid adenoma and eventually to hyperthyroidism. In contrast, blocking autoantibodies to the thyrotropin receptor would inhibit the molecular cascade and cause hypothyroidism. Similarly, hypothyroidism could arise from an inherently defective thyrotropin receptor.

With higher thyrotropin concentrations, coupling of the thyrotropin receptor to yet another G protein leads to activation of phospholipase C. Further, growth factor I, transforming growth factor ß, platelet-derived growth factor, epidermal growth factor, fibroblast growth factor, and cytokines stimulate growth of thyroid follicular cells, mainly acting through the protein tyrosine kinase signal transduction pathway.[1]

Hypothalamic secretion of thyrotropin-releasing hormone (TRH) controls TSH secretion. Free thyroid hormones in serum exert negative feedback at the level of the pituitary by inhibiting TSH secretion and antagonizing TRH by decreasing the number of receptors on the thyrotropin cell. At the level of the thyroid gland, iodine depletion enhances the responsiveness to TSH, whereas iodine enrichment inhibits the TSH response. The biosynthesis and secretion of T_3 and T_4 can be affected, not only by deficiency of dietary iodide but also by drugs such as lithium, propylthiouracil, sulfonamides, phenytoin, and nitrophenols.

Deiodination of approximately 40% of T_4 produces inactive T_3 or reverse T_3 (rT_3). Certain conditions (eg, starvation) increase the T_4 conversion to rT_3, thereby requiring occasional measurement of rT_3.

Calcitonin is a 32 amino-acid polypeptide transcribed from a locus on chromosome 11p, which is tightly linked with the gene for parathyroid hormone (PTH), parathormone. Three peptides are encoded by the calcitonin gene: calcitonin, a 21 amino-acid carboxy-terminal flanking peptide (katacalcin), and a calcitonin gene–related peptide. The action of calcitonin is independent of PTH and vitamin D. The main endocrine effect is to decrease the number and activity of osteoclasts, thereby reducing bone resorption, which is the basis for its use in the treatment of Paget's disease. Extrathyroid calcitonin synthesized in the pancreas, gastrointestinal tract, pituitary, and brain probably acts as a locally inhibitory neurotransmitter. Calcitonin is present in high circulating levels in the fetus, probably owing to the high fetal serum levels of calcium. Apparently, calcitonin affects fetal bone metabolism and skeletal growth and remains at high levels in cord blood as well as in young children. Levels are very low in older children and adults, in whom it has little consequence to metabolism. Even with very high levels in patients with medullary carcinoma of the thyroid, hypercalcemia is not observed.[2]

Laboratory Assessment of Thyroid Function

Serum Tests Related to Thyroid Hormone The most widely used serum assays of thyroid hormone are the T_4 radioimmunoassay (RIA), T_3 uptake of resin, and free T_4. Thyroxine RIA, as a measure of total serum thyroxine, has a normal range of 5 to 13 mg/dL.

States of altered thyroxine binding affect available thyroid hormone. Adequate assessment requires T_4 RIA measurement in combination with the T_3 uptake of resin as an indirect measure of T_4-binding protein. The T_3 uptake of resin test is performed by incubating serum with radioactive, labeled T_4 or T_3 in the presence of insoluble particles of resin or charcoal that absorb any labeled hormone not bound by serum protein. After removal of the serum, the percentage of radioactive hormone not taken up by serum proteins is measured in the precipitate. The result varies inversely with the concen-

tration of unoccupied sites on serum protein carriers of thyroid hormone. Triiodothyronine uptake of resin may vary in certain conditions such as pregnancy, when the T_4 is increased and the T_3 uptake is low because of increased thyroid-binding globulin.

The "free" T_4 determination (FT_4) measures the metabolically effective fraction of circulating T_4. Free T_4 provides the best assay of thyroid hormones in regard to function because it is not affected by thyroid-binding globulin. Free T_4 may be estimated from the more commonly available free T_4 index, equal to the product of the T_4 and the T_3 uptake of resin. This calculation is used to correct for abnormalities of T_4 binding. Nonetheless, this is unnecessary if the free T_4 test is available.

Thyroxine-binding globulin (TBG) RIA is a direct and specific test for T_4-binding abnormalities and is unaffected by changes in other serum proteins.

Triiodothyronine assessments made by RIA can be valuable in the diagnosis of thyrotoxicosis with normal T_4 values (T_3 thyrotoxicosis) and in some cases of toxic nodular goiter.

The measurement of serum TSH concentration by RIA is useful to evaluate hyperfunction of the thyroid gland. Most evaluations of thyroid function should start with a sensitive thyrotropin assay. In a patient with little clinical probability of thyroid dysfunction, a normal TSH screen requires no further testing. If TSH is elevated, an FT_4 and possibly a thyroid antibody should be performed to evaluate for hypothyroidism. If TSH is low, then an FT_4 and possibly a T_3 should be performed to evaluate for hyperthyroidism. However, in elderly patients, a low TSH assay may not be associated with hyperthyroidism. This approach to testing provides an efficient way to assess thyroid hormone function (Table 63–1).[3]

Direct Tests of Function The most commonly used direct test of gland function is the thyroid radioactive iodine uptake (RAIU). Historically, the RAIU involved in vivo administration of iodine 131 (^{131}I), but ^{123}I has become the agent of choice because of lower radiation dosages. It is administered orally in capsule form 24 hours before measurement of thyroid accumulation of isotope. The RAIU is also used in the thyroid suppression test, in which the RAIU is repeated after 7 days of daily administration of T_3. A decrease in the RAIU is evidence of the presence of thyroid suppression. This is of value in the assessment of glandular hyperfunction. Similarly, the thyrotropin stimulation test (RAIU performed following TSH administration) is useful to differentiate primary thyroid insufficiency from thyroid hypofunction caused by pituitary hypofunction.

Imaging of the Thyroid Gland

Closely related to the RAIU is the thyroid scintiscan by gamma camera 24 hours after administration of ^{123}I. To evaluate metastatic thyroid cancer or a substernal thyroid, ^{131}I should be used to overcome photon attenuation by bone. Several radionuclides have been studied that concentrate in thyroid tissue, including gallium 67, thallium 201, and technetium 99m-methoxy-isobutyl isonitrile (^{99m}TCMIBI) also called 99m TC-sestamibi and ^{99m}Tc tetrofosmin. Nonetheless, it is not generally possible to differentiate benign from malignant thyroid nodules by this modality alone.[4]

TABLE 63–1. Summary of Interpretations to be Drawn from the Results of a "Serum Thyroid Panel"

Results of Serum Testing	*Interpretation of Thyroid Function*
Normal TSH and no clinical suspicion of thyroid disease	Normal
Normal TSH and normal FT_4	Normal
Elevated TSH and low FT_4	Primary hypothyroidism
Elevated TSH and elevated FT_4	Secondary (pituitary) hyperthyroidism or inappropriate thyroxine ingestion
Low TSH and high FT_4	Primary hyperthyroidism
Low TSH and low FT_4	Pituitary insufficiency

TSH = thyroid-stimulating hormone; FT_4 = free thyroxine.

The thyroid also concentrates monovalent ions such as sodium ^{99m}Tc pertechnetate. Because ^{99m}Tc has a short half-life and is not organically bound, its presence in the thyroid is only transient. Thus, if radioactive iodine uptake is blocked or contraindicated, ^{99m}Tc, at a dose of approximately 5 mCi, may be administered intravenously, with a scan performed at 20 minutes. Either technique adequately assesses thyroid enlargement or cold nodules. Increased sensitivity to the planar ^{99m}Tc scan has recently been reported using single-photon emission computed tomography (SPECT) for preoperative localization of occult medullary carcinoma of the thyroid gland.[5]

In traditional management following total thyroidectomy for thyroid cancer, thyroid hormone supplement is withheld to render the patient hypothyroid with a consequential increase in TSH level. This stimulates residual thyroid tissue and functioning metastatic cells to take up radioactive iodine for diagnostic scanning or ablation. Recombinant thyrotropin has been shown effective, although not as sensitive as thyroid withdrawal, in evaluation of patients with thyroid cancer. Recombinant human thyrotropin appears more sensitive than thyroid hormone withdrawal in causing an elevation of quantitative thyroglobulin levels in metastatic thyroid cancer. Instead of discontinuing thyroid hormone, recombinant human thyrotropin can be given on two consecutive days with a scanning dose administered on the third day. On the fifth day, the scan and quantitative thyroglobulin are obtained. Recombinant human thyrotropin is still under investigation for use in the treatment of metastatic thyroid cancer but is approved for use in treatment of hypopituitarism and in patients with so much functioning residual thyroid tissue as to prevent elevation of TSH. The advantage of recombinant human thyrotropin in this setting prevents the long period of hypothyroidism that can be very symptomatic for the patient.[6]

B-mode (two-dimensional) ultrasonography is a useful adjunctive tool for the evaluation of thyroid masses.[7] This method is particularly recommended in children or pregnant women, in whom radioactive isotopes are undesirable. The chief role of this modality is to distinguish between cystic and solid lesions, particularly in goiter. However, ultrasonography and thyroid scan are complementary tools for evaluation of thyroid nodules, and some argue that both are necessary for adequate assessment.[8]

Computed tomography (CT) and magnetic resonance imaging are useful to evaluate thyroid anatomy and relationships with adjacent structures, particularly in preoperative planning. Computed tomographic scan with iodinated contrast interferes with performance of the radioactive iodide uptake and scan for up to 6 weeks. Thyroid scans and RAIU should therefore be performed before CT with contrast.

Positron emission tomography (PET) is a promising modality for investigation of thyroid lesions. Measuring uptake of [18^{F}]-2-deoxy-2-fluoro-D-glucose (FDG) in thyroid nodules, increased uptake was seen in malignant lesions, distinguishing benign from malignant disease in each of 19 cases.[9] Further evaluation of this imaging modality in larger series provides an important future research direction.

Fine-Needle Aspiration Biopsy

Fine-needle aspiration (FNA) biopsy has become the diagnostic study of choice for all solitary thyroid nodules, for nodules within multinodular goiters that grow rapidly or steadily or have a worrisome texture, or for diffusely enlarged thyroids with localized nodules. The goal set by the Papanicolao Society of Cytopathology for laboratories is to aim for a standard for FNA diagnosis of thyroid nodules of a false-negative rate of less than 2% and a false-positive rate of less than 3%.[6] Fine-needle aspiration is currently performed by 88% of otolaryngologists for evaluation of thyroid masses and by 12% for evaluation of parathyroid masses.[10] This office procedure requires a 10 mL syringe with a fine needle (21 or 22 gauge) passed through alcohol-prepared skin to a palpable nodule or area of interest in the thyroid during continuous aspiration. The specimen may be smeared on a slide and sprayed immediately with fixative or submitted in appropriate media (eg, Carbowax) for later examination. The validity of this technique depends largely on the expertise and experience of the individual cytopathologist. Reportedly, sensitivity ranges from 75 to 93.5% and specificity from 75 to 100%.[11,12] However, malignancy can be present in up to 15% of indeterminate specimens and 11% of negative specimens. Comparison of diagnostic signs with biopsy-proven malignancy

shows an association with tumor size (> 4 cm), irregular borders, fixed lesions, or heterogeneity on ultrasonography. Close follow-up with repeat FNA in 6 months is recommended in patients with any of these features despite negative FNA.[13] Beyond well-differentiated thyroid tumors, FNA enables earlier diagnosis of rare but highly malignant neoplasms such as teratoma of the thyroid, characterized by small, round cells with neuroepithelial differentiation and metastases to lymph nodes.[14]

The gross appearance of aspirated fluid often yields diagnostic information: crystal-clear fluid is suggestive of a parathyroid cyst; yellow fluid is suggestive of a transudate; chocolate, green, or turbid fluid suggests degeneration; and bloody fluid from a palpable mass can suggest a rapidly growing vascularized tumor or aspiration of a vessel. Complications of this procedure include hematoma formation and recurrent laryngeal nerve injury, and the risk of infection is low.

Embryology

The parafollicular cells of the thyroid gland are part of the amine precursor uptake and decarboxylation (APUD) cell population, which is derived from neuroectodermal or neural crest cells.

The primordium of thyroid follicles arises from endoderm in the floor of the pharynx between the first and second pharyngeal pouches. This area evaginates to form a median diverticulum by the end of the fourth week. It grows caudally as a tubular thyroglossal duct through the midline of the primordium of the hyoid bone, eventually to bifurcate and develop into the isthmus, pyramidal lobe, and lateral lobes of the thyroid gland. The foramen cecum marks the site of origin of the thyroglossal duct, often identified in examination of the dorsum of the tongue. Maldevelopment of the thyroglossal duct can result in midline cervical ectopic thyroid or even lingual thyroid. Lingual thyroid is an irregular base-of-tongue mass that may interfere with aerodigestive function as it increases in size. Remnants of the thyroglossal duct that persist can give rise to midline cysts without sex predilection and usually prior to the third decade. Because of the developmental course of the thyroid gland, effective surgical extirpation requires complete excision of the cyst and any tract up to the foramen cecum including resection of 1 to 2 cm of the central portion of the hyoid bone (as described by Sistrunk in 1920). Although anomalies of the thyroid-hyoid apparatus are rare, fusion of the thyrohyoid interval (thyroid cartilage to hyoid bone) has been reported and can render the standard Sistrunk procedure impossible.[15]

The fetal thyroid begins to collect and organify iodine by about 10 weeks gestation. Adequate fetal thyroid function becomes necessary because maternal-fetal transfer of T_3 and T_4 is minimal. By the end of the first trimester, colloid follicles are visible, and both T_4 and TSH are detectable in the fetal blood. Fetal thyroid function increases in the second trimester associated with increasing thyroid secretion and TBG. Further, an increase in fetal TSH occurs as the fetal hypothalamus matures and produces TRH. Maternal TRH crosses the placenta and may contribute to the fetal pituitary activity, but, otherwise, the fetal pituitary-thyroid axis is a functional unit distinct from the mother.

Disorders of the Thyroid Glands

Goiter Goiter (from Latin *guttur*, meaning throat) is an enlargement of the thyroid gland resulting in swelling in the front of the neck. The term struma (from Latin *struere*, to build) is occasionally applied to describe this condition. A diffusely enlarged thyroid gland becomes palpable when the volume of the gland is doubled, and a visible goiter is usually at least three times the normal thyroid mass of 20 g.

Simple (Nontoxic) Goiter During puberty and pregnancy, the thyroid gland normally undergoes a diffuse enlargement. This is related in part to increased estrogens and subsequent increase in TBG. Colloid goiter can thus arise and is typified histologically by large colloid spaces and flattened follicular epithelium. This condition is usually self-limited and rarely requires treatment.

The term endemic goiter is applicable when 10% or more of the population has generalized or localized thyroid enlargement. It generally reflects a dietary deficiency of iodide particular to a geographic region, causing insufficient thyroid hormone secretion. Five percent of the world's population have goiter or associated disorders (cretinism, impairment of growth and mental development, etc). Of these, 75% live in less developed countries, where iodine deficiency is prevalent.[16] Twenty-five percent of goiters occur in more devel-

oped countries. These so-called "sporadic goiters" arise largely from iodine-sufficient conditions including autoimmune thyroiditis, hypo- or hyperthyroidism, and thyroid carcinoma. Histologically, endemic goiter is identical with colloid goiter.

Nodular Goiter The cause of nodular goiter is poorly understood but may be related to varying levels of TSH over a lifetime. The percentage of individuals with nodular goiter increases with advancing age. Multinodular goiter has a prevalence of 4% in the United States (with a female-to-male ratio of 6.4 to 1.5). Histologically normal parenchyma is interspersed between nodules of varying size and consistency. Nodules smaller than 2 cm are rarely noticed by patients, although with quick growth or painful hemorrhage into a cyst, the patient may seek attention. Large multinodular goiters are usually asymptomatic. They can, however, cause compression of neck structures, resulting in dysphagia, cough, or respiratory distress or a feeling of constriction in the throat. This may necessitate subtotal or total bilateral thyroidectomy. Overall, the recurrence rate for multinodular goiter treated by nontotal bilateral thyroidectomy is 21%, with the average time to recurrence at 8.75 years.[17]

The molecular biology of follicular cell growth in thyroid nodules is incompletely understood. However, distinct molecular pathogenetic mechanisms exist in hyperfunctioning and nonfunctioning follicular thyroid adenomas. Activating or "gain of function" mutations of the TSH receptor gene (*TShR*) or Gs alpha genes occur with hyperfunctioning thyroid adenomas, in which case, activation of the cAMP cascade leads to proliferation but maintains differentiation. Nonfunctioning follicular adenomas probably arise from oncogenes other than the *TShR* and Gs alpha genes.[18]

Medical therapy is indicated if a goiter is a cosmetic problem or if it becomes symptomatic because of compression and displacement of vital structures. Exogenous thyroid hormone is given to suppress TSH secretion. Radioiodine is usually effective for more extensive cases.

Substernal goiter can develop from downward growth of the thyroid gland into the superior mediastinum. This can potentially cause serious symptoms from compression of tracheal or venous and arterial structures. Substernal goiter not responsive to thyroid hormone suppression is generally managed by thyroidectomy, usually through a standard cervical incision.

Thyroid cancer has reportedly been found in 4 to 17% of patients with multinodular goiter; however, this likely represents bias in selection of patients for operation. The incidence of carcinoma is probably closer to 0.2%. For multinodular goiters with behavior or findings suggestive of malignancy, evaluation by FNA biopsy cytology may be useful. In the event of sudden, continued growth of a hard nodule, whole-goiter progression, symptoms related to compression or displacement, or severe cosmetic deformity despite medical treatment, a subtotal thyroidectomy is appropriate.

Hypothyroidism Hypothyroidism results from insufficient thyroid hormone secretion. In adults, this appears as myxedema characterized by weakness, lethargy, nonpitting edema, coarse dry skin, and possibly hearing loss. Hypothyroidism is present in approximately 1 in 5,000 neonates, although cretinism may not become evident until after several months of extrauterine life. This may arise from inadequate iodide intake in the maternal diet, thyroid agenesis, or administration of propylthiouracil or radioactive iodide during pregnancy. In infants, hypothyroidism causes cretinism with lethargy, stunted growth, mental retardation, and hearing loss. Hypothyroidism during pregnancy apparently can harm the neuropsychological development of the fetus. Mild maternal hypothyroidism alone has been associated with lower IQ scores in offspring. For women of childbearing age who are hypothyroid or who undergo thyroidectomy, it is important that they maintain a euthyroid state should they become pregnant. This is an exceedingly important responsibility of the thyroid surgeon to provide proper patient education in this regard.[6]

Hypothyroidism is not in itself an indication for thyroidectomy, but it can arise from surgical or medical intervention. In particular, surgical extirpation of a lingual thyroid or a thyroglossal duct cyst without preoperative thyroid scan may unexpectedly remove all functioning thyroid tissue. Although congenital hypothyroidism can affect hearing, no auditory effects of short-term (8 weeks) T_4 depletion are detectable.[19] Patients at high risk for myxedema coma require effective teaching regarding the signs and symptoms of hypothyroidism as well as annual thyroid function tests (including serum TSH).[20]

Careful monitoring of patients on thyroid replacement is prudent. Specifically, subclinical hyperthyroidism is associated with reduced bone density in the pelvis and axial skeleton in postmenopausal women. A three-fold increase in the incidence of atrial fibrillation is seen in elderly with hyperthyroid states. Therefore, in the absence of metastatic thyroid cancer, TSH suppression only to low normal is recommended. Given the deleterious effects of iatrogenic hyperthyroidism, treatment of cold nodules and euthyroid multinodular goiter with TSH suppression is generally out of favor.[6]

Thyrotoxicosis Thyrotoxicosis is the clinical and physiologic response of the tissues to an excess supply of active thyroid hormone. Causes of this syndrome fall into one of three main categories. The most important category includes diseases in which the gland sustains overproduction of thyroid hormone. This can arise in the presence of an unregulated thyroid stimulator of extrapituitary origin, as in Graves' disease or in patients who develop hyperthyroidism in association with Hashimoto's thyroiditis.[8] Rarely, excessive secretion of TSH from a pituitary tumor causes thyrotoxicosis.[21]

The second category involves the development of thyrotoxic states secondary to subacute (de Quervain's) thyroiditis or in the syndrome of chronic thyroiditis with spontaneously resolving thyrotoxicosis. In the presence of an inflammatory process, preformed hormone leaks from the gland in excess. The gland, however, fails to produce new hormone because of the suppression of TSH by feedback inhibition; this process is therefore self-limited. This often results in a transient insufficiency of thyroid hormones.

The third category is unusual and arises when a source of excess hormone arises from other than the thyroid gland, as in the case of a community-acquired outbreak of transient thyrotoxicosis arising from ingestion of ground beef contaminated with bovine thyroid gland.[22] Rarely, functioning metastatic follicular carcinoma can account for thyrotoxicosis. Up to 25% of patients with multinodular goiter may develop Jod-Basedow disease, which is thyrotoxicosis in response to the use of exogenous iodine (eg, radiographic contrast dye or iodide expectorants). This disorder is self-limited and is treated principally by discontinuing iodide use.

Hyperthyroidism If sustained hyperfunction of the thyroid gland leads to thyrotoxicosis, the condition is properly termed hyperthyroidism. Thus, not all thyrotoxic states are associated with hyperthyroidism. In true hyperthyroidism, an increased radioactive iodine uptake is found. In contrast, nonhyperthyroid thyrotoxic states have a decreased RAIU. Furthermore, in conditions with sustained hyperfunction of the thyroid gland, the use of treatments to decrease hormone synthesis (antithyroid agents, surgery, or radioiodine) is appropriate. In the absence of a true hyperthyroid state, however, these are inappropriate and ineffective methods for treatment of thyrotoxicosis. Young patients more commonly develop thyrotoxicosis from Graves' disease, whereas older patients develop toxic nodular goiter.

Graves' Disease Graves' disease is a relatively common disorder of unknown origin consisting of a triad of hyperthyroidism with diffuse goiter, ophthalmopathy, and dermopathy. All aspects of the triad need not appear together. Graves' disease is more frequent in women and has a definite familial predisposition. The basic disorder is a disruption of homeostatic mechanisms caused by the presence of an abnormal thyroid stimulator in the plasma. This long-acting thyroid stimulator is present in the IgG class of immunoglobulins. Long-acting thyroid stimulator, however, does not appear to be the antibody responsible for ophthalmopathy. Thus, Graves' disease appears to be an autoimmune process in which an antibody to the thyroid TSH receptor is produced by the patient's own lymphocytes. Subtle variation in immunoregulatory genes may be associated with autoimmune disease states. A complex interaction between genetic and environmental factors influences susceptibility to autoimmune thyroid diseases, Graves' disease and autoimmune hypothyroidism. Human leukocyte antigen and cytotoxic T-lymphocyte–associated-4 regions appear to influence susceptibility to disease to some degree, although other genes remain unknown. As crucial factors in the regulation of inflammatory and immune responses, cytokines are also potential candidate genes for autoimmune thyroid disease.[23]

The symptoms of Graves' disease include nervousness and tremors with mood swings, tachycardia, hypertension, diarrhea, insomnia, and heat intolerance, particularly in younger patients. In older patients, weakness, dyspnea, and even cardiac

failure can arise. Thyrotoxicosis may lead to degeneration of skeletal muscle fibers, enlarged heart, fatty infiltration and fibrosis of the liver, and bony decalcification.

Ocular signs include a characteristic stare with infrequent blinking and lid lag owing to sympathetic overstimulation. In contrast to the infiltrative ophthalmopathy, these findings subside when thyrotoxicosis is treated. The ophthalmopathy of Graves' disease is caused by an inflammatory infiltrate of the orbit that spares the globe. Edema and inflammation of the muscles, with infiltration by lymphocytes, mast cells, and plasma cells, account for the increased volume of orbital contents and proptosis. Eventually, muscle fibers degenerate and fibrose.

The dermopathy of Graves' disease is characterized by a thickened dermis caused by a lymphocytic infiltration.

In the triad of Graves' disease, diffuse toxic goiter is the most common presentation. In a severe case, Graves' disease is a straightforward diagnosis. Increased radioactive iodide uptake, serum T_4 radio immuno assay (RIA), and T_3 resin uptake serve more as baseline than diagnostic aids.

The treatment of hyperthyroidism in Graves' disease has two approaches, both of which are directed to limit thyroid hormone production by the gland. The first approach uses antithyroid agents to blockade hormone synthesis chemically. The second approach is ablation of thyroid tissue either by surgery or by means of radioactive iodine.

A trial of antithyroid therapy is desirable in children and patients under 40, including pregnant women, but it may also be employed in older patients. This includes propylthiouracil or methimaxole. Propylthiouracil is preferred in pregnancy because it has less transference across the placenta. Furthermore, the drug is not excreted in breast milk. Iodide inhibits release of hormones from the thyroid gland in hyperfunction. This is useful in treatment of thyrotoxic crisis, although the response is often incomplete and transient. Large doses of glucocorticoids such as dexamethasone help to reduce the serum T_4 concentration and can contribute to management of thyrotoxicosis as well as oculopathy.

In the past, a major cause of mortality in Graves' disease was thyroid storm, characterized by hyperthermia, tachycardia, hypertension, profuse sweating, intense irritability, extreme anxiety, and eventual prostration with irreversible hypotension and death.[20] This phenomenon represents an acute adrenergic crisis caused by thyroid hormone augmentation of the effects of catecholamines, particularly by induction of additional myocardial catecholamine receptors. Treatment with sympatholytic drugs including reserpine, guanethidine, and β-adrenergic blockers (propranolol), as well as oxygen, intravenous glucose, iodide, and adrenal steroids, has brought this serious problem under control for the surgeon.

Surgical ablation of the thyroid gland in Graves' disease is recommended in patients who have had a relapse or recurrence after drug therapy, in patients with a large goiter or drug toxicity, and in patients who fail to follow a medical regimen or to return for periodic examinations.

Although a radioiodide has no known carcinogenic effects as a β-emitter in adults treated for hyperthyroidism, many physicians prefer to reserve therapy for patients over 30 years of age who are unlikely to bear children. Hypothyroidism occurs with insidious onset in 40 to 70% during the 10-year period following treatment with ^{131}I. Patients treated for Graves' disease are particularly vulnerable to serious complications of myxedema. Some clinicians recommend treatment with large doses of ^{131}I followed by permanent physiologic replacement with exogenous thyroid hormone. In younger patients or in those for whom radioiodide is contraindicated, a subtotal thyroidectomy can be performed as described in a later section of this chapter. Surgical patients must attain a stable euthyroid state through antithyroid agents before any operation.

Toxic Adenoma Thyrotoxicosis caused by thyroid hormone secreted by an autonomous follicular adenoma is termed toxic adenoma. This can follow a long period of euthyroid function in which suppression of normal gland balances the hyperfunctioning adenoma. In regions with adequate dietary iodine, hyperthyroidism owing to Graves' disease is 50 times more prevalent than hyperthyroidism owing to autonomously functioning thyroid adenoma. However, hyperthyroidism arises from different mechanisms in iodine-deficient areas. For instance, in areas of Denmark, hyperthyroidism is attributable to Graves' disease in only 39%, with autonomously functioning thyroid adenoma responsible for 10% and toxic multinodular goiter (discussed below) accounting for 48%. It appears that iodine deficiency

promotes the development of autonomously functioning thyroid nodules.[1]

Radioiodide is generally an effective treatment for toxic adenoma, often sparing the normal remaining gland, which has low uptake of iodide. Thyroid lobectomy is appropriate for patients at risk, as discussed later in this chapter in relation to the solitary nodule.

Toxic Multinodular Goiter Toxic multinodular goiter (Plummer's disease) is a disease of aging that arises in a simple (nontoxic) goiter of long standing. The development of thyrotoxicosis with slow progression of multinodular goiter is usually caused by an autonomous functioning adenoma. Approximately 10% of solitary palpable nodules have autonomous function. In multinodular goiters over 100 g, two-thirds have autonomous secretion of thyroid hormone. Treatment is primarily medical, consisting of antithyroid drugs (propylthiouracil or methimaxole), sympatholytic therapy, and radioiodide. In the absence of symptoms caused by mass effects of goiter, surgery is reserved for toxicity not controlled medically.

Thyroiditis Rarely, the thyroid gland may have a bacterial infection. Of much greater significance are the inflammatory diseases of the thyroid that do not become manifest in the acute phase. Among these is subacute (de Quervain's granulomatous) thyroiditis, a febrile disease in which the thyroid becomes acutely congested, edematous, and mildly tender, often after an upper respiratory infection. The disease is characterized by repeated remissions and exacerbations over several months without hypothyroidism or evidence of stimulating antibodies. Thyroid microsomal and thyroglobulin antibodies, however, do appear in low titers in subacute thyroiditis. Histologically, an inflammatory infiltrate with large, multinucleated giant cells containing vesicles of colloid suggests a foreign body reaction. This condition is not treated by a surgical procedure.

Hashimoto's disease (struma lymphomatosa) is the most common type of thyroiditis. This goitrous form of autoimmune thyroiditis and its variant, lymphocytic thyroiditis, appear mostly in women as symmetric, rubbery enlargement of the gland. Although microsomal and thyroglobulin antibodies are present, tissue damage appears to be cell mediated. T cells from patients with autoimmune thyroiditis react with iodinated thyroglobulin, but with the addition of iodine, they react with the noniodinated form. Addition of iodine to thyroglobulin generates or exposes epitopes. Further, thyroglobulin-reactive autoantibodies exhibit proteolytic activity on thyroglobulin.[24] Histologically, a lymphocytic infiltrate tending to form lymphoid follicles with plasma cells is seen. Follicles are progressively compressed and distorted into fused masses of cells without a colloid space. Follicular cells may eventually be transformed into Hürthle cells, which appear as nests of large pink cells with abundant granular cytoplasm packed with mitochondria. Although this condition is associated with mild thyrotoxicosis in the early stages because of release of stored hormone from damaged follicles, most patients have depressed thyroid function (RAIU low to normal). In the later and chronic stage, hypothyroidism commonly develops. Fine-needle aspiration biopsy cytology can assist in the diagnosis. Open biopsy or procedures to alleviate compression of the trachea or esophagus are sometimes necessary.

Riedel's struma (invasive fibrous thyroiditis) is a rare presentation of thyroiditis characterized by irreversible, profound hypothyroidism and a stony hard, irregular, fibrous gland. Symptoms can include cough, dyspnea, and difficulty in swallowing. Replacement thyroid hormone is the principal treatment.

Management of Thyroid Neoplasia

Solitary Nodule of the Thyroid The most common indication for thyroidectomy (50% of cases) in the United States is the presence of a solitary thyroid nodule, defined as a discrete mass, greater than or equal to 1 cm in diameter, discovered by palpation of a thyroid gland otherwise of normal size and consistency. Solitary nodules may occur in up to 4% of the population and are most often found in patients 30 to 50 years of age, with a four-fold increase in frequency in women. With ultrasonography, nodules occur in up to 50% of individuals over 50 years of age. It is important to distinguish clinically between the solitary nodule and a multinodular thyroid gland because the latter has a low incidence of malignancy. One series reports spontaneous disappearance as the most common outcome for an untreated, long-standing nodule (38.3%); on the other hand, 13% of nodules in this series were enlarging. Needle biopsy demonstrated malignancy in 26.3% of enlarging nodules, compared with 6.4%

of nodules that were of stable size.[25] In another series, interval examination at 5 years of patients with incidental findings of abnormalities on thyroid ultrasonography revealed that 12% had new nodules (all benign), 12% had grown but all were benign, 24% were diminished or disappeared, and one had a benign adenomatous nodule. They concluded that incidental ultrasonographic abnormalities of the thyroid are generally benign and clinically inconsequential.[26]

Increased risk of malignancy is associated with prior neck irradiation, and a thyroid nodule with enlarged lymph nodes or fixation of nodule to overlying strap muscles suggests malignancy.[27] The most common cause of solitary thyroid nodules in children and adolescents is follicular adenoma, although malignancy is reported in 25.5% of cold nodules in children.[28] A neck mass is the most common presenting symptom of papillary or follicular thyroid carcinoma in patients under 20 years of age.[29]

The evaluation of patients with thyroid nodules should include an assessment of laryngeal motor function, serum thyroid hormone levels, radioactive iodide uptake, and scintiscan. If a hot nodule is demonstrated by ^{99m}Tc scanning, reexamination with radioiodide is recommended. Ultrasonography to differentiate between cystic and solid thyroid nodules is accurate in 82 to 95% of cases. Fine-needle aspiration biopsy cytology often provides an accurate diagnosis and is used by some practitioners to treat cystic lesions. The reported incidence of carcinoma in solitary nodules ranges from 1.3 to 20.3%.

The series with the highest yield of carcinomas on thyroidectomy had stringent selection of patients based on risk factors. Thus, using known risk factors, patients with solitary nodules at increased risk for malignancy can be identified and selected for surgical procedures.[30] Increased risk is found in the following situations:

1. Thyroid nodules associated with vocal cord paralysis (infrequently, benign nodules can cause paresis or paralysis by compression, which is usually alleviated by excision of the mass)[31]
2. Solitary nodules in men
3. Solitary nodules in patients under 20 years old or over 50 years old
4. Solitary nodules in patients with a history of head and neck irradiation
5. Palpable lower or midcervical lymph nodes and apparently normal thyroid glands in young patients
6. Functioning or cold nodules on scintiscan or radionuclide uptake outside the thyroid gland
7. Solid or partially cystic lesions by ultrasonography
8. Recurrent cystic lesions following needle aspiration
9. Abnormal cytologic findings on needle aspiration biopsy
10. Solitary nodules that fail to disappear in response to thyroid hormone suppression therapy

In patients with chronic thyroid disease, high-risk patients who are candidates for surgery may be identified by the following criteria:

1. Multinodular goiter containing cold nodules that enlarge in response to suppression with thyroid hormone
2. Rapidly enlarging nodules in chronic goiters with or without suppression
3. Goiters with vocal cord paralysis
4. A history of head and neck irradiation
5. A history of neoplasia of the thyroid gland

Thyroid adenomas are benign, encapsulated neoplasms with a glandular cellular structure. These lesions may be classified as follicular adenomas, Hürthle cell adenomas, and papillary adenomas. If a "hot" or hyperfunctioning adenoma is demonstrated by technetium scan and this finding persists on a thyroid suppression test, this lesion can be treated with radioactive iodide. Benign teratoma can also occur in the thyroid gland.

Malignant Lesions

Malignant tumors of the thyroid include papillary adenocarcinoma, follicular adenocarcinoma, Hürthle cell carcinoma, medullary carcinoma, and undifferentiated carcinomas (small cell carcinoma and giant cell carcinoma). Various miscellaneous malignant lesions include lymphoma, sarcoma, and teratoma.

Papillary Adenocarcinoma Thyroid carcinomas comprise approximately 1% of all malignant tumors in the United States. Of these, 60 to 70% are papillary carcinomas. These account for 80 to 90% of radiation-induced thyroid carcinomas. The demonstration of G protein or thyrotropin receptor gene mutations in the presence of a synergistic,

activated *RAS* gene supports the multistep model of thyroid carcinogenesis in differentiated carcinoma.[1] Of children aged 10 to 15 years exposed to fallout in the Chernobyl accident, rearrangements of the *RET* oncogene were found in 44%. The majority of *RET* rearrangements identified arose from inversion of part of chromosome 10 involving papillary thyroid cancer genes 1 and 3 (*PTC1* and *PTC3*). *PTC1* rearrangements were associated with classic papillary carcinoma and histologically diffuse sclerosing variants, whereas *PTC3* apparently gave rise to tumors with a solid/follicular morphology.[32] Dietary iodine in childhood is also regarded as an etiologic factor. Although usually sporadic, papillary thyroid carcinoma can occur in familial form (fPTC) and is sometimes associated with nodular thyroid disease. Studies of the fPTC susceptibility gene by genetic linkage and sequence analysis reveal *MET* as the proto-oncogene of isolated fPTC.[33]

Approximately 25% of papillary carcinomas are occult, discovered incidentally at surgery, and of questionable clinical significance. The peak incidence is in the third and fourth decades, with a three-fold increase in frequency in women. Study of a large population in Taiwan with papillary thyroid carcinoma over a 15-year period found the mean age of onset at 40.4 ± 14.6 years and a 5-year survival probability of 0.9817%. The survival was improved over time, apparently owing to more aggressive surgical and medical management and, especially, earlier detection of disease.[34]

One variant of papillary carcinoma, referred to as Warthin-like tumor of the thyroid gland, has been described. These neoplasms are characterized by prominent T- and B-cell lymphocytic infiltration in stalks of papillae and oxyphilic epithelial metaplasia suggestive of Warthin salivary tumor. However, nuclear features are reminiscent of papillary thyroid carcinoma, and tumor cells are strongly positive for Leu M1 and epithelial membrane antigen, although less for thyroglobulin and negative for calcitonin. The *RET* proto-oncogene encodes a receptor tyrosine kinase for transforming growth factor-β–related neurotrophic factors. Expression of the *RET* proto-oncogene has been detected in numerous tissues, including neural crest–derived cell lines. More recently, *RET* expression has been demonstrated in papillary and follicular thyroid cancer cells.[35] Further evidence for the papillary carcinoma of this variant, Warthin-like tumor of the thyroid gland, lies in expression of the *RET/PTC* fusion gene in these tumors as evidenced by strong reactivity with antibodies against *RET/PTC*.[36]

Among several unique histologic subtypes of papillary carcinoma is tall cell variant, constituting 5 to 10% of all papillary thyroid carcinomas. In patients over 50 years old, it is associated with a more aggressive clinical course as evidenced by a higher rate of recurrence and tumor-related mortality. Tall cell variant tends to present at an advanced stage owing principally to extrathyroid invasion.[37]

Thyroglossal duct cysts are second in incidence only to enlarged cervical lymph nodes as the cause of pediatric neck masses. One percent of thyroglossal duct cysts are malignant, and 81.7% of these are papillary carcinomas.[38]

Although two-thirds of papillary thyroid carcinomas have follicular elements, their behavior is determined by the papillary aspect. These are unencapsulated tumors that tend to invade lymphatic vessels as well as normal surrounding thyroid parenchyma directly. Multicentricity of the primary neoplasm is a common feature, seen in 20 to 80% of patients.

This lesion generally has a prolonged course, and the overall mortality is estimated at 10% or less. The patient's age at diagnosis is the most important prognostic factor. The prognosis is excellent in children and young adults, even in the presence of advanced primary disease or lymph node metastases.[39] Despite the good prognosis, almost 50% of patients who die of papillary carcinoma do so because of local invasion.

Various surgical procedures have been recommended, ranging from subtotal lobectomy to total thyroidectomy. Microscopic disease in the contralateral lobe has been reported in 8 to 88% of cases by various investigators. Total ipsilateral lobectomy, isthmusectomy, and subtotal contralateral lobectomy probably comprise the most common procedure for papillary carcinoma confined to the thyroid gland. For more extensive well-differentiated disease, improved survival requires aggressive management of the primary and cervical node disease by surgery, followed at 6 weeks by thyroid scan and high-dose ^{131}I administration for ablation of the remnant.[40]

With the application of excellent surgical skills, a complete extirpation of the thyroid gland can be performed without recurrent laryngeal nerve injury

or the reported 2% incidence of permanent hypoparathyroidism. Recurrent neck disease must be treated by neck dissection. The concept of "node plucking" is controversial.[41]

Follicular Adenocarcinoma Follicular carcinoma accounts for approximately one of five thyroid cancers, peaking in the fifth decade of life with a 3 to 1 female preponderance. Like papillary carcinoma, it is well differentiated, despite a worse prognosis. These tumors can be encapsulated, or they can invade vasculature. Distant metastases, especially to lung and bone, are found commonly. Of deaths from follicular carcinoma, 75% arise from distant metastases.

For patients under 40 years of age with encapsulated follicular carcinoma, isthmusectomy and lobectomy are performed, with total thyroidectomy reserved for patients over 40 years old. Some surgeons prefer to perform total thyroidectomy for all patients because follicular carcinoma, unlike papillary carcinoma, can undergo late anaplastic transformation. For invasive follicular carcinoma, total thyroidectomy is recommended, followed by ablative radioiodine, because of the potential for distant metastases. In the presence of palpable lymph nodes in the neck, modified radical neck dissection is recommended with preservation of the eleventh cranial nerve and the sternocleidomastoid muscle. Some surgeons do remove the ipsilateral internal jugular vein.

A major difficulty is distinguishing between follicular adenoma and follicular carcinoma. In up to one-third of cases, a frozen-section diagnosis of follicular adenoma is changed to follicular carcinoma. The extent of resection with indeterminate lesions must take this into account with consideration of prognostic factors such as the patient's age and sex and the size of the lesion.[42]

Thallium 201 scintigraphy with serum thyroglobulin provides a basis for follow-up of well-differentiated thyroid cancer. Others use whole-body ^{131}I (WBI) scan at 6 months after thyroidectomy with an ablative dose of ^{131}I. One group reported detection of metastatic tumor based on whole-body ^{99m}Tc-sestamibi scan (WBMIBI) with serum thyroglobulin levels in several patients with negative WBI scans.[43] Moderate elevations in thyroglobulin may be attributable to inadequate thyroid suppression therapy and may require assessment by careful assay of TSH levels. Patients with elevated thyroglobulin levels despite adequate suppression therapy following definitive treatment of a well-differentiated carcinoma are at increased risk of recurrence.[44] Whole-body ^{99m}Tc-sestamibi scan is a potentially valuable tool not only in WBI-negative patients but also in the presence of antithyroglobulin antibodies. Another group recommends use of the FDG PET scan as the study of choice for detection of ^{131}I-negative recurrences and metastases.[45]

Hürthle Cell Carcinoma Hürthle cells observed in thyroid gland atrophy of Hashimoto's disease are a benign entity. Hürthle cells occur in clusters as a degenerative transformation of the follicular cells. Hürthle cell carcinoma can occur as a malignant thyroid neoplasm composed of this cell type; however, this is sometimes classified as a variant of follicular carcinoma. This disease behaves clinically as an intermediate between low-grade and angioinvasive follicular carcinoma. It is difficult to confirm malignancy of a Hürthle cell mass on a pathologic basis alone; therefore, any isolated nodule should be considered malignant on clinical grounds.

Treatment of Hürthle cell carcinoma involves at least an isthmusectomy and lobectomy.[46] If Hürthle cell carcinoma appears as a solitary thyroid nodule exceeding 2.0 cm, or if it is bilateral, total thyroidectomy is recommended. Radical neck dissection is further recommended in the event of cervical nodal metastases.

Medullary Carcinoma with Amyloid Stroma Medullary carcinoma accounts for 5 to 10% of thyroid carcinomas. Lesions are capable of local invasion, spread to regional lymphatic vessels, or distant metastases. They are derived from C cells of the thyroid gland and are capable of producing calcitonin. Calcitonin is an important tumor marker in medullary carcinoma, with the chemoluminescence immunoassay of monomeric calcitonin being more sensitive than nonspecific RIA.[47]

A variant of medullary carcinoma is seen in thyroid carcinoma with mixed follicular and C-cell differentiation patterns and highly heterogeneous patterns of growth. These are referred to as mixed medullary thyroid carcinomas, and their clinical behavior is the same as ordinary medullary carcinoma. Analysis of the two cell types shows that they do not arise from a common precursor as remarkable differences in the genetic profile (*RET* muta-

tions and allelic losses) are displayed. In some cases, the follicular component is oligo-/polyclonal and, thus, probably hyperplastic rather than neoplastic. The morphology of this variant may arise from follicular cells growing into the medullary carcinoma after acquiring some molecular defect. Management of mixed medullary thyroid carcinomas is determined by the true neoplastic component as medullary carcinoma.[48]

Multiple endocrine neoplasia type I (MEN-I) is characterized by parathyroid chief cell hyperplasia with pancreatic islet cell and pituitary adenomas.[49] Two further familial syndromes, multiple endocrine neoplasia type IIa (MEN-IIa) and the rarer type IIb (MEN-IIb), were found in association with medullary carcinoma of the thyroid after its first description in 1959.[50,51]

Multiple endocrine neoplasia type IIa (Sipple's syndrome) consists of medullary thyroid carcinoma with associated pheochromocytoma and parathyroid hyperplasia. Metachronous development of each aspect of Sipple's syndrome can occur. Multiple endocrine neoplasia type IIb is characterized by medullary carcinoma of the thyroid, pheochromocytoma and multiple mucosal neuromas, ganglioneuromatosis, and a Marfan's habitus. Further, a familial non-MEN (FN-MEN) is recognized in which hereditary medullary carcinoma of the thyroid occurs without associated endocrinopathy. Each of these clinical syndromes, MEN-IIa, MEN-IIb, and FN-MEN, is inherited as an autosomal dominant trait. Kindred should be screened by detection of increased serum calcitonin, either basally or following pentagastrin stimulation. Patients with medullary carcinoma or C-cell hyperplasia as a premalignant precursor require prompt and thorough surgical intervention because this malignant lesion is biologically aggressive, with onset at a young age.[52] Children with MEN-IIb should be studied shortly after birth, and those with MEN-IIa should be evaluated by the age of 1 year.[53]

The basis for these familial syndromes associated with medullary carcinoma lies in the common embryologic origin of the parafollicular cell and the adrenal medullary cell. Both are members of the APUD cell population, derived from neural crest cells. The presence of abnormal (nondiploid) DNA as determined by flow cytometric analysis of nuclear suspensions appears to be associated with a worse prognosis in sporadic and familial medullary thyroid carcinoma.[54]

Although papillary thyroid carcinoma is not a characteristic of MEN-II, the diseases have associated genes. The proto-oncogene *RET*, which has been demonstrated to be rearranged in papillary cancer, has been mapped near the predisposition locus for MEN-II, with apparent tight linkage between loci. *RET* is expressed in medullary thyroid carcinoma and pheochromocytoma, making it a candidate gene for MEN-II. The variations in expression of the MEN-II syndrome may be attributable to structural alteration of several contiguous genes in the MEN-II chromosomal region.[55] Techniques of molecular biology will likely provide the means to assess and control this form of malignant disease in the foreseeable future.

Medullary carcinoma does not take up radioiodide, so this treatment is not recommended. Because of the propensity of medullary carcinoma for local microvascular invasion, late recurrence (up to 20 years after treatment), and metastatic disease, aggressive surgical intervention is recommended. Furthermore, external ionizing radiation and chemotherapy are of limited value. A total thyroidectomy is generally recommended because of the high incidence of bilaterality, especially in familial syndromes. Most clinicians favor a radical neck dissection on the side of the primary neoplasm with a modified neck dissection on the contralateral side in the event of bilateral thyroid gland involvement. The incidence of bilateral disease in the gland or neck is especially high in familial disease. Attention to family history and evaluation for other endocrinopathies are imperative. As with well-differentiated thyroid adenocarcinomas, medullary carcinomas can eventually degenerate to anaplastic carcinoma.

Anaplastic Carcinoma Anaplastic carcinoma, largely a disease of the elderly, accounts for less than 10% of malignant thyroid tumors. This disease has shown a three-fold to four-fold decline in incidence internationally over the past 45 years, relating inversely to the increase in extirpative thyroid procedures.[56] Foci of differentiated thyroid cancer are commonly observed, suggesting a degenerative origin, based on "clonal evolution." Chromosomal aberrations 5p, 8p, and 8q, rarely found in differentiated thyroid carcinoma, are found in anaplastic carcinoma. These chromosomes may harbor gene loci affecting aggressiveness of neoplastic growth

and, thus, could play an important role in carcinogenesis.[57] Up to 20% of patients have a history of a previously treated well-differentiated carcinoma of the thyroid.

Anaplastic carcinoma is an often rapidly fatal, aggressive, enlarging, bulky mass that distorts the anterior neck and often obstructs the aerodigestive tract. In many cases, the patient undergoes biopsy by needle aspiration or open biopsy, and the neoplasm is deemed inoperable. Palliative radiation therapy and chemotherapy may prolong life.[58] Patients with the small cell type may have a slightly longer survival than those with large cell anaplastic carcinoma.

Other Malignant Lesions Sarcomas comprise less than 1% of malignant tumors of the thyroid. They have a poor prognosis despite aggressive surgical intervention. All varieties of lymphoma arise in the thyroid gland, most frequently non-Hodgkin's lymphoma. This lesion may appear as a rapidly enlarging goiter in an elderly person. Treatment usually involves lobectomy or total thyroidectomy followed by external beam radiation. A wide range of thryoid neoplasms representing metastases from distant sites has been reported.

Surgical Anatomy

In the medial aspect, the thyroid gland is deep to the sternohyoid and underlying sternothyroid (strap) muscles. Inferiorly, the sternocleidomastoid muscle overlaps the strap muscles, whereas more superiorly, it overlaps the lateralmost extent of the thyroid gland. The anterior surface of the thyroid gland faces the pretracheal layer of the deep cervical fascia, which extends around posteriorly to envelop the trachea, esophagus, and recurrent laryngeal nerves.

The deep or medial surface of the thyroid gland is adapted to the larynx and trachea. At the superior pole of each lateral lobe, the medial surface is in contact with the inferior pharyngeal constrictor and the posterior portion of the cricothyroid muscle. The medial aspect of the thyroid gland is also intimately related to the parathyroid glands and the recurrent laryngeal nerves. The correct level of dissection for dorsal mobilization is anatomically defined as lying anterior to the lamella that covers the vessel- and nerve-containing plane of the neck.[59] On occasion, the gland may extend posterior to the trachea. Appreciation of these relations contributes to surgical technique and understanding of the symptoms of thyroid masses.

Suspension of the Thyroid Gland The thyroid gland is suspended in the neck from the cricoid and thyroid cartilage by two ligaments: the anterior suspensory ligament and the surgically more important posterior suspensory (Berry's) ligament.

The anterior suspensory ligament attaches the superior border of the isthmus and superomedial aspect of each lateral lobe to the cricoid cartilage and cricothyroid muscle.

The posterior suspensory ligament is distinct from the anterior suspensory ligament and becomes a pedicle for the thyroid gland in the course of thyroid lobectomy. Berry's ligament extends from the medial and deep portions of each lateral lobe of the thyroid to the lateral surface of the cricoid cartilage as well as the posterior lateral aspect of the first and second tracheal rings. It is an important surgical landmark because the recurrent laryngeal nerve passes either deep to this ligament or between leaves of the ligament in its cephalic course. The recurrent laryngeal nerve often divides into two branches inferior to this ligament; therefore, all branches must be visualized to avoid injury. Furthermore, a portion of the lateral thyroid lobe extends deep to the ligament. In some patients, the ligament may actually be embedded in glandular tissue, demanding meticulous surgical dissection.

An occasional anomaly observed is a unilateral or paired bilateral muscle, the levator glandulae thyroideae muscle, inserting into the thyroid gland and originating from the hyoid bone.

Blood Supply The blood flow to the thyroid gland is on the order of 100 mL/minute, reflecting the importance of thyroid hormone to whole-body metabolism. The thyroid gland receives this abundant blood supply through paired inferior thryoid arteries as well as an occasional unpaired lowest thyroid artery of variable size (Figure 63–1).

The dominant artery to the thyroid is usually the inferior thyroid artery, although it can be of variable origin and relationship. Typically, it ascends along the medial side of the anterior scalene muscle behind the prevertebral fascia. Thereafter, it loops down onto the anterior surface of the longus colli to pass behind the common carotid artery in variable relation to the middle cervical ganglion of the sym-

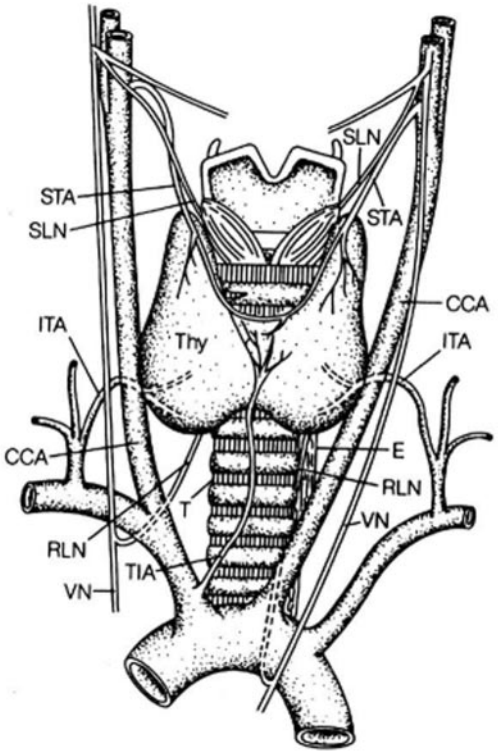


FIGURE 63–1. Arteries of the thyroid and the laryngeal nerves, anterior view. Dissection of the recurrent laryngeal nerve (RLN) begins in the triangle bounded by the trachea (T), the common carotid artery (CCA), and the thyroid gland (Thy). Branches of the inferior thyroid artery (ITA) can run close to the recurrent nerve. The esophagus (E) runs somewhat to the left of the trachea. The superior thyroid artery (STA) is closely related to the external branch of the superior laryngeal nerve (SLN). The vagus nerve (VN) courses parallel to the carotid artery. The thyroid ima artery is labeled TIA.

pathetic trunk. Medial to this point, it penetrates the prevertebral fascia and, in variable relationship to the recurrent laryngeal nerve, divides into a dominant upper branch to just below the midportion of the lateral thyroid lobe and a lower branch to the inferior border of the lobe. Small branches of the inferior thyroid artery cross deep to Berry's ligament and along its inferior edge, where clamping for hemostasis may injure the recurrent laryngeal nerve. Duplication of the inferior thyroid artery has been observed; in this case, both inferior thyroid arteries arise from the subclavian artery on one side. In the occasional absence of an inferior thyroid artery, the ima thyroid artery may replace it, or it may simply be an accessory vessel. Most commonly, the ima thyroid artery arises from the innominate (brachiocephalic) artery.

The superior thyroid artery is the first anterior branch of the external carotid artery, although it can arise from the bulb of the common carotid artery. In its descent toward the apex of the thyroid lobe, the artery is closely related to the course of the external branch of the superior laryngeal nerve. This nerve is therefore a potential source of complication in thyroid surgery. On approach to the gland, the superior thyroid artery divides into an anterior and a posterior branch. The anterior branch runs close to the medial aspect of the lobe and across the isthmus superiorly, often anastomosing with the superior thyroid artery from the opposite side. The course of the superior thyroid artery is the most constant of the arteries to the thyroid gland.

Venous drainage of the thyroid gland is by two or three pairs of veins that anastomose freely in the gland. These vessels are named by location: superior, middle, and inferior thyroid veins (Figure 63–2).

The superior thyroid vein emerges from the upper portion of the thyroid gland and accompanies the superior thyroid artery to empty into the internal jugular vein or lower end of the common facial vein at the level of the carotid bifurcation.

The middle thyroid vein is variable in presence and size and arises from the lateral surface of the thyroid, crossing in front of the common carotid artery to drain into the internal jugular vein. It may be torn in dissecting the gland from the carotid sheath. The middle thyroid vein is present in roughly 50% of patients.

The inferior thyroid vein drains the lower portion of the gland to the brachiocephalic vein of the same side. The inferior thyroid veins form the plexus thyroideus impar, a rich anastomosis in front of the trachea, which can produce troublesome bleeding in a tracheostomy. Occasionally, the inferior thyroid veins join inferiorly to drain by way of a common trunk to the left brachiocephalic vein, which is referred to as a thyroidea ima vein.

Neural Input Neural fibers reach the thyroid gland from the superior cervical ganglion of the sympathetic trunk and the superior laryngeal branch of the vagus. It is generally accepted, how-

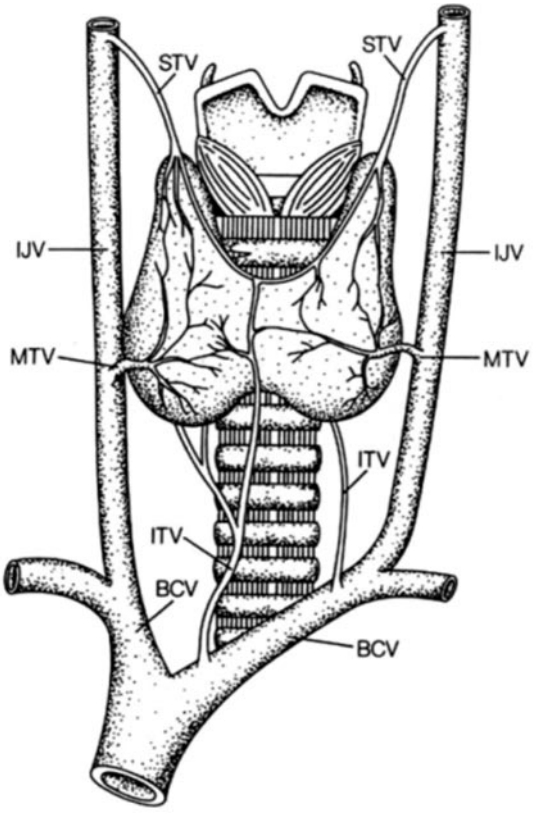


FIGURE 63–2. Venous drainage of the thyroid gland, anterior view. The superior thyroid vein (STV) closely approximates the course of the superior thyroid artery. The middle thyroid vein (MTV) occurs variably, draining into the internal jugular vein (IJV). The inferior thyroid veins (ITV) drain into their respective brachiocephalic veins (BCV).

ever, that these nerves are related functionally to blood vessels only and not directly to the function of glandular tissue.

Lymphatics Lymphatic drainage of the thyroid is chiefly by inferior and superior lymphatic vessels accompanying the arterial blood supply. The isthmus is mostly drained by inferior lymphatic channels, as are the lower portions of the lateral lobes. The lower channels then distribute to the lower deep cervical lymph nodes, the supraclavicular lymph nodes, and the pretracheal and paratracheal lymph nodes, as well as the prelaryngeal lymph nodes just above the isthmus. The Delphian node is a median prelaryngeal lymph node often palpable in the presence of thyroid disease. Drainage also occurs to the brachiocephalic lymph nodes, which are related to the thymus in the superior mediastinum. Lymph vessels from the thyroid gland may enter into the thoracic duct directly. Superior lymphatic channels drain the superior and medial surfaces of the lateral lobes as well as the superior surface of the isthmus. The upper channels empty into the upper cervical lymph nodes.

PARATHYROID GLANDS

The parathyroid glands consist of four small, ovoid, yellowish-brown structures located between the posterior border of the thyroid lobes and its capsule. These endocrine glands produce PTH, which acts to increase serum calcium. They are named anatomically with respect to laterality and superior or inferior position. Each gland is normally about 50 mg and measures on the order of 6 mm by 4 mm by 2 mm, with the least dimension in the anteroposterior orientation. Adenomas of the parathyroid glands can weigh from 120 to 800 mg. Generally, PTH production is proportional to the mass of the gland.

The parathyroid glands are endoderm derivatives of the third and fourth pharyngeal pouches. Like the thymus, the *inferior* parathyroid glands (IPGs) are derivatives of the *third* pharyngeal pouch. Therefore, IPGs are also referred to as *parathyroids III*. Similarly, the *superior* parathyroid glands (SPGs), derived from the *fourth* pharyngeal pouch, are referred to as *parathyroids IV.*

The parathyroids have a rich blood supply from the inferior thyroid artery and the anastomotic artery that connects the superior to the inferior thyroid arteries. This anastomotic artery can serve as a guide in locating parathyroid glands. Dense lymphatics drain the parathyroids in close association to lymphatics of the thyroid gland and thymus. Neural input to the parathyroids is sympathetic, arising from the superior or middle sympathetic ganglion for vasomotor control, although glandular function is controlled by serum calcium rather than bloodflow.

The SPGs are more constant in position than the IPGs. The SPGs are usually located around the middle of the posterior border of the lateral lobe of the thyroid gland. The inferior glands have more varied positions. Inferior parathyroid glands can occur within the thyroid capsule below the inferior thyroid artery near the inferior pole of the lateral thyroid lobe. Tumors arising from the IPGs in this

relation tend to grow down the inferior thyroid veins, anterior to the trachea, and into the superior mediastinum. Inferior parathyroid glands can also present posterior to the thyroid capsule and above the inferior thyroid artery, and tumors arising from these can develop posteriorly and inferiorly to grow behind the esophagus and into the posterior mediastinum. Inferior parathyroid glands can also occur within the thyroid lobe itself near the inferior aspect of the posterior border. However, parathyroid glands can exist as diffusely scattered collections of parathyroid tissue in connective tissue or fat, or represented as only three distinct glands. Generally, SPGs are *deep* and IPGs are *superficial* to the recurrent laryngeal nerves on each side.

In children, parathyroid glands consist of cords or columns of cells called chief cells with fine connective tissue septae extending from a thin capsule into the gland. Around 6 years of age, other cells appear, the oxyphil cells, which are probably derived from chief cells. The oxyphil cells are larger than chief cells with a greater volume of eosinophilic, granular cytoplasm. The oxyphil cells apparently have high metabolic activity with abundant mitochondria and glycogen particles but sparse endoplasmic reticulum. Lipid deposits accumulate progressively in parathyroids with aging. Large sinusoidal capillaries are distributed between the columns. Chief cells vary in cytoarchitecture, with the most active staining darker with more granular endoplasmic reticulum, a prominent Golgi complex, and numerous membrane-bound secretory vesicles. The chief cells secrete PTH from the cytoplasm into the pericapillary spaces by exocytosis.

Calcitonin and PTH-related peptide (PTHrP) are important for calcium regulation in the fetus. In adults and children, calcium-phosphorous homeostasis is regulated principally by vitamin D and PTH.

Parathyroid hormone–related peptide is homologous to PTH in the first 13 amino acids of the amino terminus, of which 8 amino acids are identical. There is no homology with the mid or carboxyl regions. Parathyroid hormone–related peptide is transcribed from a gene on the short arm of chromosome 12, whereas the gene for PTH is located on the short arm of chromosome 11. Parathyroid hormone–related peptide acts on PTH receptors in bone and kidney cells, increases urinary cAMP and production of 1α-hydroxylase involved in production of 1,25-dihydroxycholecalciferol ($1,25[OH]_2D_3$). Parathyroid hormone–related peptide is vital to normal fetal development, and production of PTHrP occurs in nearly every cell type, including every embryonic tissue at some point in development. For instance, production in keratinocytes is important in keratinocyte differentiation and development of hair follicles. Parathyroid hormone–related peptide gene deletion is a lethal error in animal models. Parathyroid hormone–related peptide produced by fetal parathyroids is present in cord blood at three-fold higher levels than that found in adult serum. Parathyroid hormone–related peptide is critical to maternal-fetal calcium transfer, providing the fetal requirement of 30 g of calcium, and is essential for normal fetal skeletal development and maturation. Particularly during the second trimester, maternal absorption of calcium should increase from around 150 to 400 mg daily. Parathyroid hormone–related peptide levels are also increased in lactation. Pasteurized bovine milk, as well as human breast milk, contains PTHrP at levels 10,000 times greater than normal serum.[2]

Whereas PTHrP is increased in benign breast hypertrophy, PTHrP is the most important of tumor-derived factors that act to increase renal reabsorption of calcium and increase osteoclastic bone resorption. Parathyroid hormone–related peptide can be overexpressed in squamous cell carcinomas of the head and neck. Increased circulating PTHrP derived from tumor is associated with paraneoplastic syndrome humoral hypercalcemia of malignancy, occurring in 2 to 10% of all cancer patients, although infrequently reported with head and neck cancer. Attempts to use PTHrP as a tumor marker have not been successful.[60]

The biologic activity of PTH lies in the first 34 residues of its 84 amino-acid chain (9,500 daltons). Secretion of PTH is increased in response to falling serum calcium levels. Parathormone acts on target cells of the kidney and bone via specific transmembrane receptors, activating a transduction pathway involving a G protein coupled to the adenylate cyclase system. In the kidney, PTH stimulates activity of 1α-hydroxylase involved in production of $1,25(OH)_2D_3$. Increased $1,25(OH)_2D_3$, in turn, induces synthesis of calbindin-D, a calcium-binding protein of the intestinal mucosa required for absorption of calcium. 1,25-Dihydroxycholecalciferol is also required for PTH to mobilize calcium by increasing resorption of bone.

Parathormone is synthesized in the parathyroid glands from an initial 115 amino-acid chain, pre-pro-PTH. This is converted to a 90 amino-acid chain, pro-PTH, which is hydrolyzed to PTH as the main secretory product. Once released, PTH is rapidly broken down in the liver and kidney into three smaller fragments, of which the 1 to 34 amino-terminal chain has biologic activity, although present in very small serum concentrations. The existence of these three derivatives is the basis for numerous assays. The carboxyl-terminal chain is more slowly cleared through glomerular filtration and is present at serum levels 50- to 500-fold greater than active hormone. Assay of the carboxyl-end fragment is useful in testing for hyperparathyroidism, although less applicable to evaluate for secondary hyperparathyroidism in patients with renal failure. Differentiation of subnormal concentrations that occur in hypoparathyroidism requires certain specific sensitive RIAs for PTH. Of particular importance is a sensitive 15-minute immunochemiluminometric assay that has been developed for intraoperative use.[2]

With inadequate serum levels of PTH (hypoparathyroidism) and consequential hypocalcemia, convulsive spasms of muscles (tetany) ensue. If laryngeal and respiratory musculature become involved, tetany can cause death. Treatment involves administration of intravenous or enteric calcium and calcitriol or vitamin D, and possibly magnesium.

The most common cause of surgical hypercalcemia is primary hyperparathyroidism, excessive PTH production from a primary defect of the parathyroid glands. Parathyroid adenoma is the cause of primary hyperparathyroidism in 80 to 85% of cases.[61] Symptoms arise from effects on several systems: renal, with polyuria, and renal colic owing to lithiasis; rheumatic, with bone and joint pain; neuronal, with fatigue, memory loss, and depression; and gastrointestinal, with dyspepsia, anorexia, and nausea. Thus arises the medical student's memory aid, rhyming the symptoms: "aching bones, renal stones, mental moans, abdominal groans." In hyperparathyroidism, excessive PTH causes demineralization of bone from calcium ionization. The resultant hypercalcemia may be associated with metastatic or aberrant calcification in vital organs including the kidneys.[2]

Calciphylaxis is a rare disorder typically affecting renal failure patients. This condition manifests in vascular calcification with subsequent skin necrosis, gangrene, and often death from sepsis. Skin lesions arise from acute local calcification beginning with mottled discolorations that indurate. Skin biopsy can establish diagnosis early. Dry gangrene of fingers and toes can occur, and a case of lingual necrosis has been documented. Parathormone is thought to act as a tissue sensitizer leading to soft tissue changes. Survival is 65% when parathyroidectomy is used to control this complicated condition compared with 35% without surgery.[62]

Management of Hyperparathyroidism

The majority of patients in the United States with primary hyperparathyroidism are asymptomatic. The usual treatment is surgical removal of the abnormal parathyroid gland(s). There is currently no effective medical treatment. Oral phosphate administration is dangerous with the risk of metastatic calcification. Estrogen therapy in postmenopausal women with mild hyperparathyroidism decreases serum calcium concentrations slightly without changing PTH levels. Although bisphosphonates can inhibit bone resorption, they have had, at most, a transient effect. Parathyroid cells have a calcium-sensing receptor that regulates synthesis and secretion of PTH. When stimulated by increased extracellular calcium, the activated receptor signals the cell via a G protein–transducing pathway that raises intracellular calcium concentration and inhibits secretion of PTH. This has encouraged a search for calcimimetic drugs that might serve to inhibit parathyroid cell function by activating the calcium receptor. One such molecule is phenylalkylamine (R)-N-(3-methoxy-alpha-phenylethyl)-3-(2-chlorophenyl)-1-propylamine or R-568. This drug inhibits secretion of PTH in postmenopausal women with mild primary hyperparathyroidism with a consequential decrease in serum calcium. There is also an associated hypercalciuric response, presumably from alleviation of parathormone-driven, renal tubular resorption of calcium.[63] Such drugs may lead to medical treatments, which can lower serum calcium and PTH secretion without surgery.

Indications for parathyroidectomy with hypercalcemia include serum calcium levels consistently greater than 1 to 1.4 mg/dL above normal in patients under 50 years of age (eg, serum calcium

12.0 mg/dL); calcium renal calculi, urinary calcium greater than 400 mg/24 hours, or reduced renal function; bone density greater than 2 standard deviations from normal; or coexisting disease that would make observation inappropriate. Assessment of bone density has become more accurate, and bone density improves up to 20% after parathyroidectomy. Certainly, parathyroidectomy prior to development of serious bone demineralization is beneficial. Improvement in glucose control attendant to parathyroidectomy suggests that diabetic patients with hyperparathyroidism should have surgery. Surgical treatment for nonspecific symptoms such as weakness and fatigability is debatable.[64]

Preoperative scanning with ^{99m}Tc-sestamibi scintigraphy has become the preferred localization study. Sestamibi scan can also be done intraoperatively, as described in more detail below.[65] These scans are low cost, specific, and sensitive. Accurate localization allows parathyroidectomy of the specific abnormal gland. This diagnostic tool has not reduced the failure rate of parathyroidectomy as sestamibi scans can miss multiglandular disease and have a small rate of false-positive results. Nonetheless, sestamibi scans are necessary for unilateral or focused parathyroidectomy and useful in patients with thyroid disease or previous thyroidectomy.

Parathyroid Masses

There have been relatively few clinically significant though nonfunctional cysts of the parathyroid glands. Typically, these solitary masses average 4 cm in diameter and present in the region of the inferior aspect of either lateral thyroid lobe, which is pushed forward by the cyst. Generally, these cysts are relatively avascular and "shell out" surgically. They can occur in all ages after childhood and without sexual predilection. Such cysts may be asymptomatic or may cause tracheal deviation, symptoms of respiratory restriction, or even hoarseness owing to pressure on the recurrent laryngeal nerve. It is fairly rare for benign, noncystic parathyroid lesions to present as palpable masses; however, up to 45% of parathyroid carcinomas are reportedly palpable.[66] Parathyroid carcinoma is rarely encountered but can be associated with uncontrollable hypercalcemia with metastases. Only 0.017% of cases of hyperparathyroidism arise from parathyroid carcinoma.[64] Parathyroid carcinoma should be surgically excised with appropriate margins of normal tissue and regional dissection as necessary.

Surgical Techniques

Total Thyroid Lobectomy and Isthmusectomy

The patient is placed in a supine position on a shoulder roll with hyperextension of the neck. A transverse incision is made through skin and platysma muscle in the anterior neck within 3 cm above the suprasternal notch and extended laterally as necessary. Hereafter, meticulous hemostasis should be maintained throughout the procedure. By subplatysmal dissection, skin flaps are developed vertically from the sternoclavicular border to the eminence of the thyroid cartilage and laterally from the midline to the anterior border of the sternocleidomastoid muscle. The strap muscles, sternohyoid and sternothyroid, are divided vertically in the midline as needed for exposure; dissection progresses in the plane of the pretracheal fascia, allowing retraction of the strap muscles to expose the true capsule of the isthmus and entire lateral lobe on the superficial surface. Superiorly and laterally, this requires dissection underneath the sternocleidomastoid muscle (Figure 63–3). The middle thyroid vein, if present, may be divided and ligated at the lateral border of the lateral lobe of the thyroid gland at this time. Dissection of the deep surface of the thyroid gland begins inferiorly and laterally with exposure of the main trunk of the recurrent laryngeal nerve. This is bounded by a triangle consisting of the trachea and esophagus medially, the common carotid artery laterally, and the inferior border of the thyroid gland superiorly. The apex of this triangle faces inferiorly, and blunt dissection begins at this area. The recurrent laryngeal nerve is at a depth at or just anterior to the tracheoesophageal sulcus and may lie in the sulcus or several centimeters lateral to it. Occasionally, the right laryngeal nerve enters the triangle laterally from behind the common carotid and passes superiorly and medially in an oblique course. Because the esophagus lies to the left of midline, the left recurrent laryngeal nerve may be close to the esophagus.

The parathyroid glands should now be identified. They appear as vascular, brownish fat. Whereas the inferior parathyroid glands have a variable location, the superior parathyroid glands are located more consistently at the midportion of the deep sur-

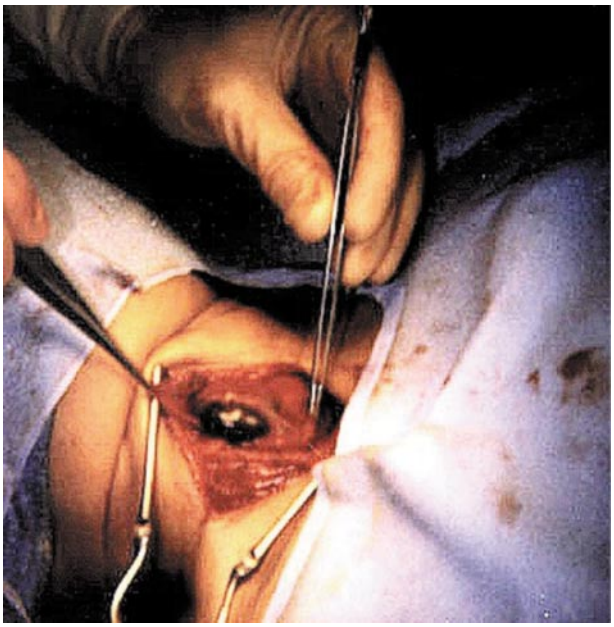


FIGURE 63–3. Photograph from hemithyroidectomy for a predominantly right-sided goiter displacing the trachea to the left and causing dysphagia. Cephalad is toward the left edge of the image. The transverse skin incision has been made and skin-platysma flaps are held back with a self-retaining retractor. The strap muscles have been divided in the midline, and the goiter has just become visible as a dark, largely cystic mass.

face of the thyroid lobe. As the dissection progresses, the inferior thyroid vein is divided and ligated. With the recurrent laryngeal nerve identified, the inferior thyroid artery is now divided and ligated close to the gland to avoid damaging the blood supply to the parathyroid gland.

The isthmus is now transected on the contralateral side and oversewn with a running 0 silk suture. The anterior suspensory ligament is transected, freeing the superior border of the isthmus and superior pole medially from the cricoid cartilage (Figure 63–4). With the course of the recurrent laryngeal nerve exposed nearly to the superior border of the thyroid lobe, the posterior suspensory ligament is identified by lateral retraction of the lobe. The recurrent laryngeal nerve runs between leaves of this ligament or deep to it (Figure 63–5). As with dissection of the facial nerve, the recurrent laryngeal nerve should be exposed and in full view when dividing any tissue.[67] The posterior ligament is sharply transected close to the trachea and cricoid cartilage. The gland may be adherent in this area,

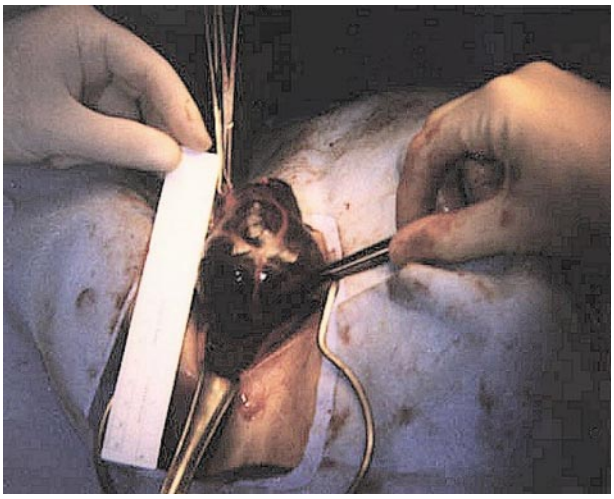


FIGURE 63–4. The right thyroid lobe has been dissected laterally and inferiorly and an isthmusectomy has been performed. Cephalad is toward the left edge of the image. The lobe is still attached by Berry's ligament and the pedicle of the superior thyroid artery and vein. At this point, the tumor is nearly delivered from the field, although injury to the superiormost aspect of the recurrent laryngeal nerve could occur until release from Berry's ligament.

and a postoperative scan often reveals uptake from residual gland.

The parathyroid glands may be superficial to the recurrent laryngeal nerve in the plane of dissection or even attached or contained by the thyroid capsule; therefore, it is prudent to examine the capsule after removal of the thyroid lobe. Parathyroid glands may appear as dark tissue suggestive of a subcapsular hematoma. Parathyroid tissue sinks in saline, which provides an intraoperative means to test suspect tissue. Frozen section of a small portion of the tissue in question should be used to confirm the presence of a parathyroid gland. If a parathyroid gland is unintentionally removed, reimplantation, as described later, is recommended.

At this point, the superior thyroid artery and vein should be skeletonized by blunt dissection while held on tension with slight downward traction of the thyroid lobe. The veins and arteries should be ligated separately with free ligatures to avoid superior laryngeal nerve injury and arteriovenous fistula. To avoid injury, some recommend that the superior laryngeal nerve be skeletonized with periapical dissection in the visceroverterbral angle.[68] It is possible

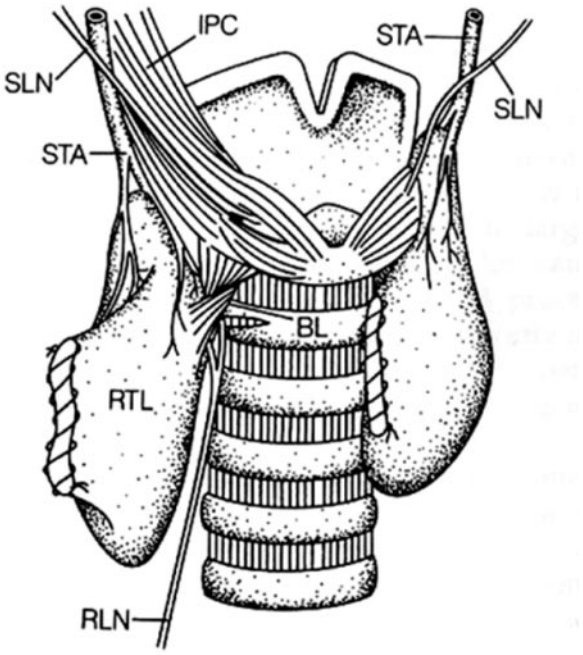


FIGURE 63–5. Total thyroid lobectomy and isthmusectomy, oblique view. The right thyroid lobe (RTL) and isthmus are under traction laterally and inferiorly, which tends to separate the superior thyroid artery (STA) from the external branch of the superior laryngeal nerve (SLN). Note the relation of the superior pole of the lateral lobe to the inferior pharyngeal constrictor (IPC). The recurrent laryngeal nerve (RLN) often branches as it runs deep to or through the posterolateral suspensory (Berry's) ligament (BL).

to find thymus tissue attached to the superior pole of the thyroid gland. The lateral lobe and isthmus are then removed from the surgical field.

Figure 63–6 demonstrates the right recurrent laryngeal nerve following hemithyroidectomy to remove a right-sided goiter (Figure 63–7). In certain situations, the recurrent laryngeal nerve is appropriately sacrificed. These situations include ipsilateral vocal cord paralysis with reasonable certainty of malignant disease grossly invading the nerve, so the patient's life would be compromised by preserving the nerve.

Assuming that the parathyroid glands have either been preserved or reimplanted, the wound is closed in layers, with good cosmetic technique. Suction drainage of the surgical field is practiced by most surgeons, usually through a separate stab incision. Postoperative management without drainage of the surgical field has been reported in large numbers without evidence of increased risk, however.[69,70] Some surgeons perform flexible fiberoptic endoscopy in the operating room immediately following extubation to document intact laryngeal motor function.

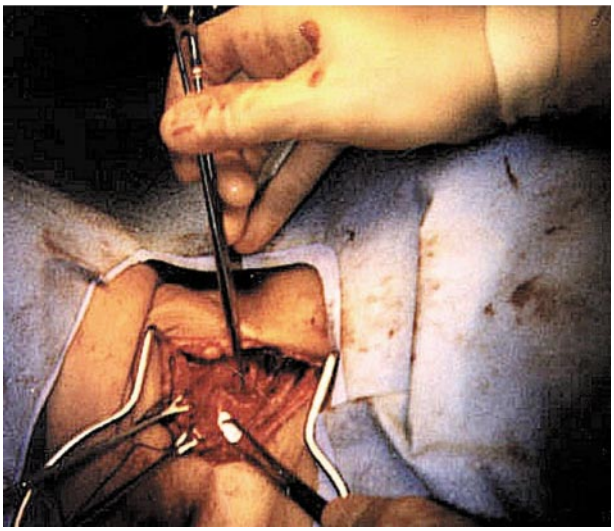


FIGURE 63–6. The lobectomy and isthmusectomy are completed. Cephalad is toward the left edge of the image. The trachea is being retracted slightly with a hemostat, and, near its tip, the right recurrent laryngeal nerve is demonstrated in the tracheoesophageal groove.

FIGURE 63–7. The dark, largely cystic right thyroid lobe and isthmus after excision, longitudinal dimension approximately 8 cm. Were parathyroid glands not identified, a search for them by subcapsular exploration of the posterior aspect of the specimen would be appropriate, and reimplantation would be performed.

Although cervical incision is usually adequate, in operations for recurrent goiter, very large posterior goiters, or carcinoma requiring mediastinal dissection, sternotomy may be indicated. Recurrent laryngeal nerve injury is more easily avoided in these more challenging cases by the following: systematic identification of the recurrent nerve (see Figure 63–6), avoidance of enucleating the intrathoracic side of the goiter first, and meticulous dissection to avoid stretching the nerve, especially on the left side. Rarely, tracheomalacia secondary to prolonged compression of the trachea by the goitrous mass requires careful postoperative observation.[71]

Technique for Reimplantation of Parathyroid Tissue Parathyroid glands removed inadvertently should be cleaned carefully in saline and diced into small pieces measuring approximately 1 × 1 mm.[72] Up to 40 pieces of parathyroid grafting material may be obtained by this technique. The parathyroid gland may be implanted into the sternocleidomastoid muscle[73] or into the volar surface of the previously prepared and draped nondominant forearm. Small pieces of parathyroid tissue can be implanted into the brachial, radialis, or flexor muscle group, closing the overlying muscle and fascia with interrupted suture of 5-0 mersilene to prevent extrusion and further to serve as a marker for the implantation site. Bleeding in the implantation site should be carefully avoided. The total surface area of the parathyroid transplantation site is generally 5 × 5 cm. This is closed in standard fashion. In the immediate postoperative period, PTH levels can be undetectable but generally recover to preoperative levels within about 2 weeks.[74]

Techniques of Parathyroidectomy Currently, two markedly different techniques of parathyroidectomy are practiced: bilateral, open neck exploration and minimally invasive radioguided parathyroidectomy (MIRP).[75]

Traditionally, surgery for biochemically proven primary hyperparathyroidism involved bilateral neck exploration following the general approach described earlier for thyroidectomy. This technique is still appropriate for parathyroid hyperplasia or clinical situations in which MIRP is contraindicated. At least 97% of parathyroid glands can be reached through the neck, so mediastinal dissection is seldom necessary. The safest neck exploration is performed in a highly controlled, bloodless surgical field. The inferior thyroid artery is usually preserved and often provides the main blood supply to adenomas. The lateral thyroid lobe is reflected medially and anteriorly to gain access to the posterior capsule and prevertebral fascia.

For management of solitary parathyroid adenoma, unilateral, focused parathyroidectomy is a rapidly advancing and appealing alternative. Currently, only about 50% of patients are candidates for focused parathyroidectomy. Several important technological tools have become available to the parathyroid surgeon that make unilateral surgery feasible: imaging of parathyroid tissue by ^{99m}Tc-MIBI scintigraphy,[76] use of an intraoperative gamma probe,[75] and intraoperative rapid PTH assay.

By 1997, endoscopic approaches to neck exploration and parathyroidectomy had been developed in animal models and were applied to human cadavers.[77] In early 1998, successful endoscopic parathyroidectomy with resection of 3^1/$_2$ parathyroid glands for hyperplasia was reported.[78] Published 1 month earlier was a report of four patients with parathyroid adenoma removed by unilateral endoscopic parathyroidectomy. In each patient, preoperative localization imaging techniques had been applied, and endoscopic exploration took place through one 11 mm and two 5 mm ports. The adenoma was endoscopically located and successfully removed in each case with rapid correction of hypercalcemia. The patients did not require analgesics, and the scars were cosmetically minimal, with excellent patient acceptance.[79] Subsequently, there has been a plethora of reported experience worldwide with endoscopic parathyroidectomy by various techniques, including "pure" endoscopic approaches using constant gas insufflation and video-assisted gasless techniques, unilateral approaches via neck or extracervical (thoracic or axillary) ports, and bilateral endoscopic explorations by anterior chest incisions. The postoperative pain is generally very limited. Cosmetically, visible surgical scars are minimized or essentially avoided.[80] Application of intraoperative sestamibi scans by gamma probe further facilitates endoscopic localization of adenoma with or without video imaging in minimally invasive surgical fields. Focused parathyroidectomy and, in particular, MIRP have made outpatient parathyroidectomy possible for

many patients.[81] Operating time by MIRP is comparable to that of open parathyroidectomy, with a typical report stating mean duration of 65 to 84 minutes with a range from 30 to 180 minutes.[82,83]

The following situations may present contraindications to MIRP: presence of coexisting nodular thyroid disease, previous neck surgery or irradiation, suspicion of parathyroid hyperplasia, history of familial primary hyperparathyroidism, or other anatomic or medical contraindications. In one report, one in six patients undergoing MIRP required conversion to open parathyroidectomy.[84] Other authors reported greater success, however, emphasizing it to be a feasible surgical procedure with normal postoperative serum calcium and PTH levels in all cases, reduced postoperative pain compared to conventional parathyroidectomy, and an acceptable length of operating time and cosmetic results.[82,83]

Strictly speaking, MIRP involves use of a handheld intraoperative detector such as the Navigator miniature handheld probe, which can be used for endoscopic procedures to guide the surgeon to a parathyroid adenoma. This requires injection of 20 mCi ^{99m}Tc sestamibi 2 hours preoperatively and performing a parathyroid scan. If the scan confirms a single adenoma, patients are taken to the operating room straight away, and a general anesthesia or local anesthesia with sedation is administered.[85]

The procedure by a lateral neck approach begins with a 15 mm transverse skin incision on the anterior border of the sternocleidomastoid muscle. The fascia connecting the thyroid lobe and lateral aspect of the strap muscles is divided adequately to permit visualization of the prevertebral fascia. After the surgical field is adequately opened, three trocars are inserted: a 12 mm trocar through the incision and two 2.5 mm trocars, one above and one below the larger trocar. Carbon dioxide is insufflated at 8 mm Hg. Unilateral, parathyroid exploration is performed through a 10 mm 0-degree videoendoscope. Magnification afforded by the fiber-optic system is excellent for identifying structures in the neck. The working space in an endoscopic surgical field is usually adequate as long as good hemostasis is maintained. Endoscopic mobilization of the lateral thyroid lobe for access to the tracheoesophageal groove can be technically difficult. The handheld probe is used to guide the surgeon to the adenoma. With the adenoma identified, trocars are withdrawn. Working directly through the skin incision, the lateral thyroid lobe is retracted medially and the adenoma is excised, clipping its pedicle.[83] Review of the literature suggests that most surgeons conclude the procedure once the adenoma is excised. However, one recommends identification and biopsy of the second gland on the same side and, if normal, concluding the operation. If the second nonadenomatous gland is histologically abnormal or if the ^{99m}Tc sestamibi scan suggests multiple lesions, then bilateral neck exploration is recommended.[86]

Although expensive, intraoperative rapid PTH assay is helpful in focused parathyroidectomy. Excision of sufficient abnormal parathyroid tissue is correlated with assays performed prior to incision, and at 5 and 10 minutes after excision, that demonstrate a measured decrease of 59%. Intraoperative rapid PTH assay is most reliable in treatment of a single adenoma but is not as reliable in parathyroidectomy for hyperplasia.[6]

Sestamibi has good affinity for hyperfunctioning parathyroid lesions as *hot nodules*. A two-phase protocol with early (10 minutes) and late (120 minutes) imaging is used. Lesions with mass greater than 250 mg are detected with this scintigraphy (sensitivity 85 to 90% with a specificity of nearly 100%), although hyperplastic glands are not imaged as effectively as adenomas. This modality is better than ultrasonography (sensitivity 82.9%) to detect ectopic parathyroid tissue and recurrent disease. It is also superior to imaging by other radionuclide tests such as subtraction thallium 201 (sensitivity 84.6%), ^{99m}Tc pertechnetate, or tetrofosmin scanning.[61,87,88] A reduced time-window scan is recommended for patients with a high degree of suspicion for parathyroid adenoma not visualized by a standard sestamibi scan. In one study in which 3 of 31 patients with adenomas had negative ^{99m}Tc-sestamibi scans, reduced time-window scans, single-view scans at 15, 30, 45, and 120 minutes, and tomograms at 60 minutes revealed two of three adenomas that would otherwise have escaped detection.[89] Whereas preoperative localization can be achieved in 85 to 90% of cases, sensitivity can be boosted to 95% using the ^{99m}Tc-sestamibi SPECT technique and a three-dimensional, volume-rendered reprojection for visualization.[76]

In one experience, 44% of patients were discharged on the day of MIRP and 83% within 23

hours after the neck exploration.[75] Another author reported 65% of MIRP patients discharged within 5 hours of surgery compared to a mean stay of 1.35 days in traditional, open parathyroidectomy.[85] Further, comparison of charges for traditional open neck exploration revealed lower costs for MIRP owing to decreased duration of the operation, anesthesia, and hospitalization and elimination of the need for histologic evaluation of frozen sections by use of intraoperative rapid PTH assay. Numerous authors concluded that MIRP is a logical approach to primary hyperparathyroidism owing to solitary adenoma.[85,90,91]

In one study of previously operated necks, preoperative ^{99m}Tc-sestamibi scanning and ultrasonography localized only 64% of hyperfunctioning glands. Intraoperative ^{99m}Tc-sestamibi scanning was found in several studies to localize correctly at least 91% of hyperfunctioning glands. Intraoperative sestamibi scan was recommended as a helpful and effective guide for exploration in necks with previous parathyroidectomy.[65,75]

Recently, intraoperative assessment of adequacy of parathyroidectomy has been made by intraoperative radioactivity ratios. Sestamibi injection is administered 3.5 to 1.5 hours preoperatively. After excision, tissue samples are counted, and a simple percentage of background radioactivity is calculated. Whereas other tissues contained at most 2% of background, hyperplastic parathyroids contained 7.5% ± 0.8%, with a maximum of 16%. Adenomas contained 59% ± 9%, with a range from 18 to 136%. This method eliminates the necessity of histopathologic evaluation of frozen sections of specimens.[92]

With successful localization, minimally invasive surgery with or without video assistance can be performed feasibly under local or general anesthesia with a high success rate. For multiglandular disease, however, neither rapid intraoperative PTH assay nor the sestamibi scan is as accurate. Even in difficult cases of reoperation, the overall success rate has not improved despite these techniques. Nonetheless, it appears that MIRP will remain important as an established though evolving technique for parathyroidectomy, especially for parathyroid adenomas.[6,92]

Complications of Thyroid Surgery

Complications of thyroid surgery and parathyroidectomy are best avoided by a thorough working knowledge of the anatomy and embryology of the visceral compartment of the neck, careful preoperative planning, and excellent surgical technique.[93] Although numerous reports advocate subtotal thyroidectomy, total lobectomy (uni- or bilateral) is recommended in most situations. Although improved local control of disease is achieved with more radical procedures, no convincing evidence of improved survival exists.[94] General anesthesia is most commonly used, and it is important for the surgeon to be present at induction to secure an airway by emergency open bronchoscopy or tracheostomy if the patient has respiratory embarrassment owing to a large goiter. A trend exists toward increased use of local anesthesia and outpatient surgery for thyroid procedures. No significant difference is found in the rate of major complications for hemithyroidectomy performed under local anesthesia as compared with general anesthesia.[95]

Thyroid operations including total thyroidectomy have been performed in large numbers as outpatient or short-stay procedures in several centers in the United States and Europe. Extensive perioperative patient teaching is required for the layperson to recognize and react properly to the symptoms and signs of surgical complications. More complex procedures such as reoperation, neck dissection, and sternal splits are more safely managed on an inpatient basis.[96]

Good exposure is an absolute necessity. Collar incisions allow the best cosmetic result for the exposure needed. In some cases, exposure is improved by division of the strap muscles approximately halfway between the hyoid bone and the sternum. This avoids injury to their innervation from the ansa cervicalis. Early in the operation, the trachea should be exposed caudal to the isthmus so emergency access to the airway can be obtained in the event of obstruction.

Exposure of the recurrent laryngeal nerve is mandatory when a lateralized procedure is performed. The nerve should be preserved from the thoracic inlet to its entry into the larynx at the cricothyroid membrane. Unilateral laryngeal nerve injury is reported in 2.3% of cases, largely because of lack of nerve identification.[97] One must remember that at least 1 in 400 cases may not have a recurrent nerve. An anomalous inferior laryngeal nerve occurs almost always on the right and is always associated with an anomalous retroesophageal subclavian

artery, that is, a right subclavian artery arising from the left descending aorta. In these cases, the inferior laryngeal nerve passes directly from the vagus to the larynx. Digital subtraction angiography can confirm this anomaly and is recommended in any reoperation on the right side of the thyroid, especially if the nerve was not identified in the previous surgery.[98]

Iatrogenic injury to the superior laryngeal nerve occurs in 2.4% of patients undergoing surgery in the thyroid periapical region. Such injury can cause temporary or permanent change in voice quality or deglutition. Mobilization of the viscerovertebral angle can provide access to the superior laryngeal nerve, and use of a nerve stimulator can be made in difficult situations to minimize the risk of this complication.[68]

If reoperation is required within 2 weeks, anatomic landmarks will be obliterated because of the inflammatory response. Blunt dissection should be used, although if the reoperation is months to years later, fibrosis often requires sharp dissection. Operating away from the primary site of surgery is recommended. The risk of morbidity is roughly doubled in a second thyroidectomy procedure.

Hemorrhage is usually not a problem. The lowest thyroid artery is divided and ligated first and then the middle thyroid vein if present. The inferior thyroid artery is best divided near the capsule of the gland to prevent injury to the sympathetic trunk (with a resultant Horner's syndrome) and to preserve blood supply to the parathyroid glands. On the left, the thoracic duct can be injured, with resulting chyle fistula. The safest way to divide the superior laryngeal artery and vein is with careful downward traction and skeletonization of the vessels. Caution is advisable if retraction stretches or places pressure on the carotid artery because this has been associated with embolic occlusion of the retinal artery.[99]

Infection occurs in less than 1% of thyroidectomies; however, suction drainage is traditionally recommended to prevent hematoma. Some surgeons have experienced no increased risk by not placing drains or using drainage only in select cases.[69,70] Tracheostomy may be necessary because of tracheal malacia or injury to the recurrent laryngeal nerves. If necessary, a suture can be placed in the second tracheal ring and brought out through the wound as a quick guide to the anterior tracheal wall. Rarely, pneumothorax can occur, and a postoperative chest radiograph is recommended.

Postoperative hemorrhage occurs in fewer than 1% of patients. The resulting hematoma remains in the midline below the strap muscles, creating an increasing pressure and causing collapse of the airway and respiratory embarrassment. In this event, the wound must be opened immediately "wherever you are." Thereafter, the patient can return to the operating room for re-exploration, control of the bleeding vessel, and repeat closure of the wound. Subsequently, edema of the airway can be expected and may require tracheostomy. Tracheostomy may also be necessary in the presence of marked tracheal malacia secondary to large goiter or for incomplete resection of cancer with anticipated postoperative radiation therapy, as in patients with anaplastic carcinoma. The need for elective tracheostomy to prevent postoperative airway obstruction occurs in less than 1% of patients.

During thyroidectomy, the parathyroid glands should be looked for and protected when found. In one-fourth of thyroidectomies, the patient has a transient mild decrease in serum calcium, which is generally asymptomatic. This is probably caused by trauma to the glands and their blood supply. Transient hypocalcemia can be severe in 8% of patients over the first 96 hours postoperatively.[100] Thus, the critical period for calcium monitoring is 24 to 96 hours postoperatively. If calcium infusion is not required within the first 72 hours, it will probably be unnecessary. It is also important to monitor serum magnesium levels postoperatively and correct as necessary.[101] Permanent hypoparathyroidism occurs in up to 8% after total thyroidectomy.[97]

Tetany with hypocalcemia can be treated with calcium gluconate intravenously or calcium carbonate orally in patients without achlorhydria. In patients with achlorhydria, calcium lactate in water and often vitamin D are required.

Myxedema can occur in 4 to 6 weeks after a thyroidectomy, particularly if the gland has been totally ablated. Hashimoto's disease also progresses to hypothyroidism eventually. This can be treated with a variety of drugs for thyroid replacement. Some of these include 2 to 3 g of desiccated thyroid extract, up to 0.2 mg of tetraiodinated levothyroxine sodium (Synthroid), or 25 to 75 μg of triiodinated liothyronine sodium (Cytomel) daily. Treatment by thyroid replacement may be achieved with T_4 alone. However, it has been suggested that partial substitution of T_3 may actually improve

mood and neuropsychological function, especially in patients who still do not feel well despite apparently adequate levels of T_4. Administration of 10 μg of T_3 to enough T_4 to ensure a euthyroid state is recommended.[6] The greatest cause of death in the past was thyroid storm, which can now be effectively controlled. The mortality rate of thyroidectomy today is approximately 0.1%.

Complications of Parathyroid Surgery

Complications of open exploration-type parathyroidectomy are essentially the same as those described for thyroidectomy except for problems arising from ablation of thyroid tissue. For MIRP, regarded as a safe and effective method of parathyroidectomy, the complication rate is described as low. One author reported only minor local complications occurring in 8% of patients.[75] Problems associated with MIRP can include hemorrhage or hematoma, infection, scarring, subcutaneous emphysema (although this is less likely with a gasless endoscopic approach), recurrent laryngeal nerve injury, failure to locate the parathyroid gland of interest (eg, in a previously operated neck), and, conceivably, pneumothorax.

REFERENCES

1. Paschke R, Ludgate M. The thyrotropin receptor in thyroid diseases. N Engl J Med 1997;337:1675–81.
2. DiGeorge AM, LaFranchi S. Disorders of the parathyroid glands. In: Nelson WE, Behrman RE, Kliegman RM, et al, editors. Nelson textbook of pediatrics. 15th ed. Philadelphia: WB Saunders; 1996. p. 1605–12.
3. Nordyke RA, Reppun TS, Madanay LD, et al. Alternative sequences of thyrotropin and free thyroxine assays for routine thyroid function testing. Quality and cost. Arch Intern Med 1998;158:266–72.
4. Kanmaz B, Erdil TY, Yardi OF, et al. The role of ^{99m}Tc-tetrofosmin in the evaluation of thyroid nodules. Nucl Med Commun 2000;21:333–9.
5. Udelsman R, Ball D, Baylin SB, et al. Preoperative localization of occult medullary carcinoma of the thyroid gland with single-photon emission tomography dimercaptosuccinic acid. Surgery 1993;114:1083–9.
6. Numann PJ. Endocrine surgery. J Am Coll Surg 2000;190:129–33.
7. Rosen IB, Walfish PG, Miskin M. The ultrasound of thyroid masses. Surg Clin North Am 1979;59:19–33.
8. Vierhapper H, Hulsmann M. Sonography and scintigraphy are necessary in diagnostics of cystic thyroid lesions. Horm Res 1999;52:256–8.
9. Bloom AD, Adler LP, Shuck JM. Determination of malignancy of thyroid nodules with positron emission tomography. Surgery 1993;114:728–34.
10. Cannon CR, Replogle WH. Fine-needle aspiration: survey of clinical utility. Otolaryngol Head Neck Surg 2000;123:563–5.
11. Bouvet M, Feldman JI, Gill GN, et al. Surgical management of the thyroid nodule: patient selection based on the results of fine-needle aspiration cytology. Laryngoscope 1992;102:1353–6.
12. Bapat RD, Shah SH, Relekar RG, et al. Analysis of 105 uninodular goiters. J Postgrad Med 1992;68:60–1.
13. Carrillo JF, Frias-Mendivil M, Ochoa-Carrillo FJ, Ibarra M. Accuracy of fine-needle aspiration biopsy of the thyroid combined with an evaluation of clinical and radiologic factors. Otolaryngol Head Neck Surg 2000;122:917–21.
14. Jayaram G, Cheah PL, Yip CH. Malignant teratoma of the thyroid with predominantly neuroepithelial differentiation. Fine needle aspiration cytologic, histologic and immunocytochemical features of a case. Acta Cytol 2000;44:375–9.
15. Urben SL, Ransom ER. Fusion of the thyrohyoid interval in a patient with a thyroglossal duct cyst. Otolaryngol Head Neck Surg 1999;120:757–9.
16. Gaitan E, Nelson NC, Poole GV. Endemic goiter and endemic thyroid disorders. World J Surg 1991;15: 205–15.
17. Cohen-Kerem F, Schachter P, Sheinfeld M, et al. Multinodular goiter: the surgical procedure of choice. Otolaryngol Head Neck Surg 2000;122:848–50.
18. Tonacchera M, Vitti P, Agretti P, et al. Functioning and nonfunctioning thyroid adenomas involve different molecular pathogenetic mechanisms. J Clin Endocrinol Metab 1999;84:4155–8.
19. Mra Z, Wax MK. Effects of acute thyroxine depletion on hearing in humans. Laryngoscope 1999;109: 343–50.
20. Gavin LA. Thyroid crisis. Med Clin North Am 1991;75:179–93.
21. Roitt IM, Doniach D, Campbell RN, Hudson RV. Autoantibodies in Hashimoto's disease (lymphadenoid goitre). Lancet 1956;2:820.
22. Kinney JS, Hurwitz ES, Fishbein DB, et al. Community outbreak of thyrotoxicosis: epidemiology, immunogenetic characteristics, and long-term outcome. Am J Med 1988;84:10–8.

23. Hunt PJ, Marshall SE, Weetman AP, et al. Cytokine gene polymorphisms in autoimmune thyroid disease. J Clin Endocrinol Metab 2000;85:1984–8.
24. Rose NR, Burek CL. Autoantibodies to thyroglobulin in health and disease. Appl Biochem Biotechnol 2000;83:245–51.
25. Kuma K, Matsuzuka F, Kobayashi A, et al. Outcome of long standing solitary thyroid nodules. World J Surg 1992;16:583–7.
26. Brander AE, Viikinkoski VP, Nickels JI, Kivisaari LM. Importance of thyroid abnormalities detected at US screening: a 5-year follow-up. Radiology 2000;215: 801–6.
27. McCaffrey TV. Evaluation of the thyroid nodule. Cancer Control 2000;7:223–8.
28. Hung W, Anderson KD, Chandra RS, et al. Solitary thyroid nodules in 71 children and adolescents. J Pediatr Surg 1992;27:1407–9.
29. Frankenthaler RA, Sellin RV, Cangir A, Goepfert H. Lymph node metastasis from papillary-follicular thyroid carcinoma in young patients. Am J Surg 1990;160:341–3.
30. Lawson W, Biller HF. The solitary thyroid nodule: diagnosis and management of malignant disease. Am J Otolaryngol 1983;4:43–73.
31. Habashi S. Benign thyroid adenoma causing recurrent laryngeal nerve palsy in a child. J Laryngol Otol 1991;105:141–2.
32. Santoro M, Thomas GA, Vecchio G, et al. Gene rearrangement and Chernobyl related thyroid cancers. Br J Cancer 2000;82:315–22.
33. Malchoff CD, Sarfarazi M, Tendler B, et al. Papillary thyroid carcinoma associated with papillary renal neoplasia: genetic linkage analysis of a distinct heritable tumor syndrome. J Clin Endocrinol Metab 2000;85:1758–64.
34. Lin JD, Chao TC, Sun JH, Weng HF. Trends in the clinical characteristics of patients with papillary thyroid carcinoma in Taiwan. Oncology 2000;58:280–5.
35. Bunone G, Uggeri M, Mondellini P, et al. RET receptor expression in thyroid follicular epithelial cell-derived tumors. Cancer Res 2000;60:2845–9.
36. D'Antonio A, De Chiara A, Santoro M, et al. Warthin-like tumour of the thyroid gland: RET/PTC expression indicates it is a variant of papillary carcinoma. Histopathology 2000;36:493–8.
37. Prendiville S, Burman KD, Ringel MD, et al. Tall cell variant: an aggressive form of papillary thyroid carcinoma. Otolaryngol Head Neck Surg 2000;122: 352–7.
38. Weiss SD, Orlich CC. Primary papillary carcinoma of a thyroglossal duct cyst: report of a case and literature review. Br J Surg 1991;78:87–9.
39. Cady B. Papillary carcinoma of the thyroid. Semin Surg Oncol 1991;7:81–6.
40. Doi SA, Woodhouse NJ. Ablation of the thyroid remnant and 131I dose in differentiated thyroid cancer. Clin Endocrinol (Oxf) 2000;52:765–73.
41. Nicolosi A, Malloci A, Esu S, et al. The role of node-picking lymphadenectomy in the treatment of differentiated carcinoma of the thyroid. Minerva Chir 1993;48:459–63.
42. Shaha AR, DiMaio T, Webber C, Jaffe BM. Intraoperative decision making during thyroid surgery based on the results of preoperative needle biopsy and frozen section. Surgery 1990;108:964–7.
43. Almeida-Filho P, Ravizzini GC, Almeida C, Borges-Neto S. Whole-body Tc-99m sestamibi scintigraphy in the follow-up of differentiated thyroid carcinoma. Clin Nucl Med 2000;25:443–6.
44. Black EG, Sheppard MC. Serum thyroglobulin measurements in thyroid cancer: evaluation of "false" positive results. Clin Endocrinol (Oxf) 1991; 35:519–20.
45. Lind P, Kumnig G, Matschnig S, et al. The role of F-18FDG PET in thyroid cancer. Acta Med Austriaca 2000;27:38–41.
46. McLeod MK, Thompson NW. Hürthle cell neoplasms of the thyroid. Otolaryngol Clin North Am 1990;23:441–52.
47. Engelbach M, Gorges R, Forst T, et al. Improved diagnostic methods in the follow-up of medullary thyroid carcinoma by highly specific calcitonin measurements. Clin Endocrinol Metab 2000;85: 1890–4.
48. Papotti M, Volante M, Komminoth P, et al. Thyroid carcinomas with mixed follicular and C-cell differentiation patterns. Semin Diagn Pathol 2000;17:109–19.
49. Wermer P. Endocrine adenomatosis and peptic ulcer in a large kindred. Am J Med 1963;35:205.
50. Block MA, Horn RC, Miller JM, et al. Familial medullary carcinoma of the thyroid. Ann Surg 1967;166:403–12.
51. Sipple JH. The association of pheochromocytoma with carcinoma of the thyroid gland. Am J Med 1961;31:163.
52. Decker RA, Toyama WM, O'Neal LW, et al. Evaluation of children with multiple endocrine neoplasia type IIb following thyroidectomy. J Pediatr Surg 1990;25:939–43.

53. Telander RL, Zimmerman D, Sizemore GW, et al. Medullary carcinoma in children: results of early detection and surgery. Arch Surg 1989;124:841–3.
54. Hay ID, Ryan JJ, Grant CS, et al. Prognostic significance of nondiploid DNA determined by flow cytometry in sporadic and familial medullary thyroid carcinoma. Surgery 1990;108:972.
55. Decker RA. Expression of papillary thyroid carcinoma in multiple endocrine neoplasia type 2A. Surgery 1993;114:1059–63.
56. Lampertico P. Anaplastic (sarcomatoid) carcinoma of the thyroid gland. Semin Diagn Pathol 1993; 10:159–68.
57. Wilkens L, Benten D, Tchinda J, et al. Aberrations of chromosomes 5 and 8 as recurrent cytogenetic events in anaplastic carcinoma of the thyroid as detected by fluorescence in situ hybridisation and comparative genomic hybridisation. Virchows Arch 2000;436:312–8.
58. Ekman ET, Lundell G, Tennvall J, Wallin G. Chemotherapy and multimodality treatment in thyroid carcinoma. Otolaryngol Clin North Am 1990; 23:523–7.
59. Gemsenjager E. Significance of the limiting lamella in struma surgery. Helv Chir Acta 1993;59:815–8.
60. Sellers RS, Schuller DE, Sharma PK, et al. Head and neck squamous cell carcinoma: measurement of plasma parathyroid hormone-related protein and serum and urine calcium concentrations. Otolaryngol Head Neck Surg 2000;123:558–62.
61. Lumachi F, Zucchetta P, Angelina F, et al. Tumors of the parathyroid glands. Changes in clinical features and in noninvasive localization studies sensitivity. J Exp Clin Cancer Res 2000;19:7–11.
62. Kriskovich MD, Holman JM, Haller JR. Calciphylaxis: is there a role for parathyroidectomy? Laryngoscope 2000;110:603–7.
63. Silverberg SJ, Bone HG, Marriott TB, et al. Short-term inhibition of parathyroid hormone secretion by a calcium-receptor agonist in patients with primary hyperparathyroidism. N Engl J Med 1997;337: 1506–10.
64. Fraker DL. Update on the management of parathyroid tumors. Curr Opin Oncol 2000;12:41–8.
65. Rossi HL, Ali A, Prinz RA. Intraoperative sestamibi scanning in reoperative parathyroidectomy. Surgery 2000;128:744–50.
66. Batsakis JG. Parenchymal cysts of the neck. In: Batsakis JG, editor. Tumors of the head and neck. 2nd ed. Baltimore: Williams and Wilkins; 1982. p. 234–8..
67. Lore JM. Surgery of the thyroid gland. Otolaryngol Clin North Am 1980;13:69–83.
68. el-Guindy A, Abdel-Aziz M. Superior laryngeal nerve preservation in peri-apical surgery by mobilization of the viscerovertebral angle. J Laryngol Otol 2000; 114:268–73.
69. Shaha AR, Jaffe BM. Selective use of drains in thyroid surgery. J Surg Oncol 1993;52:241–3.
70. Ariyanayagam DC, Naraynsingh V, Busby D, et al. Thyroid surgery without drainage: 15 years of clinical experience. J R Coll Surg Edinb 1993;38:69–70.
71. Rodriguez JM, Hernandez Q, Piñero A, et al. Substernal goiter: clinical experience of 72 cases. Ann Otol Rhinol Laryngol 1999;108:501–4.
72. Wells SA, Gunnells JC, Shelburne JD, et al. Transplantation of the parathyroid glands in man: clinical indications and results. Surgery 1975;78:34–44.
73. Shaha AR, Burnett C, Jaffe BM. Parathyroid autotransplantation during thyroid surgery. J Surg Oncol 1991;46:21–4.
74. Funahashi H, Satoh Y, Imai T, et al. Our technique of parathyroid autotransplantation in operation for papillary thyroid carcinoma. Surgery 1993;114:92–6.
75. Flynn MB, Bumpous JM, Schill K, McMasters KM. Minimally invasive radioguided parathyroidectomy. J Am Coll Surg 2000;191:24–31.
76. Moka D, Voth E, Dietlein M, et al. Technetium 99m-MIBI-SPECT: a highly sensitive diagnostic tool for localization of parathyroid adenomas. Surgery 2000;128:29–35.
77. Brunt LM, Jones DB, Wu JS, et al. Experimental development of an endoscopic approach to neck exploration and parathyroidectomy. Surgery 1997;122:893–901.
78. Naitoh G, Gagner M, Garcia-Ruiz A, Heniford BT. Endoscopic endocrine surgery in the neck. An initial report of endoscopic subtotal parathyroidectomy. Surg Endosc 1998;12:202–5.
79. Yeung GH, Ng JH. The technique of endoscopic exploration for parathyroid adenoma of the neck. Aust N Z J Surg 1998;68:147–50.
80. Ikeda Y, Takami H, Sasaki Y, et al. Endoscopic neck surgery by axillary approach. J Am Coll Surg 2000;191:336–40.
81. Howe JR. Minimally invasive parathyroid surgery. Surg Clin North Am 2000;80:1399–426.
82. Miccoli P, Bendinelli C, Vignali E, et al. Endoscopic parathyroidectomy: report of an initial experience. Surgery 1998;124:1077–9.

83. Henry JF, Defechereux T, Gramatica L, de Boissezon C. Minimally invasive videoscopic parathyroidectomy by lateral approach. Langenbecks Arch Surg 1999;384:298–301.
84. Gauger PG, Reeve TS, Delbridge LW. Endoscopically assisted, minimally invasive parathyroidectomy. Br J Surg 1999;86:1563–6.
85. Goldstein RE, Blevins L, Delbeke D, Martin WH. Effect of minimally invasive radioguided parathyroidectomy on efficacy, length of stay, and costs in the management of primary hyperparathyroidism. Ann Surg 2000;231:732–42.
86. Kountakis SE, Maillard AJ. Parathyroid adenomas: is bilateral neck exploration necessary? Am J Otolaryngol 1999;20:396–9.
87. Fujii H, Kubo A. Parathyroid imaging with technetium-99m sestamibi. Biomed Pharmacother 2000; 54 Suppl 1:12s–6s.
88. Norman J. The technique of intraoperative nuclear mapping to facilitate minimally invasive parathyroidectomy. Cancer Control 1997;4:500–4.
89. Kacker A, Scharf S, Komisar A. Reduced-time-window sestamibi scanning for nonlocalized primary hyperparathyroidism. Otolaryngol Head Neck Surg 2000;123:456–8.
90. Kumar A, Cozens NJ, Nash JR. Sestamibi scan-directed unilateral neck exploration for primary hyperparathyroidism due to a solitary adenoma. Eur J Surg Oncol 2000;26:785–8.
91. Dackiw AP, Sussman JJ, Fritsche HA, et al. Relative contributions of technetium Tc99m sestamibi scintigraphy, gamma probe detection, and the rapid parathyroid hormone assay to the surgical management of hyperparathyroidism. Arch Surg 2000;135: 550–7.
92. Murphy C, Norman J. The 20% rule: a simple, instantaneous radioactivity measurement defines cure and allows elimination of frozen sections and hormone assays during parathyroidectomy. Surgery 1999;126:1023–8.
93. Lore JM. Complications in management of thyroid cancer. Semin Surg Oncol 1991;7:120–5.
94. Staunton MD, Bourne H. Thyroid cancer in the 1980s: a decade of change. Ann Acad Med Singapore 1993;22:613–6.
95. Hochman M, Fee WE. Thyroidectomy under local anesthesia. Arch Otolaryngol Head Neck Surg 1991;117:405–7.
96. Lo-Gerfo P, Gates F, Gazetas P. Outpatient and short-stay thyroid surgery. Head Neck 1991;13: 97–101.
97. Herranz-Gonzalez J, Gavilan J, Matinez-Vidal J, Gavilan C. Complications following thyroid surgery. Arch Otolaryngol Head Neck Surg 1991;117:516–8.
98. Campbell PR, Serpell JW, Young AE. Non-recurrent laryngeal nerves: the role of digital subtraction angiography to identify subjects. Aust N Z J Surg 1991;61:358–9.
99. Kihara S, Ohtuka H, Ikenaga T, et al. Branch retinal artery occlusion following thyroidectomy for papillary carcinoma of the thyroid: report of a case. Surg Today 1993;23:750–4.
100. Bourrel C, Uzzan B, Tison P, et al. Transient hypocalcemia after thyroidectomy. Ann Otol Rhinol Laryngol 1993;102:496–501.
101. Szubin L, Kacker A, Kakani R, et al. The management of post-thyroidectomy hypocalcemia. Ear Nose Throat J 1996;75:612–4.

SUGGESTED READING

DiGeorge AM, LaFranchi S. Section 2: disorders of the thyroid gland. Section 3: disorders of the parathyroid glands. In: Nelson WE, Behrman RE, Kliegman RM, Arvin AM, editors. Nelson's textbook of pediatrics. 15th ed. Philadelphia: WB Saunders; 1996. p. 1587–1605.

Falk SA, editor. Thyroid disease: endocrinology, surgery, nuclear medicine, and radiotherapy. 2nd ed. Baltimore: Lippincott Williams & Wilkins; 1997.

Schoem SR, Khan A, Wenig BM. Evaluation and management of the thyroid nodule: a practical approach. Alexandria (VA): Continuing Education Program of the American Academy of Otolaryngology-Head and Neck Surgery Foundation, Inc.; March 1998.

Singer PA, Cooper DS, Daniels GH, et al. Treatment guidelines for patients with thyroid nodules and well-differentiated thyroid cancer. American Thyroid Association. Arch Intern Med 1996;156: 2165–72. http://www.ama-assn.org/sci-pubs/journals/archive/inte/vol_156/no_19/60162.htm.

CHAPTER 64

Laser Surgery in the Head and Neck

Lou Reinisch, PhD, Robert H. Ossoff, MD

The use of lasers in otolaryngology has resulted in the addition of new surgical procedures and the refinement of conventional surgical procedures. The carbon dioxide laser is still the most commonly used laser for surgery of the upper aerodigestive tract. The use of the carbon dioxide laser for treating lesions of the oral cavity, pharynx, larynx, and tracheobronchial tree has been facilitated by previous experience with the techniques of endolaryngeal microsurgery.[1] The introduction of the microspot for carbon dioxide laser microsurgery has further refined the use of the carbon dioxide laser for endoscopic surgery.[2,3] The superpulsed carbon dioxide laser and the argon and the potassium titanyl phosphate/532 (KTP/532) lasers are being used for stapedotomy with increasing frequency.

The neodymium:yttrium-aluminum-garnet (Nd:YAG) laser, with and without contact probes, continues to be used for obstructing lesions of the tracheobronchial tree[4,5] and for cavernous hemangiomas of the head and neck.[6] The use of the Nd:YAG contact probes for head and neck surgery is still being explored. The KTP/532 laser has been used by many otolaryngologists for pharyngeal tonsillectomy, although there appear to be no major benefits of using the laser over conventional techniques in routine tonsillectomies.[7–10] The continuous and flash-pumped yellow dye lasers are used to treat port-wine stains with excellent results.[11–13] The concept of photodynamic therapy (PDT) to treat malignancies of the upper aerodigestive tract is still under active investigation.[14–34]

It is apparent that many lasers are being used in otolaryngology for a wide variety of surgical diseases. This chapter focuses on these lasers and their potential uses. The laser is not a panacea, but when it is used to advantage for a specific disease process, an improvement in patient care may be expected.

LASER BIOPHYSICS

LASER is an acronym for light amplification by the stimulated emission of radiation. A laser is an optical device containing three fundamental parts. First, there is the laser medium. In the first laser made, the laser medium was a ruby crystal.[35] Second, there is an excitation source. In the first laser made, the excitation source was the flash from a camera. Third, one needs two mirrors to provide the optical feedback, involving a laser medium, an excitation source, and a resonant cavity.[36,37]

What happens in the lasing process can be explained in the following series of steps. The lasing medium (eg, the ruby crystal) is placed between two parallel mirrors facing each other, as shown in Figure 64–1. The lasing medium is then excited. The excitation, perhaps from the flash lamp or from an electrical arc, will promote the molecules or atoms in the lasing medium from the ground state up to an excited state.

Atoms or molecules in an excited state will return to the lower energy state in the process that is called "spontaneous emission." In spontaneous emission, the excited atoms or molecules spontaneously decay from their excited state to their lower energy states. The energy is emitted as light. The emission is in all directions, and some of the emission can be reflected back into the laser medium.

If the light is reflected by the mirror and reenters the laser medium, the light can then stimulate additional radiation. In other words, as the light travels through the lasing medium, if it finds a mol-

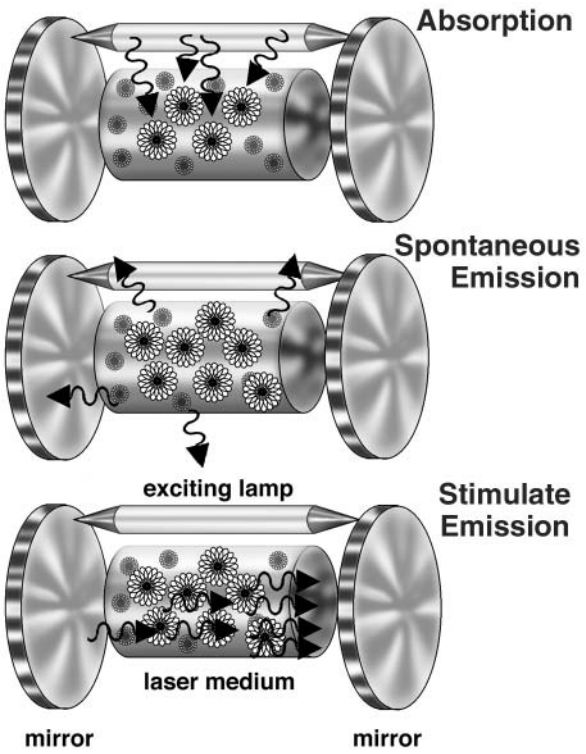


FIGURE 64–1. Schematic illustration of the laser, showing the mirrors, the laser medium, and the exciting lamp. Light from the lamp can excite atoms in the medium through the process of absorption. The excited atoms can return to their ground state and emit the absorbed energy as light in a process called spontaneous emission. If emitted light interacts with other atoms in the excited state, it will stimulate them to return to their ground states and emit light in a process called stimulated emission. The amplification of the light by stimulated emission is the basis for the laser.

ecule or an atom in an excited state, it will tickle or stimulate that atom to release its energy and make the transition to the ground state. So one wave of light becomes two waves of light. Both of these waves of light continue to propagate through the lasing medium. To achieve a net amplification of the light, more atoms or molecules must be in the excited state than those in the ground state. This is termed a "population inversion."

The stimulated emission continues as the light reaches the front mirror and then is reflected back through the laser medium yet another time. The stimulated light is emitted in the same direction as the incident light, so all of the rays of light are parallel. The light continues to be amplified by the stimulated emission as long as more atoms or molecules are in the excited state than the ground state. The "population inversion" is crucial to the lasing process. These events provide the origin of the laser acronym.

The laser light is then released from the resonant cavity by a partially reflecting mirror. A fraction of the light is transmitted out of the laser cavity while the remaining fraction of the light stays in the cavity to maintain the lasing process. The lasing process continues in the resonant cavity as long as the excitation source keeps the molecules or atoms in the laser medium excited. This light is emitted either as a continuous wave or as a pulse depending on whether the excitation is continuous or pulsed.

Typically, surgical lasers have an optical resonating chamber (cavity) with two end mirrors; the space between these mirrors is filled with a lasing medium such as argon, Nd:YAG, or carbon dioxide.[38] The radiant energy emitted is extremely intense and unidirectional or collimated.[39] The light is coherent, both temporally and spatially. Temporal coherence means that the photons alternate sinusoidally in phase with one another, whereas spatial coherence means that the photons are equal and parallel across the wave front. These properties of monochromaticity, intensity, collimation, and coherence distinguish the organized radiant energy of a laser from the disorganized radiant energy of a light bulb.

After emission of the radiant energy through the partially reflective mirror, the laser beam can be passed through a fiber. The fiber allows transmission of the laser energy at the distal end of the fiber. The total internal reflection of the optical fiber causes the collimation of the laser beam to be lost. The result is a divergent beam at the fiber tip.

The laser beam can also be passed through a lens that focuses the radiant energy to a small beam waist, or spot size, ranging from approximately 0.1 to 2.0 mm. The focusing lens is usually treated with special optical coatings to allow the helium-neon aiming beam and the invisible carbon dioxide beam to be focused coplanar. This lens is characterized by a specific focal length that determines the distance from the lens to the target tissue for focused use.

Surgical lasers typically permit end-user control of three variables: (1) focal length spot size

(mm), (2) power (watts), and (3) exposure time (seconds). Power, measured in watts, may be kept constant with widely varying effects, depending on the spot size and duration of exposure. For example, if the time of exposure is kept constant, the relationship between power and depth of tissue injury becomes logarithmic as the spot size is varied.

Power density or irradiance, on the other hand, is a useful way of monitoring the surgical effects of the laser. It is a measure of the power output of the laser in watts over the cross-sectional area of the laser beam in square centimeters (in the equation below, r is the radius of the spot):

$$\text{Power density} = \text{Power (watts)}/\pi \times r^2\ (\text{cm}^2)$$

Furthermore, power density is a measure of the concentration of radiant energy of the laser beam. Here, if the time of exposure is kept constant, the relationship between power density and depth of injury is linear as the spot size is varied. Therefore, surgeons should calculate the appropriate power density for various procedures; such calculations allow the surgeon to control tissue effects predictably when changing from one focal length to another or when using different surgical lasers.

Radiant exposure (RE) or fluence expressed as joules/cm^2 is equal to the power density (watts/cm^2) multiplied by the time of exposure in seconds:

$$\text{RE} = \text{power density} \times \text{time}$$

This is a measure of the total amount of laser energy per unit area of exposed target tissue.

Carbon dioxide lasers can emit radiant energy with different beam configurations that ultimately determine the depth of tissue damage and vaporization across the spot size. Transverse electromagnetic mode (TEM) refers to the intensity distribution over the spot area and determines the shape of the focused laser spot. The most fundamental TEM is TEM_{00}. In this mode, if the laser beam is cut in cross-section, it appears circular, with the power density of the beam following a gaussian distribution (Figure 64–2, A). In other words, the power density is greatest at the center of the beam and diminishes progressively toward the periphery. Thus, the center of the beam can be used for vaporization and the periphery for coagulation.

$TEM_{01}{}^{\star}$ (Figure 64–2, B) and $TEM_{11}{}^{\star}$ (Figure 64–2, C) have a more complex distribution of energy and cause characteristic variations in tissue ablation depth. In fact, there are cold spots in the center of their beams. Additionally, they cannot be focused to as small a spot size as TEM_{00} lasers. Most carbon dioxide surgical lasers manufactured in the last 10 years produce a TEM_{00} beam profile.

Although simple ray diagrams normally show parallel light to be focused to a point, the actual situation is more complicated. A lens will focus a gaussian beam to a beam waist or a finite size. This beam waist is the minimum spot diameter, d, and can be written as

$$d \sim \frac{2 f \lambda}{D}$$

where f is the focal length of the lens, λ is the wavelength of light, and D is the diameter of the laser beam incident on the lens (Figure 64–3). The beam waist occurs not at one distance from the lens but over a range of distances. This range is termed the depth of focus and can be written as

$$\text{depth of focus} \sim \frac{\pi d^2}{2 \lambda}$$

We realize the depth of focus every time we focus a camera. With a camera, a range of objects is in focus, and we can set the focus without carefully measuring the distance between the object and the lens. Notice from the above equations that a long focal length lens (a large f) leads to a large beam waist. A large beam waist also translates as a large depth of focus.

The size of the laser beam on the tissue (spot size) can therefore be varied in two ways. Since the minimum beam diameter of the focal spot increases directly with increasing the focal length of the laser-focusing lens, the surgeon can change the focal length of the lens to obtain a particular beam diameter. As the focal length becomes smaller, there is a corresponding decrease in the size of the focal spot; also, the smaller the spot size is for any given power output, the greater the corresponding power density. The second way the surgeon can vary the spot size is by working either in or out of focus. The minimum beam diameter and highest power concentration occur at the focal plane, where much of the precise cutting and vaporization are carried out (Figure 64–4, A). As the distance from the focal plane increases, the laser beam diverges or becomes defocused (Figure 64–4, B). Here the cross-sectional

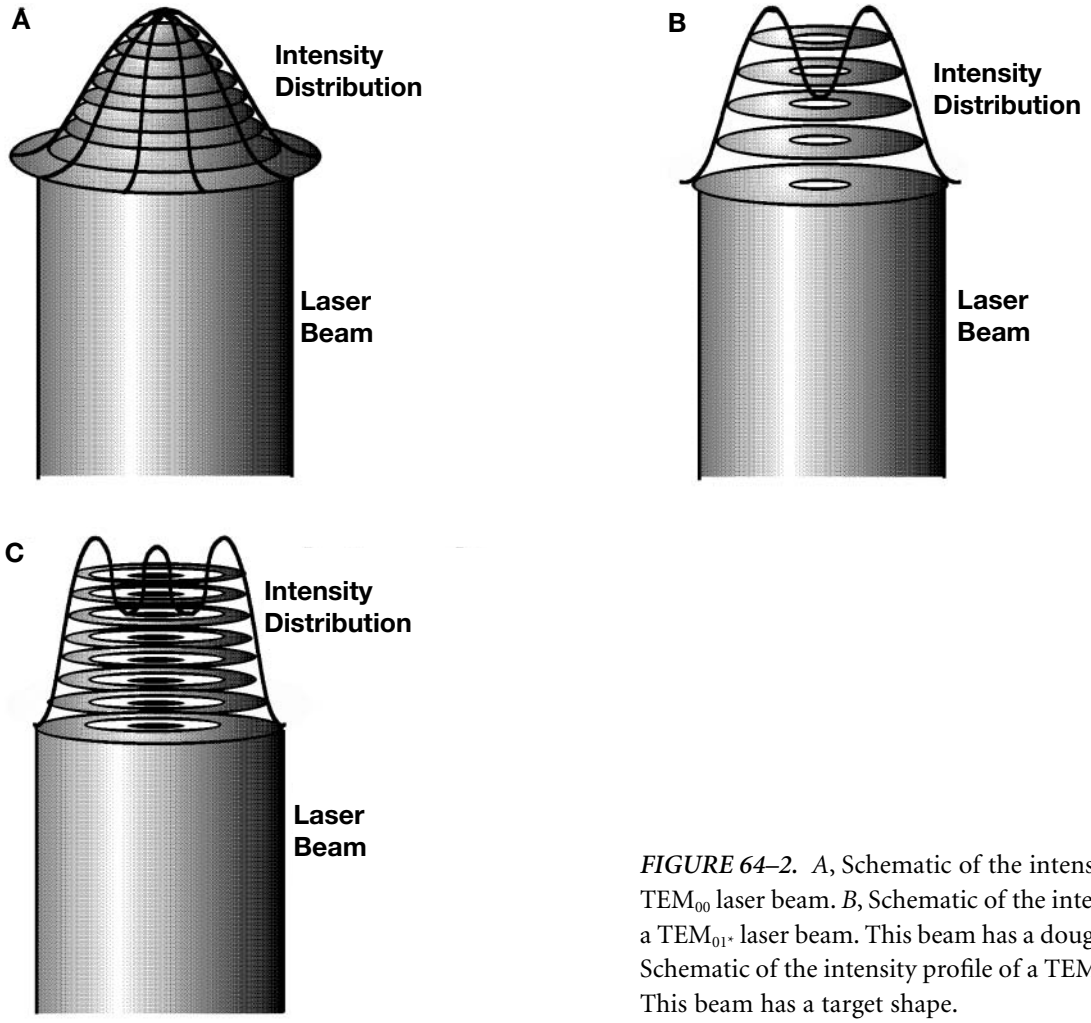


FIGURE 64–2. *A*, Schematic of the intensity profile of a TEM_{00} laser beam. *B*, Schematic of the intensity profile of a TEM_{01^*} laser beam. This beam has a doughnut shape. *C*, Schematic of the intensity profile of a TEM_{11^*} laser beam. This beam has a target shape.

area of the spot grows larger and thus lowers the power density for a given output. As one can readily see, the size of the focal spot depends on both the focal length of the laser lens and whether the surgeon is working in or out of focus.

Figure 64–5 demonstrates these concepts using arbitrary ratios accurate for a current model TEM_{00} carbon dioxide laser. The laser lens setting (focal length) and working distance (focus/defocus) combinations shown here determine the spot size of the laser beam. The height of the various cylinders represents the relative fluence of the light incident on the tissue.

The equations for power density and radiant exposure demonstrate that the surgeon can vary the amount of energy delivered to the target tissue by varying the power output, cross-section area of the beam, and time of exposure. The power output of the typical medical carbon dioxide laser can be varied between 1 and 100 watts. Tissues with less water content require a greater power output to achieve the desired effect.

The use of a fiber to deliver the laser energy to tissue causes the laser beam to lose its collimation (unidirectionality). The result is that the closer the fiber tip is to the tissue, the smaller is the spot size or area of laser energy distribution. As the fiber is moved away from the tissue, the spot size will increase, causing a dilution of the laser energy over the larger area.

The third means by which the surgeon can vary the amount of energy delivered to the target tissue is by varying the exposure time. The longer the tissue is exposed to the laser beam, the greater the amount of radiant exposure (joules/cm^2) that is delivered to the target tissue. The time can be varied by working in either the continuous or the pulsed mode, with the pulse duration ranging from 0.05 to 0.5 second.

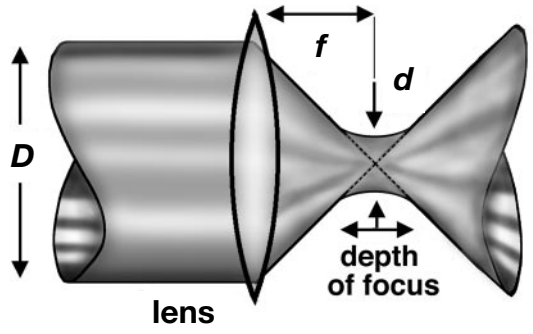


FIGURE 64–3. Schematic of the diffraction limited spot size or beam waist, d, of a parallel beam of light of diameter, D, focused by a lens with a focal length of f.

Laser energy can also be delivered over a short time duration, from milliseconds to femtosecond ranges. A continuous-wave laser can be "pulsed" in the millisecond range. A noncontinuous or pulsed laser can deliver true pulses of laser energy in a duration as short as several femtoseconds (10^{-15} seconds) (Table 64–1). The pulsed laser delivers energy with high peak powers. One needs to visualize a train of pulses with high peaks but short time durations (Figure 64–6). The high peaks are the peak intensities in watts that can be obtained by pulsed lasers. The average power delivered to tissue is expressed by the pulses/second times peak power times the pulse width. On a continuous-wave laser, the meter reads in watts. On a pulsed laser, however, the meter reads in joules (watts times seconds). Pulsed lasers and superpulsing of a continuous-wave laser permit the laser to produce a high peak power. A pulsed Nd:YAG ophthalmologic laser can produce 6.7×10^6 watts per pulse. The pulse duration can be as short as 30×10^{-12} seconds. If we remember that power density or irradiance is the number of watts divided by the area, we can see that the pulsing of the laser permits high wattage with resultant very high-power densities. These high-power densities can cause nonthermal tissue interactions. The flash-pumped yellow dye laser is an example of a pulsed laser used in otolaryngology to treat port-wine stains. Some carbon dioxide lasers and the KTP/532 laser deliver laser energy in a quasicontinuous wave or pulsed mode.

In summary, by varying the power output of the laser unit, the spot size or cross-sectional area of the beam, or the time of exposure, the surgeon can adjust the laser to incise, coagulate, or vaporize the target tissue.

TISSUE EFFECTS

The actual tissue effects produced by the radiant energy of a laser vary with the specific wavelength of the laser used. The interaction of laser energy with

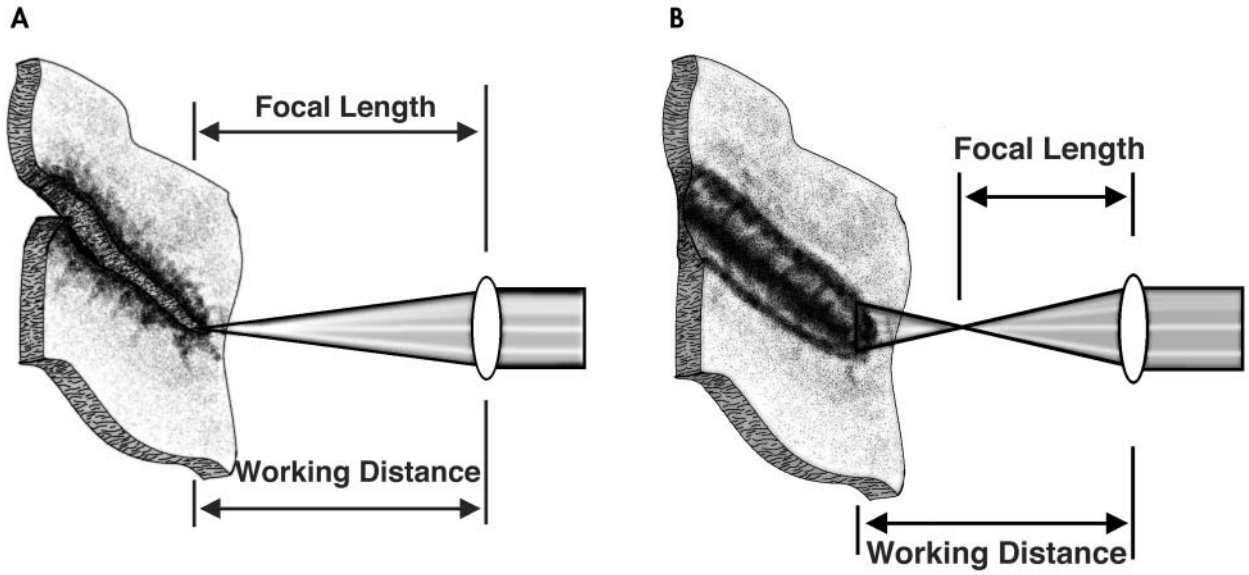


FIGURE 64–4. A, Schematic illustration of the laser incising tissue with the working distance equal to the focal length of the lens. *B*, Schematic illustration of the laser excising tissue with the working distance not equal to the focal length of the lens.

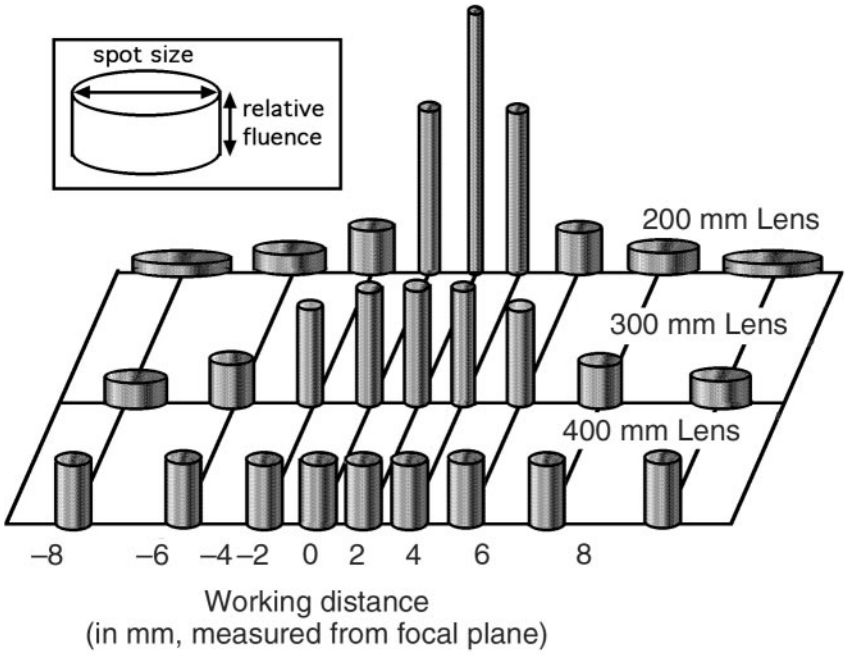


FIGURE 64–5. Graph of the fluence and spot size for three different focal length lenses, working at and near the focal length of the lens. The cylinder diameter is proportional to the spot diameter for a current model carbon dioxide laser. The height of the cylinder is proportional to the fluence of the laser. The lowest fluence has arbitrarily been set equal to one.

living tissue can produce three distinct reactions. First, the laser energy can be absorbed by chromophores within the tissue. The resulting effect can be the production of heat. This is the thermal effect seen in most conventional laser systems in use today.

Second, the radiant energy of a laser can stimulate or react with specific molecules within a cell. This reaction can cause a chemical change to occur within the cell. This effect is termed photochemical. An example is the reaction that occurs with injection of a photosensitizing drug into tissue and the biochemical effect that is produced when the drug is activated by the stimulating effect of radiant laser energy.

Third, the use of short pulses of high-wattage laser energy can disrupt cellular architecture because of the production of sound waves or photoacoustic shock waves.[40–42] The mechanical disruption of tissue is an example of a nonthermal tissue effect.

The radiant energy of a laser can interact with tissue in four ways. The radiant laser energy can be transmitted through the tissue with little of the energy absorbed by the surrounding tissue. The first interaction is transmission. The transmission of laser energy can be understood by remembering the effect of argon laser energy on water. The energy of the argon laser passes through or is transmitted through clear liquids without being absorbed significantly. The second interaction is absorption. The absorption of laser energy by tissue is a complex process depending on the laser wavelength and the color of the tissue. The length of the laser wavelength and the amount of color have an effect on energy absorption. The carbon dioxide laser is essentially colorblind. The visible lasers, however, are color dependent in energy absorption. The third interaction is scattering. Laser energy is also scattered for-

TABLE 64–1. Time Scales

1 second (s)					
1 millisecond (ms)	10^{-3} s				
1 microsecond (μs)	10^{-3} ms	10^{-6} s			
1 nanosecond (ns)	10^{-3} μs	10^{-6} ms	10^{-9} s		
1 picosecond (ps)	10^{-3} ns	10^{-6} μs	10^{-9} ms	10^{-12} s	
1 femtosecond (fs)	10^{-3} ps	10^{-6} ns	10^{-9} μs	10^{-12} ms	10^{-15} s

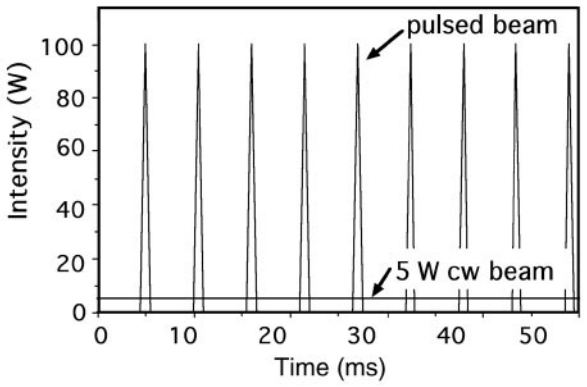


FIGURE 64–6. Schematic of the temporal profile of a continuous-wave laser and a pulsed laser. The continuous-wave laser produces a constant 5 W laser output. The pulsed laser produces pulses of light, each 100 W in peak intensity and on for 5 ms with a repetition rate of approximately 18 Hz.

ward into tissue and backward away from tissue. The scattering coefficient depends on the wavelength. The shorter the wavelength, the more strongly scattered is the laser energy. Therefore, the laser energy of visible lasers is more strongly scattered than for lasers in the infrared region of the electromagnetic spectrum. The last laser tissue effect is reflection of laser energy from the tissue surface in addition to limiting the penetration of the laser into the tissue.

The radiant energy of a laser produces a thermal effect when it is absorbed by the tissue and converted to heat.[43–45] When the target absorbs a specific amount of radiant energy to raise its temperature to approximately 60° to 65°C, protein denaturation starts to occur. Blanching of the tissue surface is readily visible, and the deep structural integrity of the tissue is disturbed. When the absorbed laser light heats the tissue to approximately 100°C, vaporization of intracellular water occurs. This causes vacuole formation, cratering, and tissue shrinkage (Figure 64–7). Carbonization, disintegration, smoke, and gas generation with destruction of the laser-radiated tissue occur at several hundred degrees Celsius.

If the radiant energy is poorly absorbed by the target tissue, excess thermal damage to adjacent tissues occurs.[44,45] Reflection of the energy at the tissue surface or transmission of poorly absorbed energy through the target tissue makes it necessary to prolong the time of exposure of the target tissue to the laser energy to achieve the desired ablation. This prolonged exposure of the target tissue causes an increased thermal effect with resultant damage to the surrounding nontarget tissue.

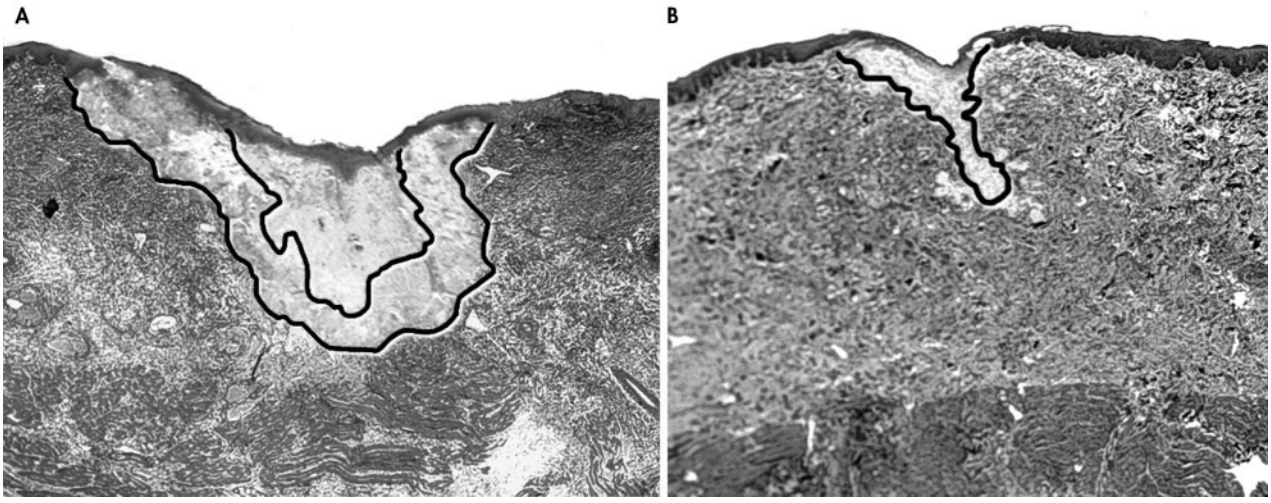


FIGURE 64–7. Laser incisions of the oral mucosa of a dog 1 week after the incisions were made. *A* shows the incision that was made with a continuous-wave carbon dioxide laser (5 W, 0.2 s repeat pulses). The large area outlined is thermally denatured collagen. The central outlined area is a region containing neutrophils still within the inflammatory phase of wound healing. *B* shows the incision that was made with the pulsed carbon dioxide laser (60 μs, 0.5 J/pulse, 10 Hz repetition rate). There is no thermally denatured collagen still present. The outlined area is a region containing neutrophils still within the inflammatory phase of wound healing. The stain is Masson trichrome.

In the center of the wound is a volume of tissue vaporization; here just a few flakes of carbon debris are noted. Immediately adjacent to this volume is a zone of thermal necrosis measuring approximately 100 μm wide. Next is a volume of thermal conductivity and repair, usually 300 to 500 μm wide. Small vessels, nerves, and lymphatics are sealed in the zone of thermal necrosis; the minimal operative trauma combined with the vascular seal probably accounts for the notable absence of postoperative edema characteristic of laser wounds.

Studies comparing the histologic properties of healing and the tensile strength of the healing wound after laser- and scalpel-produced incisions in experimental animals have been performed. In 1971, Hall demonstrated that the tensile strength in a carbon dioxide laser–produced incision was less up to the twentieth day of healing and became the same by the fortieth day.[46] In 1981, Norris and Mullarry studied the healing properties of carbon dioxide laser incisions in pigs histologically.[47] They showed that scalpel-produced incisions exhibited better tissue reconstruction than laser-produced incisions up to the thirtieth day, after which time both incisions exhibited similar results.

In 1982, Finsterbush et al measured the tensile strength of carbon dioxide laser incisions in rabbits and compared them to scalpel wounds.[48] They gently removed the charring and debris on the wound edges before closure. The laser beam–produced wounds were significantly stronger than those made with a scalpel for the first 19 days. In 1983, Buell and Schuller created carbon dioxide laser incisions in pigs and compared them with scalpel incisions.[49] They found the laser wound to be weaker in tensile strength than the scalpel wounds for the first 3 weeks. Research on wound healing from lasers has confirmed the delay in the healing of laser incisions.[50,51] Analysis of the healing process and the effect of the laser on components of the healing process is beginning to provide details about the laser wound-healing delay.[52,53]

Studies in our laboratory have shown that the delay in wound healing is secondary to thermal damage to tissues surrounding the incision and thus produces debris that must be removed before wounds can be repaired.[54–60] This delay has been observed through the first 30 days of wound healing. During the ablation process, some of the thermal energy is conducted to surrounding tissues and can damage the cellular and extracellular structures. Additionally, some of the tissue at the perimeter of the laser beam is exposed to low levels of laser energy, and this subablative energy also damages surrounding tissue. This excess thermal damage can be observed at the borders of laser incisions. The thickness of this thermal damage has been correlated with the delay in wound healing.[51,54]

The laser pulse structure and duration have been shown to be an important parameter in minimizing thermal damage. These studies identified the importance of the pulse length in relation to the thermal relaxation time, a time constant that is the amount of time required for thermal energy to diffuse to surrounding tissues, thus heating and injuring them. This time has been estimated by several models as ranging from 7 to 700 μs.[56,57,60] Thus, a shorter pulse duration should show less thermal damage because the laser energy is delivered so quickly that it cannot diffuse to surrounding tissues. Studies in canine oral mucosa have shown decreases in wound healing delays when using a 60 to 100 μs pulse duration.[57]

By developing lasers that have more favorable wound-healing properties, the major disadvantage of laser incisions could be minimized while still offering the unique advantages that the laser affords for surgery.

LASER TYPES AND APPLICATIONS

Six types of lasers are currently in use in otolaryngology-head and neck surgery, and many more are in various states of development. The characteristic potential for the clinical application of a particular surgical laser is determined by its light emission paradigm, that is, pulse character and wavelength, and by its absorptive characteristics in tissue. The surgeon should therefore consider the properties of each wavelength at the time that he or she chooses to use a particular laser. This facilitates the accomplishment of his or her surgical goal with minimal morbidity and maximal efficiency.

Argon Lasers

Argon lasers produce a visible blue-green light with discrete wavelengths of 488 and 514 nm. The radiant energy from an argon laser is readily transmitted through clear aqueous tissues such as those of the

cornea, lens, and vitreous humor because it has a low water absorption coefficient. Certain tissue pigments, such as melanin and hemoglobin, effectively absorb the argon laser light. When low levels of this blue-green laser light interact with highly pigmented tissues, a localized coagulation within these tissues takes place. The clinician uses this selective absorption of the argon laser by pigmented tissue to photocoagulate pigmented lesions such as port-wine hemangiomas and telangiectasias.[61–65] Used in this application, the radiant energy from the argon laser passes through the overlying skin with some absorption by melanin and beta-carotene and reaches the pigmented target tissue, causing protein coagulation. A second-degree burn occurs in the overlying epidermis. Gradual blanching of the laser-photocoagulated area occurs over several months.

Focusing the argon beam to a small spot size causes its power density to increase sufficiently to result in vaporization of the target tissue. Otologists have used this laser to perform stapedotomies because of its ability to be focused into a very small beam. This procedure was popularized by Perkins.[66] Because bone reflects most visible light, including the blue-green radiant energy of argon laser, it becomes necessary to place a small drop of blood on the stapes to initiate absorption. Other middle ear applications of the argon laser include lysis of middle ear adhesions, ossicle sculpturing, and spot welding of grafts in tympanoplasty.[67–73] Advances in fiber-optic technology have allowed development of a series of handheld argon laser microprobes, which have been called Endo-Otoprobes (HGM Medical Laser, Salt Lake City, Utah).[74–76] The optical fiber-transmitting laser energy is contained within the Endo-Otoprobe, which is handheld like a surgical instrument in the operative field. Because Endo-Otoprobes are handheld and available in a variety of angles, use of these probes allows the surgeon to use the laser in a direction different from the visual field of the microscope, facilitating aiming the laser beam around a corner without having to use an expensive and delicate micromirror.

The advantages of using the argon laser with the Endo-Otoprobes in otology include permitting the surgeon to vaporize or coagulate tissue while avoiding excessive mobilization of the ossicular chain. The blue-green color of the laser light also decreases the risk of damage to the facial nerve and inner ear structures.[77]

Because the blue-green light from the argon laser readily penetrates the eye to the retina, special amber-colored safety glasses are required for all personnel in the operating room.

Neodymium:Yttrium-Aluminum-Garnet Lasers

Neodymium:yttrium-aluminum-garnet lasers produce light that has a wavelength of 1064 nm and is therefore invisible. Because the water absorption coefficient of this laser is low, its radiant energy can be transmitted through clear liquids, which facilitates its use in the eye or other water-filled cavities, such as the urinary bladder or synovial capsule. The absorption of its energy by tissue depends on pigment; therefore, most of the energy is scattered both forward and backward through the tissue. The zone of damage produced by impact with the Nd:YAG laser is not limited as with the carbon dioxide laser. A homogeneous zone of thermal coagulation and necrosis may extend 4 mm from the impact site, and precise control is not possible.

These characteristics make the Nd:YAG laser an excellent surgical tool for tissue coagulation. Vaporization and incision can be performed with this wavelength, but precision is lacking, and tissue damage is difficult to control. Attachment of synthetic sapphire contact tips to the end of the quartz fiber-optic delivery system allows the Nd:YAG laser to be used for incision and vaporization with increased precision and diminished tissue damage.[6,78,79]

The Nd:YAG laser beam can be transmitted through flexible fiber-optic endoscopes. A separate fiber passed down the biopsy channel of the endoscope transmits the laser beam. Extruded quartz or sapphire tips in many shapes are also available.[78,80,81] These are sometimes called sculpted tips, and they are used once and discarded. They appear to be reasonably effective in delivering the laser light to a small volume and providing some limitation to the lateral thermal damage. The major applications for Nd:YAG lasers to date are treatment of acute gastrointestinal hemorrhage[82] and palliation of obstructing bronchial carcinomas[83–85] and esophageal carcinomas.[86–89] Other applications include photocoagulation of hemangiomas[90,91] of the head and neck and intranasal hereditary hemorrhagic telangiectasia.[65,92]

Laser-induced interstitial thermotherapy (LITT) is a technique whereby laser energy is directly applied into tumors with flexible Nd:YAG laser optical fibers passed through needles placed into the tumor. This technique is attractive since it is minimally invasive and carries a low morbidity. It may also allow the treatment of deep and difficult to reach tumors in the head and neck and other areas as improved noninvasive monitoring techniques of laser–tissue interactions are developed. The treatment is enhanced with ultrasonography and magnetic resonance imaging for real-time interstitial needle placement in tumors, identification of vessels, and monitoring and quantifying laser-induced tissue damages. Additionally, laser-induced tissue photoablation can be monitored using real-time color flow Doppler ultrasonography.[93]

Lymphangiomas of the tongue and neck are vascular lesions that are difficult to treat and frequently recur. Magnetic resonance–controlled LITT with the Nd:YAG laser has been shown to be a safe treatment. Real-time monitoring of tissue temperature with thermosensitive sequences allows controlled coagulation necrosis.[94]

It has also been shown in a group of 121 patients with symptoms of nasal obstruction owing to hypertrophied inferior turbinates that an endonasal laser technique, the wedge turbinectomy using the Nd:YAG laser, is effective and safe. A steady improvement in nasal patency was observed in 104 patients (85.9%). Additionally, 65% of the patients who experienced long-term failure in this study were affected by allergic rhinitis.[95]

Neodymium:yttrium-aluminum-garnet lasers are typically available in 15- to 110-watt units. Because the invisible light from the Nd:YAG laser can easily cause retinal damage, specially colored safety glasses are required for all personnel in the operating room.

Pulsed Yellow Dye Laser

The pulsed yellow dye laser operates at a visible wavelength of 585 nm. The laser beam is pulsed to deliver energy for only 360 μs. This laser was specifically developed for use in treating port-wine stains.[11–13,49,96] Oxyhemoglobin has an absorption peak near 577 nm. Although this absorption peak is less than at 418 nm, it does permit deeper laser penetration into tissue. Tan et al showed that at 585 nm there is maximal hemoglobin absorption with a minimum of scattering and minimal absorption by melanin and other pigments.[97] When the duration of the laser pulse is shorter than the time for the heat energy to be conducted away to adjacent tissues, the thermal relaxation time, excessive tissue heating can result. It was found for cutaneous microvessels that 20 μs pulses from a pulsed dye laser cause vessel wall fragmentation and hemorrhage. Garden et al found that longer pulse widths caused no significant hemorrhage.[11] The pulse width of 360 μs caused thermal denaturation of the vessel, producing a more desirable clinical outcome. The epidermis in their study showed no evidence of damage when the laser was used at recommended threshold exposures.

The flash pump dye laser has also been used for the treatment of laryngeal papillomas. In a trial comparing flash pump dye laser to the carbon dioxide laser treatment of laryngeal papillomas in a children's hospital, none of the patients described major discomfort, and five patients described their voice as being the same as or better than it was after prior carbon dioxide laser procedures. A treatment effect was noted in all patients and was similar to the results noted on the carbon dioxide laser–treated side. Because the flash pump dye laser coagulates rather than vaporizes tissue, potential advantages may include decreased scarring relative to carbon dioxide laser treatment and improved patient and operator safety.[98]

An innovative use for the flashlamp pumped dye laser tuned to slightly shorter wavelengths was suggested by Ito and Baba.[99] An in vivo study was performed to evaluate the use of pulsed dye laser beam for the fragmentation of salivary calculi. Optimal fragmentation was achieved with a wavelength of 504 nm. Under continuous endoscopic monitoring, laser-induced shock wave lithotripsy is performed. Laser lithotripsy of salivary stones with endoscopic monitoring permits treatment on an outpatient basis with little inconvenience to the patient.

The yellow dye laser has special electrical and plumbing requirements. Special safety eye wear is required.

Argon Tunable Dye Laser

The argon tunable dye laser system works on the principle of the argon laser, making a high-intensity

beam that is focused on dye that is continuously circulating in a second laser optically coupled to the argon laser. The argon laser beam energizes the dye, causing it to emit laser energy at a longer wavelength than the pump beam. By varying the type of dye and using a tuning system, different wavelengths can be obtained. The laser energy from this dye laser can then be transmitted through flexible fiber-optics and delivered through endoscopic systems or inserted directly into tumors. The major clinical use of this laser is with selective PDT of malignant tumors following the intravenous injection of a photosensitizer, typically, hematoporphyrin derivative.[18–25] The argon tunable dye laser is normally tuned to emit laser light at 630 nm.

After the intravenous injection, the photosensitizer disseminates to all of the cells of the body, rapidly moving out of normal tissue but remaining longer in neoplastic tissue. After a few days, a differential in concentration exists between the tumor cells and the normal cells. When the tumor is exposed to red light (630 nm), the photosensitizer absorbs the light; the absorption of this red light causes a photochemical reaction to occur. Toxic oxygen radicals, such as singlet oxygen, are produced within the exposed cells, causing selective tissue destruction and cellular death. Since there is less photosensitizer in the normal tissues, a much less severe or no reaction occurs. The main technical problem is getting enough light to the target area. Here the argon tunable dye laser system has helped to solve this problem. Additional research to increase the laser intensity and simplify the sometimes cumbersome setup is being conducted with gold vapor lasers.[21,25]

Results obtained by many investigators in this country demonstrate that the premise of treating selected neoplasms with a photosensitizer followed by activation with light is valid. The overall potential and exact place of maximum value of this form of treatment remain to be established. Areas that appear to be very promising include carcinoma of the urinary bladder,[27] endobronchial lesions of the lung,[28] selected carcinomas of the upper aerodigestive tract,[29] skin cancers,[30] and cutaneous metastases of breast cancers.[31] Trials are now being conducted in certain specialties on intraoperative PDT in conjunction with conventional surgery.[28]

Manyak et al reported in 1988 that patients who underwent PDT for recurrence of a primary head and neck neoplasm generally did not have improved survival despite successful local tumor eradication.[19] Photodynamic therapy may be indicated for either curative or palliative treatment of endobronchial malignancies. Prospective studies are under way to examine the efficacy of PDT for endobronchial neoplasms.

Photodynamic therapy has also been used as a palliative treatment for recurrence (after chemotherapy, external beam irradiation, and surgery) or in elderly patients to avoid inappropriately extensive and mutilating surgery.[33] The photosensitizing dye was Photosan 3. The light source was an argon dye laser with a wavelength of 630 nm and a power of 100 mW/cm^2, coupled with a cylindrical or spherical applicator and a lens fiber tip. It has been used for patients suffering from cancer of the palatine arch, tonsil, nasal septum, and glottis.

In many instances, PDT is used as an adjuvant therapy with surgery to reduce tumor recurrence in the head and neck. It is given intraoperatively after resection. A concern with the use of intraoperative PDT is the possible effect on wound healing, especially on the healing of myocutaneous flaps, which are widely used to reconstruct defects following resections for head and neck cancer. In a laboratory study, rats receiving porfimer sodium (Photofrin) and laser irradiation had serous effusion, epidermal necrosis, and weaker tensile strength of the healing of a flap based on the inferior epigastric artery.[34] These results should be considered when contemplating the use of PDT as adjuvant intraoperative therapy for tumor surgery requiring flap reconstruction after ablative surgery.

The argon dye laser requires special electrical and plumbing installation. Special eyewear is required. In addition, prolonged photosensitivity is the most common side effect encountered by the patients.

KTP/532 Laser

The KTP/532 laser is an optically pumped, solid-state laser that produces a visible 532 nm beam with an emerald green color. This laser is a frequency-doubled Nd:YAG laser producing all of its output at 532 nm. The Nd:YAG laser rod is continuously pumped with a krypton arc lamp and Q-switched. The Nd:YAG laser wavelength is frequency-doubled by placing a KTP crystal in the beam path, which halves the wavelength, producing the characteristic

532 nm output. Like the argon laser, the radiant energy from the KTP laser is readily transmitted through clear aqueous tissues because it has a low water absorption coefficient. Certain tissue pigments, such as melanin and hemoglobin, absorb the KTP laser light effectively. When low levels of green laser light interact with highly pigmented tissues, a localized coagulation takes place within these tissues. The KTP/532 laser can be selected for procedures requiring precise surgical excision with minimal damage to surrounding tissue, vaporization, or photocoagulation. The power density chosen for a given application determines the tissue interaction achieved at the operative site.

The KTP/532 laser is transmitted through a flexible fiber-optic delivery system that can be used in association with a micromanipulator attached to an operating microscope or freehand in association with various handheld delivery probes having several different tip angles. These handheld probes facilitate use of the KTP/532 laser for benign and malignant oral cavity applications,[10,100] functional endoscopic sinus surgery and other intranasal applications,[101,102] and otologic applications.[73,103–106] Examples of handheld KTP/532 laser applications include tonsillectomy, turbinectomy, stapedotomy, excision of acoustic neuroma, and excision of benign and malignant laryngeal lesions. Use of the micromanipulator facilitates middle ear and microlaryngeal laser surgery because it is more conventionally performed using a carbon dioxide laser.

Different dyes are being tested for PDT that absorb strongly in the visible region of the spectrum. The KTP/532 laser becomes the obvious light source. It was reported that cultured human P3 squamous cells incubated 2 hours with daunomycin exhibited enhanced cytotoxicity after exposure to the monochromatic 532 nm green light from a KTP laser.[107] When the KTP laser output was varied from 0 to 120 joules in daunomycin-sensitized tumor cells, a linear phototherapy response was seen with energy as low as 12 J inducing drug phototoxicity.

The KTP/532 laser is typically available as a unit that delivers between 0.5 and 20 watts with fiber diameters of 200, 400, and 600 microns. Because the visible green light from the KTP/532 laser can penetrate the eye and cause retinal damage, special wavelength-specific safety glasses must be worn by the surgeon and all personnel in the operating room.

Other Lasers

In an effort to have a more controlled laser effect with less damage to adjacent tissue, several lasers in the near- to mid-infrared region are being investigated. These include the erbium:YAG and the holmium:YAG. The erbium:YAG lases at the infrared peak of water absorption at 2.94 μm. Here the extinction length in water is less than 2 μm. The laser produces very clean incisions with a minimal amount of thermal damage to the adjacent tissue.[108] The erbium:YAG laser has been quite successful in ablating bone and cartilage in laboratory investigations.[109,110] There are some concerns about the negative consequences of the photoacoustic signal that accompanies the ablation.[111-116] There are two negative aspects to this laser. The wavelength is too long to be transmitted through normal optical fibers. This gives a distinct advantage to lasers that produce light that can be transmitted through fibers. More importantly, the thermal propagation is so short that there is practically no tissue coagulation and no hemostasis. This laser is therefore unsuitable to use in highly vascular tissue.

The holmium:YAG laser operates at 2.1 μm. This wavelength can be effectively transmitted through fibers. The extinction length in water is about 0.4 mm, which suggests that this laser light should interact with tissue very similar to the carbon dioxide laser. The holmium:YAG has been combined with the fiber-optic endoscope for sinus surgery.[117,118] The hemostasis is good, and the soft bone ablation is readily controlled.[119–121] Adjacent thermal damage zones vary from 130 to 220 μm.

A study demonstrated the clinical efficacy of the holmium:YAG laser for endoscopic sinus surgery.[122] In a prospective, randomized, controlled, single-blinded study, 32 patients underwent endoscopic sinus surgery using the holmium:YAG laser on one side of the nose and conventional endoscopic instrumentation on the other side. Patients were followed for 2 years after surgery. The mean intraoperative blood loss was 24.6 mL less on the laser-treated side of the nose than on the conventionally treated side ($p < .001$). Improvements in symptoms of pain, congestion, and drainage were equivalent for both treatment modalities. Microscopic analysis demonstrated the ability of the holmium:YAG laser to remove tissue in relatively thin layers with ablation depths of 260 to 340 microns per pulse.

A study also demonstrated the utility of the holmium:YAG laser in the management of pediatric airway disorders. In the study, safety, precision, hemostasis, bone-cutting properties, and accessibility to the lesion of the equipment were compared to standard therapies. Additionally, postoperative outcomes were compared to standard therapies. It was concluded that the holmium:YAG laser is a safe, effective tool in the treatment of pediatric airway disorders and offers the advantage of a flexible fiber-optic system, good hemostasis, and better bone-cutting characteristics than the carbon dioxide laser.[123]

There is also work with other materials that emit light in the near-infrared region of the spectrum, such as the cobalt:magnesium fluoride laser (tunable from 1.8 to 2.14 μm).[124] Alexandrite lasers (750 nm)[125] and titanium sapphire lasers (tunable from 0.6 to 1.0 μm) have also been considered.[126]

The diode laser (600 to 1,000 nm) has also been used for various surgical applications. The diode laser is small and lightweight, and the light is delivered by an optical fiber. The interaction of the diode laser light is similar to the Nd:YAG or the KTP/532 laser. There is moderate absorption by melanin and hemoglobin. It has been considered for turbinate reduction,[127] laser-assisted stapedectomy,[128] and mucosal intact laser tonsillar ablation[129] and is to be used with laser welding techniques to prevent fistula formation.[130]

Ultimately, many parameters such as cost, reliability, and size, in addition to the tissue response, will influence choice of lasers in medicine.

Carbon Dioxide Laser

Carbon dioxide lasers produce light with a wavelength of 10.6 μm in the mid-infrared, and therefore invisible, range of the spectrum. The site where the invisible carbon dioxide laser beam will impact the target tissue is indicated by a built-in coaxial helium-neon laser beam that is red. To increase the visibility of the aiming beam, some lasers have a green aiming beam. The carbon dioxide laser has a high water absorption coefficient, is independent of tissue color, and is well absorbed by all soft tissues that are high in water content; the thermal effects of adjacent nontarget tissue are minimal. The extinction length of this wavelength is about 0.03 mm in water and in soft tissue; reflection and scattering are negligible. These properties make the carbon dioxide laser versatile for use in otolaryngology-head and neck surgery.

At present, the invisible radiant energy emitted from the optical resonating chamber of a carbon dioxide laser can be mechanically conveyed by means of a series of mirrors through an articulating arm to the target tissue.[131] Transmission through rigid nonfiber-optic endoscopes facilitates use of the carbon dioxide laser for bronchoscopy, laparoscopy, and arthroscopy. Hollow and flexible waveguides are available for clinical use, but there is considerable loss of energy and limitations on maximum incident energy. However, these flexible waveguides are shown to be effective in a number of procedures.

Carbon dioxide surgical lasers can operate in both continuous and pulsed modes. They can be used freehand for macroscopic surgery, adapted to the operating microscope for microlaryngoscopy, or adapted to a bronchoscopic coupler for rigid bronchoscopy.[132–135]

Carbon dioxide lasers are typically available in 25- to 100-watt units. They are self-contained and have no special electrical or plumbing requirements.

Although the beam cannot penetrate the eye to the retina, it can cause corneal or scleral burns.[136,137] Therefore, protective glasses are required; ordinary glass or plastic correctional lenses have been satisfactory. Clear plastic safety glasses have also provided adequate protection. Hard contact lenses stop the beam, but all of the uncovered portions of the eye are unprotected from any misdirected laser radiation.

Macroscopic Applications Macroscopic applications of the carbon dioxide laser are numerous. In the oral cavity, benign tumors can be excised or ablated with the laser.[138-143] A one-stage tongue release can be effectively performed; this procedure is helpful in rehabilitating patients following composite resection with tongue flap reconstruction.[144] Speech and deglutition can be improved in selected cases. Multiple areas of leukoplakia can be precisely excised and ablated[29,138,145–148]; in most instances, a graft is not necessary to resurface the operative field. Selected superficial carcinomas can be excised with the use of the laser, and large recurrent or inoperable tumors can be debulked for palliation.[88,143,149,150]

Within the nasal cavities and paranasal sinuses, the carbon dioxide laser is used to treat choanal atresia, hypertrophic inferior turbinates, squamous

papilloma, recurrent inverted papilloma, and hereditary hemorrhagic telangiectasia.[151–156] For this last condition, we believe the Nd:YAG laser to be a more efficacious instrument. The optical waveguides have improved the treatment for hypertrophic inferior turbinates[157] and have permitted a laser partial epiglottidectomy to be performed in those patients who have not been improved with uvulopalatopharyngoplasty.[158] Also, the Swiftlase scanner for the continuous-wave carbon dioxide laser has been used in the management of 387 patients with hypertrophy of the inferior turbinate mucosa. The 1-year postoperative follow-up revealed good results in 261 (81%) of 321 patients.[159] In at least one study, patients preferred the carbon dioxide laser to the Nd:YAG laser for treatment of the inferior turbinates.[160]

Laser surgery of the inferior turbinates has been used as an effective surgical treatment for patients with allergic rhinitis, particularly for those who have persistent nasal obstruction and do not respond well to pharmacologic therapy. There is evidence that the expression of local inflammatory cytokines can be attenuated by carbon dioxide laser treatment.[161] This may be closely related to the clinical effectiveness of this procedure. Treatment with the carbon dioxide laser has been shown to be more effective with a laser output of 5 W than with an output of 3 W. Additionally, treatment was judged less effective in patients aged 15 years or less than in older patients.[162]

Lingual tonsillectomy, excision of tonsillar neoplasms, serial excision of the palatine tonsils, and excisions of neoplasms of the posterior pharyngeal wall can be performed with the help of this laser.[139,163,164] Peritubal adenoid tissue and recurrent papilloma in the nasopharynx can be ablated by reflecting the radiant energy of the carbon dioxide laser off the front surface of a stainless steel mirror.

Facial plastic surgical applications include excision of rhinophyma,[165–168] excision of benign and malignant skin tumors,[169] and ablation or dermabrasion of nevi and tattoos.[170–173]

Procedures have been introduced to the United States from Europe that allow a serial ablation of the uvula and soft palate to treat patients with snoring problems successfully.[174] Performed in an ambulatory setting under local anesthesia, this procedure is associated with much less pain and perioperative morbidity for the patient than with conventional surgery. Additionally, there is no risk of airway obstruction, no hospitalization, and no lost time from work. For these reasons, it has become very popular. The treatment was shown to be nearly five times more effective in patients with a body mass index (see Chapter 58) less than 28 kg/m^2.[175]

A handheld otoscope has been combined with a flashscanner of the carbon dioxide laser for pressure equalization tube insertion in an office setting.[176] Laser-assisted tympanic membrane fenestration was performed with the laser set at single pulse. Insertion of grommets was accomplished using the otomicroscope and an "alligator" microforceps. Although the data indicate high parent and physician satisfaction, along with strong incentives for physicians to use this technique, there are few data that compare this technique to general anesthesia and surgical insertion of tubes.

Microscopic Applications *Benign Laryngeal Disease* The carbon dioxide laser has been used more in microlaryngeal surgery than in any other area of interest to the otolaryngologist-head and neck surgeon. This laser has universal applications in microscopic surgery of benign laryngeal disease.[134] Its advantages of increased precision and decreased postoperative edema must be weighed against the disadvantages of risks of laser-associated complications and increased operating time. With experience, the drawbacks become limited, and the advantages open new horizons. Although knowledge of its use permits the otolaryngologist to apply the carbon dioxide laser in practically all microlaryngeal procedures for benign laryngeal disease in a safe and effective manner, it is not the only surgical tool to be used for any given operative procedure.

Surgery for recurrent respiratory papillomatosis has advanced with the use of the laser.[153,154,177] Initial disappointment in its inability to cure the disease has been tempered by its effectiveness in preserving normal laryngeal structures and maintaining the translaryngeal airway with an efficiency previously unattained.[178,179] Surgery in children for webs, subglottic stenosis, capillary hemangiomas, and other lesions is also significantly enhanced by the precision, preservation of normal tissue, and predictable minimal postoperative edema associated with judicious use of the carbon dioxide laser.[180–184]

Although laryngeal hemangiomas in adults are uncommon, there have been four cases reported in

2000.[185] All masses were detected at endoscopy, and all were biopsied and ablated uneventfully with a carbon dioxide laser. Laryngeal hemangiomas that affect only the glottic area can be treated endoscopically with a carbon-dioxide laser with little to no risk.

Surgery in adults for polyps, nodules, leukoplakia, papillomas, cysts, and other benign laryngeal conditions also finds an advantage with this laser.[186–188] In fact, the laser has created an era of conservation surgery for benign disease. Classically, laryngeal surgery has consisted of vocal cord stripping with healing by remucosalization. Now more normal tissue can be preserved and more limited spot ablations performed; even flaps of mucosa may be mobilized and advanced in the larynx with endoscopic laser techniques.[134] The laser has lowered the risks to the integrity of the anterior commissure of the larynx during one-stage operations on both cords by its ability to ablate areas and to leave adjacent tissue 0.1 mm away undisturbed by mechanical or other effects.

Biopsy may be performed with the laser, even with a small spot size machine, although there is some tissue vaporization. Scissors and cup forceps still have a role for small lesion biopsy. Surgical judgment is obviously important. In extensive areas of suspicious disease, the laser (with a small spot size) is appropriately used for excisional biopsy.

Webs and subglottic stenoses are also effectively incised with the laser.[189–192] Fewer than 50% of anterior commissure webs recur after laser resection. Yet some lasers have such large spot sizes that their use is inappropriate for the procedure. Stenosis can also be treated as long as treatment requires simple tissue ablation with minimal associated edema. If skeletal support is needed, the laser does not provide this, and other techniques should be used.

Endoscopic arytenoidectomy has been performed for many years.[193–196] The addition of the laser allows the surgeon to vaporize the mucosa and underlying arytenoid cartilage layer by layer precisely in a dry field. The precision associated with its use facilitates performance of this operation even by surgeons who had difficulty in mastering the conventional techniques of endoscopic arytenoidectomy. Other microscopic applications of the carbon dioxide laser include excision of internal laryngoceles and excision of hypopharyngeal diverticula.[197–201] In the latter application, a specially designed endoscope facilitates performance of the operation.

The introduction of the adult subglottiscope has extended the microscopic control and delivery of the carbon dioxide laser radiation to the subglottis and upper trachea.[202] Selected stenoses in this area have been successfully treated using microtrapdoor flaps.[203,204]

In summary, the carbon dioxide laser has proved to be a tool with numerous applications for microscopic laryngeal surgery. As the instrument becomes more familiar and friendly to the operating team, the results will continue to improve.

Malignant Laryngeal Disease. Endoscopic management of malignant laryngeal disease is not a new concept. In 1920, Lynch reported 39 patients with early glottic cancers that he successfully treated with transoral excision.[205] In 1941, New and Dorton reported a 90% cure rate with transoral excision and diathermy.[206] In 1973, Lillie and DeSanto reported 98 patients with early glottic carcinoma who were treated with transoral excision; all were cured, although 5 of these patients required further treatment.[207] Hemostasis was achieved with electromicrocautery.

In carcinoma in situ, the tumor's malignant cells have not penetrated the basement membrane and therefore have no metastatic potential. In a study reported in 2000, 29 patients with glottic carcinoma in situ with a minimum follow-up of 2 years were treated initially with transoral carbon dioxide laser surgery.[208] A complete removal of the tumor was possible with superficial laser cordectomy in 21 patients and with subligamental laser cordectomy in 8 cases. There was no tumor-related death in this series. Repeated laser resections were performed in 4 patients for local recurrences. No patient required total laryngectomy or radiotherapy during a follow-up ranging from 25 to 143 months. The local control rates were superior to those previously reported with conventional surgery and similar to those after radiotherapy. The ultimate rate of larynx preservation was seven times higher than reported after radiotherapy. Carbon dioxide laser surgery is an option for the treatment of carcinoma in situ of the larynx.

The transoral treatment of squamous cell carcinoma of the larynx using the carbon dioxide laser is therefore an obvious extension of the application of this surgical instrument.[209–211] The advantages of precision, hemostasis, and decreased postoperative edema allow the laryngologist to perform exquisitely

accurate and relatively bloodless endoscopic surgery of the larynx. The laser has facilitated a renaissance in the endoscopic management of early vocal cord carcinomas.[212,213]

Three distinct advantages exist with this modality of treatment. First, with the laser, it is possible to remove bulky tumors of the anterior commissure area or vocal fold to obtain an accurate assessment of its extent for purposes of staging and treatment. Additionally, previously biopsied tumors of the vocal cord may represent a diagnostic problem with respect to their extent. Here the laser may be used to excise the area surrounding the biopsy site.

Inexact staging after direct laryngoscopy is fairly common. By definition, a T1 carcinoma of the true vocal cords requires cord mobility as defined by the indirect examination. This implies that the tumor is confined to the surface epithelium and has not deeply invaded the vocalis muscle. Vocal cord mobility by itself is a poor indicator of tumor invasion of the underlying muscle.[214] Deep invasion of the vocalis muscle by the cancer has been found after excisional biopsy in many patients who were staged T1 by the indirect examination. These patients should realistically be reclassified as T3 squamous cell carcinoma.

Airway re-establishment is a second advantage of the carbon dioxide laser in treating patients with laryngeal carcinoma. The laser can be used to reduce the amount of tumor obstructing the airway, thereby avoiding the need for a preoperative tracheostomy.

The third advantage of using the carbon dioxide laser in the treatment of laryngeal cancers is that endoscopic excision of laryngeal carcinoma using the carbon dioxide laser may be curative. The experience of Strong et al,[143] Ossoff et al,[214] Blakeslee et al,[215] and Hirano et al[216] appears to support the previous work of Lillie and DeSanto,[207] New and Dorton,[206] and Lynch.[205]

Our current treatment plan for the endoscopic management of T1 vocal cord carcinoma includes an excisional biopsy with the carbon dioxide laser.[214] Supravital staining with toluidine blue can be performed as a diagnostic aid before biopsy. The surgical specimen is labeled and sent for frozen-section examination. Any questionable margins are controlled by frozen section. If the tumor is found to be histologically T3 or T4, the patient is later treated by conventional surgical techniques (partial or total laryngectomy).

Advantages of Carbon Dioxide Laser Surgery. The greatest advantage of the carbon dioxide laser is precision, which allows the surgeon to make accurate lesions. When the carbon dioxide laser is coupled to the operating microscope with a micromanipulator, it becomes possible to achieve surgical accuracy within 0.1 mm. The smallest spot sizes are in the range of 0.10 (200 mm objective) to 0.3 mm (400 mm objective) with current technology. This control, provided by the micromanipulator, is significantly more than is available with handheld microlaryngeal cupped forceps. The second major advantage of the carbon dioxide laser is hemostasis. Characteristically, the laser coagulates vessels up to 0.5 mm in diameter but has the ability to coagulate larger vessels up to 2 mm in diameter under certain circumstances. Vessels larger than 2 mm must be clamped and ligated. The third advantage of the carbon dioxide laser is decreased postoperative edema. A fourth advantage is decreased postoperative pain. Pain has been reported to be reduced or absent after use of the carbon dioxide laser. This probably results from the thermal sealing of nerve endings in the operative wound. Other advantages that have been attributed to the carbon dioxide laser include decreased postoperative scarring and uncomplicated wound healing.

Cold Steel versus the Carbon Dioxide Laser. We are increasing our preference for cold steel over the carbon dioxide laser for critical surgery that involves the vocal folds and can affect the voice. Both tools are effective. However, we find that with sufficient training and skill, the cold steel can yield a slightly better voice, probably owing to the lack of lateral thermal damage.

A similar conclusion was found by a group in Germany.[217] They reported on a prospective randomized study to examine the postoperative functional results after laser phonosurgery in comparison with cold phonosurgery. Conventional cold phonosurgery was performed in 23 patients, and 21 patients were treated by laser phonosurgery. To determine vocal function, examinations included direct videolaryngoscopy, determination of maximal phonation, speech voice field, and assessment of their singing voice. The results 4 months after surgery showed an improvement of vocal function in both treatment groups in comparison with the

preoperative findings. However, the improvement is only statistically significant after cold surgery.

Otosclerosis. The carbon dioxide laser has been shown to possess ideal tissue interaction characteristics for surgery in the oval window for the performance of primary stapedotomy and revision stapedectomy.[218] A small spot size microscopic coupler should be used in these cases.

Bronchoscopic Applications. Successfully coupling the carbon dioxide laser to a ventilating bronchoscope represented a logical extension of the clinical application of the carbon dioxide laser.[149,219] Specifically, use of the carbon dioxide laser coupled to a ventilating bronchoscope allows the surgeon to perform hands-off endoscopic surgery through the relatively long and narrow operative field of a ventilating bronchoscope. The advantages of precision and hemostasis facilitate visualization, and decreased postoperative edema and scarring facilitate control over the final result of the operation.

Current indications for bronchoscopic carbon dioxide laser surgery include management of recurrent respiratory papillomatosis[153,154,177,178] or granulation tissue involving the trachea[220] and resection of selected areas of tracheal stenosis.[162,191,192,221,222] Tracheal and proximal endobronchial adenomas and webs can also be resected using the carbon dioxide laser.[223] Finally, patients with selected obstructing tracheal and proximal endobronchial cancers can have their airways reestablished as a means of palliation using the laser bronchoscope.[224] The major contraindication to this procedure is extraluminal compression of the trachea by an extrinsic tumor. The Nd:YAG laser is preferred over the carbon dioxide laser for this latter application.[225,226]

Anesthetic Techniques for Carbon Dioxide Laser Surgery. Optimal anesthetic management of patients for laser surgery of the larynx must include attention to the safety of the patient, the requirements of the surgeon, and the hazards of the equipment.[227,228]

Because of the time required to align accurately and manipulate the laser, most patients require general anesthesia for this type of operation. Any nonflammable general anesthetic is suitable, and both halothane and enflurane fall into this category. Because of the fire hazard, the inspired concentration of oxygen is important. Mixtures of helium, nitrogen, or air plus oxygen are commonly used to maintain the FIO_2 below 40% but also to ensure adequate oxygenation. Intravenous supplementation of inhalation anesthesia with small doses of narcotic and/or tranquilizers is commonly used to shorten the emergence period following anesthesia. Total intravenous anesthesia with reversible short-acting narcotics and tranquilizers plus jet ventilation is another acceptable anesthetic technique for microlaryngeal laser surgery. The use of the adult subglottiscope requires the use of jet ventilation and total paralysis.

Precautions for Carbon Dioxide Laser Surgery. Certain precautions are necessary when using the carbon dioxide laser. First, to reduce the risk of ocular damage, the patient's eyes should be protected by a double layer of moistened eye pads, and all operating room personnel should wear protective glasses. Additionally, a sign should be placed outside the operating room door warning all persons entering the room to wear protective glasses because the laser is in use. Corneal damage can occur from direct or reflected laser beam irradiation to the eyes. Second, all exposed skin and mucous membranes of the patient outside the surgical field should be protected by a double layer of moistened surgical towels.

Protection of the endotracheal tube from either direct or reflected laser beam irradiation is of primary importance.[229–233] Should the laser beam strike an unprotected endotracheal tube carrying oxygen, ignition of the tube could cause a catastrophic, intraluminal, blowtorch-type airway fire. Red rubber endotracheal tubes wrapped circumferentially with reflective tape have been used successfully in the past to reduce the risk of intraluminal fire. Metallic endotracheal tubes and silicone endotracheal tubes specially designed for use in microlaryngeal laser surgery also reduce the risk of intraluminal fire.[234–240] The cuff of the endotracheal tube should be filled with methylene blue–colored saline to act as a heat sink. The placement of water-saturated neurosurgical cottonoids below the vocal cords to protect the cuff of the endotracheal tube further reduces this risk. These cottonoids need to be moistened frequently during the procedure. The

development of the operating platform placed in the subglottic larynx can minimize the possibility of an airway fire even further. Should the cuff become deflated by an errant hit of the laser beam, the tube should be removed and replaced with a new one. The cottonoids must be counted like surgical sponges at the completion of the procedure. Perhaps the single most important safety precaution related to reducing the risk of an intraluminal endotracheal tube fire is using mixtures of oxygen (FIO_2 40% or less) and helium for ventilating the patient during microlaryngeal laser surgery.

Additional precautions must be taken to provide for adequate smoke evacuation from the laryngoscope, bronchoscope, or oral cavity to allow for safety as well as visualization of the operative field.[241–244]

Complications of Carbon Dioxide Laser Surgery. Many of the possible complications of carbon dioxide laser surgery have been mentioned in the general discussion of precautions for laser surgery. These include corneal burns, skin burns, misdirected burns from direct or reflected energy, and endotracheal tube ignition.[245–247]

Other complications associated with the use of the carbon dioxide laser include trauma to the vocal cords during intubation with the aluminum foil–wrapped endotracheal tube, intraoperative airway obstruction secondary to a foreign body (aluminum foil tape or cottonoids), tracheal perforation, acquired glottic web, subglottic stenosis, vocal cord fibrosis, arytenoid perichondritis, delayed bleeding, and delayed postoperative airway obstruction.[248–253]

Great care must be exercised when wrapping the red rubber endotracheal tube circumferentially with the reflective metal tape. We recommend using quarter-inch tape because it wraps better than larger sizes with fewer subsequent folds, kinks, or rough spots. Additionally, the endotracheal tube should be inspected for areas of possible dislodged aluminum tape.

Tracheal perforation is a distinct possibility when performing laser bronchoscopy. Using the laser in the single- or repeat-pulse mode should allow the surgeon to differentiate among mucosa, perichondrium, and tracheal cartilage. Using the laser in the continuous mode often makes that necessary distinction rather difficult.

Acquired glottic web can occur if laser surgery is performed bilaterally at the anterior commissure.[254,255] Subglottic stenosis can result from accidental laser irradiation to the subglottic larynx while working at the anterior or posterior commissure.[256] We recommend using the operating platform to protect the mucosal tissues of the subglottis. Vocal cord fibrosis results from vaporizing vocalis muscle during the process of laser surgery to the true vocal cords.[254,257] This complication can be avoided by proper control of the power density while working on the true vocal cords. Close observation also aids in the prevention of this complication; a small amount of tissue retraction occurs within the laser impact point when the vocalis muscle is penetrated by the laser beam. Arytenoid perichondritis can occur after deliberate or accidental cartilage exposure during endoscopic laser surgery. Parenteral antibiotics help to minimize this complication.

Delayed postoperative bleeding often occurs if vessels larger than 1 mm are treated by laser coagulation. In dealing with bleeding vessels, those that would ordinarily need to be clamped and ligated if the laser were not being used should still be clamped and ligated. Only vessels less than 1 mm should be coagulated with the carbon dioxide laser. Although delayed postoperative airway obstruction associated with using the carbon dioxide laser is rare, it can occur within 18 hours of the operative procedure. The most frequent time for airway edema to occur in our experience has been between 1 and 6 hours postoperatively. We do not recommend performing outpatient endoscopic microlaryngeal laser procedures. When performing laser surgery in the oral cavity under local anesthesia, we keep our patients in the day surgery suite for up to 6 hours postoperatively to observe them for edema or possible airway obstruction before discharge.

REFERENCES

1. Andrews AH, Polanyi TG, editors. Microscopic and endoscopic surgery with the CO_2 laser. Littleton (MA): John Wright-PSG; 1982.
2. Ossoff RH, Werkhaven JA, Raif J, Abraham M. Advanced microspot microslad for the CO_2 laser. Otolaryngol Head Neck Surg 1991;105:411–4.
3. Shapshay SM, Wallace RA, Kveton JF, et al. New microspot micromanipulator for CO_2 laser applica-

tion in otolaryngology—head and neck surgery. Otolaryngol Head Neck Surg 1988;98:179–81.
4. Ossoff RH. Bronchoscopic laser surgery: which laser when and why. Otolaryngol Head Neck Surg 1986;94:378–81.
5. Unger M. Neodymium:YAG laser therapy for malignant and benign endobronchial obstructions. Clin Chest Med 1985;6:277–90.
6. Apfelberg DB, Maser MR, White DN, et al. Combination treatment for massive cavernous hemangioma of the face: YAG laser photocoagulation plus direct steroid injection followed by YAG laser resection with sapphire scalpel tips, aided by superselective embolization. Lasers Surg Med 1990;10:217–23.
7. Joseph M, Reardon E, Goodman M. Lingual tonsillectomy: a treatment for inflammatory lesions of the lingual tonsil. Laryngoscope 1984;94:179–84.
8. Krespi YP, Har-El G, Levine TM, et al. Laser laryngeal tonsillectomy. Laryngoscope 1989;99:131–5.
9. Kuhn F. The KTP/532 laser in tonsillectomy. KTP/532 clinical update. Laserscope No. 06, San Jose, CA, 1988.
10. Oas RE, Bartels JP. KTP-532 laser tonsillectomy: a comparison with standard technique. Laryngoscope 1990;100:385–8.
11. Garden JM, Polla LL, Tan OT. The treatment of port-wine stains by the pulsed dye laser. Analysis of pulse duration and long-term therapy. Arch Dermatol 1988;124:889–96.
12. Tan OT, Kershmann R, Parrish JA. The effect of epidermal pigmentation on selective vascular effects of pulsed laser. Lasers Surg Med 1984;4:365–74.
13. Tan OT, Stafford TJ, Murray S, Kurban AK. Histologic comparison of the pulsed dye laser and copper vapor laser effects on pig skin. Lasers Surg Med 1990;10:551–8.
14. Schweitzer VG, Visscher D. Photodynamic therapy for treatment of AIDS-related oral Kaposi's sarcoma. Otolaryngol Head Neck Surg 1990;102:639–49.
15. Cortese DA, Kinsey JH. Hematoporphyrin-derivative phototherapy for local treatment of cancer of the tracheobronchial tree. Ann Otol Rhinol Laryngol 1982;91:652–5.
16. Dougherty TJ. Photodynamic therapy. Photochem Photobiol 1993;58:895–900.
17. Davis RK. Photodynamic therapy in otolaryngology-head and neck surgery. Otolaryngol Clin North Am 1990;23:107–19.
18. Gluckman JL. Hematoporphyrin photodynamic therapy: is there truly a future in head and neck oncology? Reflections on a 5-year experience. Laryngoscope 1991;101:36–42.
19. Manyak MJ, Russo A, Smith PD, Glatstein E. Photodynamic therapy. J Clin Oncol 1988;6:380–91.
20. Gluckman JL, Zitsch RP. Photodynamic therapy in the management of head and neck cancer. Cancer Treat Res 1990;52:95–113.
21. Petrucco OM, Sathananden M, Petrucco MF, et al. Ablation of endometriotic implants in rabbits by hematoporphyrin derivative photoradiation therapy using the gold vapor laser. Lasers Surg Med 1990;10:344–8.
22. Wenig BL, Kurtzman DM, Grossweiner LI, et al. Photodynamic therapy in the treatment of squamous cell carcinoma of the head and neck. Arch Otolaryngol Head Neck Surg 1990;116:1267–70.
23. Rausch PC, Rolfs F, Winkler MR, et al. Pulsed versus continuous wave excitation mechanisms in photodynamic therapy of differently graded squamous cell carcinomas in tumor-implanted nude mice. Eur Arch Otorhinolaryngol 1993;250:82–7.
24. Okunaka T, Kato H, Conaka C, et al. Photodynamic therapy of esophageal carcinoma. Surg Endosc 1990;4:150–3.
25. Shikowitz MJ. Comparison of pulsed and continuous wave light in photodynamic therapy of papillomas: an experimental study. Laryngoscope 1992; 102:300–10.
26. Donaldson DR, Emami AJ, Wax MK. Endoscopically monitored percutaneous dilational tracheotomy in a residency program. Laryngoscope 2000;110: 1142–6.
27. Dougherty TJ, Marcus SL. Photodynamic therapy. Eur J Cancer 1992;28A:1734–42.
28. Pass HI. Photodynamic therapy in oncology: mechanisms and clinical use. J Natl Cancer Inst 1993; 85:443–56.
29. Grant WE, Hopper C, Speight PM, et al. Photodynamic therapy of malignant and premalignant lesions in patients with 'field cancerization' of the oral cavity. J Laryngol Otol 1993;107:1140–5.
30. Lui H, Anderson RR. Photodynamic therapy in dermatology: recent developments. Dermatol Clin 1993;11:1–13.
31. Khan SA, Dougherty TJ, Mang TS. An evaluation of photodynamic therapy in the management of cutaneous metastases of breast cancer. Eur J Cancer 1993;29A:1686–90.
32. Biel MA. Photodynamic therapy of head and neck cancers. Semin Surg Oncol 1995;11:355–9.

33. Ofner JG, Bartl B, Konig S, Thumfart WF. Photodynamic therapy in selected cases at the ENT Clinic, Innsbruck: case reports. J Photochem Photobiol B 1996;36:185–7.
34. Kubler A, Finley RK, Born IA, Mang TS. Effect of photodynamic therapy on the healing of a rat skin flap and its implication for head and neck reconstructive surgery. Lasers Surg Med 1996;18:397–405.
35. Maiman TH. Stimulated optical radiation in ruby. Nature 1960;187:493.
36. Reinisch L. Laser physics and tissue interactions. Otolaryngol Clin North Am 1996;29:893–914.
37. Schawlow AL, Townes CH. Infrared and optical lasers. Physiol Rev 1958;112:1940.
38. Fuller TA. The physics of surgical lasers. Lasers Surg Med 1980;1:5–14.
39. Fuller TA. The characteristics of operation of surgical lasers. Surg Clin North Am 1984;64:843–9.
40. Doukas AG, McAuliffe DJ, Flotte TJ. Biological effects of laser-induced shock waves: structural and functional cell damage in vitro. Ultrasound Med Biol 1993;19:137–46.
41. LeCarpentier GL, Motamedi M, McMath LP, et al. Continuous wave laser ablation of tissue: analysis of thermal and mechanical events. IEEE Trans Biomed Eng 1993;40:188–200.
42. Verdaasdonk RM, Borst C, van Gemert MJ. Explosive onset of continuous wave laser tissue ablation. Phys Med Biol 1990;35:1129–44.
43. Reinisch L. Laser induced heating and thermal propagation: a model of tissue interaction with light. Proc Conf Lasers and Elect Optics 1989;11:TuR4.
44. McKenzie AL. How far does thermal damage extend beneath the surface of CO_2 laser incisions? Phys Med Biol 1983;28:905–12.
45. Walsh JT, Flotte TJ, Anderson RR, Deutsch TF. Pulsed CO_2 laser tissue ablation: effect of tissue type and pulse duration on thermal damage. Lasers Surg Med 1988;8:108–18.
46. Hall RR. The healing of tissues incised by carbon dioxide laser. Br J Surg 1971;58:222–5.
47. Norris CW, Mullarry MB. Experimental skin incision made with the carbon dioxide laser. Laryngoscope 1982;92:416–9.
48. Finsterbush A, Rousso M, Ashur H. Healing and tensile strength of CO_2 laser incisions and scalpel wounds in rabbits. Plast Reconstr Surg 1982;70:360–2.
49. Buell BR, Schuller DE. Comparison of tensile strength in CO_2 laser and scalpel skin incisions. Arch Otolaryngol 1983;109:465–7.
50. Middleton WG, Tees DA, Ostrowski M. Comparative gross and histological effects of the CO_2 laser, Nd-YAG laser, scalpel, Shaw scalpel and cutting cautery on skin in rats. J Otolaryngol 1993;22:167–70.
51. Kyzer MD, Aly AS, Davidson JM, et al. Sub ablation effects of the KTP laser on wound healing. Lasers Surg Med 1993;13:62–71.
52. Pogrel MA, Pham HD, Guntenhoner M, Stern R. Profile of hyaluronidase activity distinguishes carbon dioxide laser from scalpel wound healing. Ann Surg 1993;217:196–200.
53. Luomanen M, Virtanen I. Fibronectins in healing incision, excision and laser wounds. J Oral Pathol Med 1991;20:133–8.
54. Speyer M, Joe J, Davidson JM, et al. Thermal injury patterns and tensile strength of canine oral mucosa incisions. Laryngoscope 1996;106:845–50.
55. Howard J, Arango P, Ossoff RH, Reinisch L. Healing of laser incisions in rat dermis: comparisons of the carbon dioxide laser under manual and computer control and the scalpel. Lasers Surg Med 1997;20:90–6.
56. Bryant GL, Davidson JM, Ossoff RH, et al. Histological study of oral mucosa wound healing: a comparison of a 6.0-6.8 μm pulsed laser to a carbon dioxide laser. Laryngoscope 1998;108:13–7.
57. Fortune DS, Huang S, Soto J, et al. The effect of pulse duration on wound healing using a carbon dioxide laser. Laryngoscope 1998;108:843–8.
58 Kuo T, Speyer MT, Ries WR, Reinisch L. Collagen thermal damage and collagen synthesis after cutaneous laser resurfacing. Lasers Surg Med 1998;23: 66–71.
59. Speyer MT, Reinisch L, Cooper KA, Ries WR. Erythema after cutaneous laser resurfacing using a porcine model. Arch Otolaryngol Head Neck Surg 1998;124:1008–13.
60. Sanders D, Reinisch L. Wound healing and collagen thermal damage in 7.5 μs pulsed CO_2 laser skin incisions. Lasers Surg Med 2000;26:22–32.
61. Apfelberg DB, Maser MR, Lash H, Rivers J. The argon laser for cutaneus lesions. JAMA 1981;245: 2073–5.
62. Cosman B. Experience in the argon laser therapy for port-wine stains. Plast Reconstr Surg 1980;65: 119–29.
63. Tan OT, Carney JM, Margolis R, et al. Histologic responses of port-wine stains treated by argon, carbon dioxide, and tunable dye lasers. A preliminary report. Arch Dermatol 1986;122:1016–22.

64. Parkin JL, Dixon JA. Argon laser treatment of head and neck vascular lesions. Otolaryngol Head Neck Surg 1981;93:211–6.
65. Kluger PB, Shapshay SM, Hybels RL, Bohigian RK. Neodymium-YAG laser intranasal photocoagulation in hereditary hemorrhagic telangiectasia: an update report. Laryngoscope 1987;97:1397–401.
66. Perkins RC. Laser stapedotomy for otosclerosis. Laryngoscope 1980;90:228–40.
67. Hodgson RS, Wilson DF. Argon laser stapedotomy. Laryngoscope 1991;101:230–3.
68. Vollrath M, Schreiner C. Influence of argon laser stapedotomy on inner ear function and temperature. Otolaryngol Head Neck Surg 1983;91:521–6.
69. McGee TM. The argon laser in surgery for chronic ear disease and otosclerosis. Laryngoscope 1983;93: 1177–82.
70. DiBartolomeo JR, Ellis M. The argon laser in otology. Laryngoscope 1980;90:1786–96.
71. Escudero LH, Castro AO, Drumond M, et al. Argon laser in human tympanoplasty. Arch Otolaryngol 1979;105:252–3.
72. Hanna E, Eliachar I, Cothren R, et al. Laser welding of fascial grafts and its potential application in tympanoplasty: an animal model. Otolaryngol Head Neck Surg 1993;108:356–66.
73. Wanamaker HH, Silverstein H. Compatibility of the argon and KTP lasers with middle ear implants. Laryngoscope 1993;103:609–13.
74. Horn KL, Gherini S. Otologic argon laser surgery using an endolaryngeal-otoprobe system. In: Abstracts of the Second International Laser Surgery Congress, Nashville (TN); 1986.
75. Horn KL, Gherini S, Griffin GM. Argon laser stapedectomy using an Endo-otoprobe system. Otolaryngol Head Neck Surg 1990;102:193–8.
76. Rauch SD, Bartley ML. Argon laser stapedectomy: comparison to traditional fenestration techniques. Am J Otol 1992;13:556–60.
77. Vincent R, Gratacap B, Causse JB, Gherini S. Argon laser and Gherini-Causse Endo-Otoprobe in otologic surgery. Ear Nose Throat J 1996;75:773–8.
78. Apfelberg DB, Lane B, Marx MP. Combined (team) approach to hemangioma management: arteriography with superselective embolization plus YAG laser/sapphire-tip resection. Plast Reconstr Surg 1991;88:71–82.
79. Mehta AC, Livingston DR, Golish JA. Artificial sapphire contact endoprobe with Nd-YAG laser in the treatment of subglottic stenosis. Chest 1987;91:473–4.
80. Hukki J, Krogerus L, Costren M, Schroder T. Effects of different contact laser scalpels on skin and subcutaneous fat. Lasers Surg Med 1988;8:276–82.
81. Iwasaki M, Sasako M, Konishi T, et al. Nd-YAG laser for general surgery. Lasers Surg Med 1985;5:429–38.
82. Matthewson K, Swain CP, Bland M, et al. Randomized comparison of Nd YAG laser, heater probe, and no endoscopic therapy for bleeding peptic ulcers. Gastroenterology 1990;98:1239–44.
83. Castro DJ, Saxton RE, Ward PH, et al. Flexible Nd:YAG laser palliation of obstructive tracheal metastatic malignancies. Laryngoscope 1990;100: 1208–14.
84. Dumon JF, Reboud E, Garbe L, et al. Treatment of tracheobronchial lesions by laser photoresection. Chest 1982;81:278–84.
85. Taber SW, Buschemeyer WC, Fingar VH, Wieman TJ. The treatment of malignant endobronchial obstruction with laser ablation. Surgery 1999;126:730–5.
86. Eckhauser ML. The neodymium-YAG laser and gastrointestinal malignancy. Arch Surg 1990;125:1152–4.
87. Eckhauser ML. Palliative therapy of upper gastrointestinal malignancies using the Nd-YAG laser. Am Surg 1990;56:158–62.
88. Rontal E, Rontal M, Jacob HJ, Klass A. Laser palliation for esophageal carcinoma. Laryngoscope 1986;96:846–50.
89. Dumon JF, Shapshay S, Bourcereau J, et al. Principles for safety in application of neodymium-YAG laser in bronchology. Chest 1984;86:163–8.
90. Shapshay SM, David LM, Zeitels S. Neodymium:YAG laser photocoagulation of hemangiomas of the head and neck. Laryngoscope 1987;97:323–30.
91. Apfelberg DB, Maser MR, White DN, et al. Benefits of contact and noncontact YAG laser for periorbital hemangiomas. Ann Plast Surg 1990;24:397–408.
92. Illum P, Bjerring P. Hereditary hemorrhagic telangiectasia treated by laser surgery. Rhinology 1988; 26:19–24.
93. Blackwell KE, Castro DJ, Saxton RE, et al. MRI and ultrasound guided interstitial Nd:YAG laser phototherapy for palliative treatment of advanced head and neck tumors: clinical experience. J Clin Laser Med Surg 1993;11:7–14.
94. Vagnetti A, Gobbi E, Algieri GM, D'Ambrosio L. Wedge turbinectomy: a new combined photocoagulative Nd:YAG laser technique. Laryngoscope 2000; 110:1034–6.
95. Eyrich GK, Bruder E, Hilfiker P, et al. Temperature mapping of magnetic resonance-guided laser inter-

stitial thermal therapy (LITT) in lymphangiomas of the head and neck. Lasers Surg Med 2000;26:467–76.
96. Garden JM, Bakus AD. Clinical efficacy of the pulsed dye laser in the treatment of vascular lesions. J Dermatol Surg Oncol 1993;19:321–6.
97. Tan OT, Morrison P, Kurban AK. 585 nm for the treatment of port-wine stains. Plast Reconstr Surg 1991;88:547–8.
98. Bower CM, Waner M, Flock S, Schaeffer R. Flash pump dye laser treatment of laryngeal papillomas. Ann Otol Rhinol Laryngol 1998;107:1001–5.
99. Ito H, Baba S. Pulsed dye laser lithotripsy of submandibular gland salivary calculus. J Laryngol Otol 1996;110:942–6.
100. Ikeda M, Takahashi H, Karaho T, et al. Amelanotic melanoma metastatic to the epiglottis. J Laryngol Otol 1991;105:776–9.
101. Kass EG, Massaro BM, Komorowski RA, Toohill RJ. Wound healing of KTP and argon laser lesions in the canine nasal cavity. Otolaryngol Head Neck Surg 1993;108:283–92.
102. Levine HL. Endoscopy and the KTP/532 laser for nasal sinus disease. Ann Otol Rhinol Laryngol 1989;98:46–51.
103. Bartels LJ. KTP laser stapedotomy: is it safe? Otolaryngol Head Neck Surg 1990;103:685–92.
104. Thedinger BS. Applications of the KTP laser in chronic ear surgery. Am J Otol 1990;11:79–84.
105. Lesinski SG, Palmer A. Lasers for otosclerosis: CO_2 vs. argon and KTP-532. Laryngoscope 1989;99:1–8.
106. Pillsbury HC. Comparison of small fenestra stapedotomies with and without KTP 532 laser. Arch Otolaryngol Head Neck Surg 1989;115:1027.
107. Paiva MB, Saxton RE, Graeber IP, et al. Improved photochemotherapy of malignant cells with daunomycin and the KTP laser. Lasers Surg Med 1998;23:33–9.
108. Herdman RC, Charlton A, Hinton AE, Freemont AJ. An in vitro comparison of the erbium:YAG laser and the carbon dioxide laser in laryngeal surgery. J Laryngol Otol 1993;107:908–11.
109. Nelson JS, Orenstein A, Liaw LH, Berns MW. Mid-infrared erbium:YAG laser ablation of bone: the effect of laser osteotomy on bone healing. Lasers Surg Med 1989;9:362–74.
110. Gonzalez C, van de Merwe WP, Smith M, Reinisch L. Comparison of the erbium-yttrium aluminum garnet and carbon dioxide lasers for in vitro bone and cartilage ablation. Laryngoscope 1990;100: 14–7.
111. Li ZZ, Reinisch L, van de Merwe WP. Bone ablation with Er:YAG and CO_2 laser: study of thermal and acoustic effects. Lasers Surg Med 1992;12:79–85.
112. Wong BJ, Dickinson MR, Berns MW, Neev JJ. Identification of photoacoustic transients during pulsed laser ablation of the human temporal bone: an experimental model. Clin Laser Med Surg 1996; 14:385–92.
113. Jovanovic S, Jamali Y, Anft D, et al. Influence of pulsed laser irradiation on the morphology and function of the guinea pig cochlea. Hear Res 2000;144:97–108.
114. Hausler R, Schar PJ, Pratisto H, et al. Advantages and dangers of erbium laser application in stapedotomy. Acta Otolaryngol (Stockh) 1999;119: 207–13.
115. Jovanovic S, Schonfeld U, Prapavat V, et al. Effects of pulsed laser systems on stapes footplate. Lasers Surg Med 1997;21:341–50.
116. Shah UK, Poe DS, Rebeiz EE, et al. Erbium laser in middle ear surgery: in vitro and in vivo animal study. Laryngoscope 1996;106:418–22.
117. Shapshay SM, Rebeiz EE, Bohigian RK, et al. Holmium:yttrium aluminum garnet laser-assisted endoscopic sinus surgery: laboratory experience. Laryngoscope 1991;101:142–9.
118. Shapshay SM, Rebeiz EE, Pankratov MM. Holmium:yttrium aluminum garnet laser-assisted endoscopic sinus surgery: clinical experience. Laryngoscope 1992;102:1177–80.
119. Kautzky M, Trodhan A, Susani M, Schenk P. Infrared laser stapedotomy. Eur Arch Otorhinolaryngol 1991;248:449–51.
120. Treat MR, Trokel SL, Reynolds RD, et al. Preliminary evaluation of a pulsed 2.15-micron laser system for fiberoptic endoscopic surgery. Lasers Surg Med 1988;8:322–6.
121. April MM, Rebeiz EE, Aretz HT, Shapshay SM. Endoscopic holmium laser laryngotracheoplasty in animal models. Ann Otol Rhinol Laryngol 1991;100:503–7.
122. Metson R. Holmium:YAG laser endoscopic sinus surgery: a randomized, controlled study. Laryngoscope 1996;106:1–18.
123. Fong M, Clarke K, Cron C. Clinical applications of the holmium:YAG laser in disorders of the paediatric airway. J Otolaryngol 1999;28:337–43.
124. Schomacker KT, Domankevitz Y, Flotte TJ, Deutsch TF. Co:MgF2 laser ablation of tissue: effect of wavelength on ablation threshold and thermal damage. Lasers Surg Med 1991;11:141–51.

125. Schlenk E, Profeta G, Nelson JS, et al. Laser assisted fixation of ear prostheses after stapedectomy. Lasers Surg Med 1990;10:444–7.
126. Frederickson KS, White WE, Wheeland RG, Slaughter DR. Precise ablation of skin with reduced collateral damage using the femtosecond-pulsed, terawatt titanium-sapphire laser. Arch Dermatol 1993;129:989–93.
127. Janda P, Sroka R, Tauber S, et al. Diode laser treatment of hyperplastic inferior nasal turbinates. Lasers Surg Med 2000;27:129–39.
128. Poe DS. Laser-assisted endoscopic stapedectomy: a prospective study. Laryngoscope 2000;110:1–37.
129. Volk MS, Wang Z, Pankratov MM, et al. Mucosal intact laser tonsillar ablation. Arch Otolaryngol Head Neck Surg 1996;122:1355–9.
130. Shohet JA, Reinisch L, Ossoff RH. Prevention of pharyngocutaneous fistulas by means of laser-weld techniques. Laryngoscope 1995;105:717–22.
131. Ossoff RH, Karlan MS. Instrumentation for CO_2 laser surgery of the larynx and tracheobronchial tree. Surg Clin North Am 1984;64:973–80.
132. Ossoff RH, Karlan MS. Instrumentation for microlaryngeal laser surgery. Otolaryngol Head Neck Surg 1983;91:456–60.
133. Ossoff RH, Karlan MS. A set of bronchoscopes for carbon dioxide laser surgery. Otolaryngol Head Neck Surg 1983;91:336–7.
134. Karlan MS, Ossoff RH. Laser surgery for benign laryngeal disease. Conservation and ergonomics. Surg Clin North Am 1984;64:981–94.
135. Ossoff RH, Karlan MS. Universal endoscopic coupler for carbon dioxide laser surgery. Ann Otol Rhinol Laryngol 1982;91:608–9.
136. Leibowitz HM, Peacock GR. Corneal injury produced by carbon dioxide laser radiation. Arch Ophthalmol 1969;87:713–21.
137. Fine BS, Fine S, Peacock GR, et al. Preliminary observations on ocular effects of high-power, continuous CO_2 laser irradiation. Am J Ophthalmol 1967;64:209–22.
138. Chu FWK, Silverman S, Dedo HH. CO_2 laser treatment of oral leukoplakia. Laryngoscope 1988;98: 125–30.
139. Duncavage JA, Ossoff RH. Use of the CO_2 laser for malignant disease of the oral cavity. Lasers Surg Med 1986;6:442–4.
140. Guerry TL, Silverman S, Dedo HH. Carbon dioxide laser resection of superficial oral carcinoma: indications, technique, and results. Ann Otol Rhinol Laryngol 1986;95:547–55.
141. McDonald GA, Simpson GT. Transoral resection of lesions of the oral cavity with the carbon dioxide laser. Otolaryngol Clin North Am 1983;16:839–47.
142. Nagorsky MJ, Sessions DG. Laser resection for early oral cavity cancer: results and complications. Ann Otol Rhinol Laryngol 1987;96:556–60.
143. Strong MS, Vaughan CW, Jako GJ, Polanyi T. Transoral resection of cancer of the oral cavity: the role of the CO_2 laser. Otolaryngol Clin North Am 1979; 12:207–18.
144. Liston SL, Giordano A. Tongue release using the CO_2 laser. Laryngoscope 1981;91:1010–1.
145. Roodenburg JL, Panders AK, Vermey A. Carbon dioxide laser surgery of oral leukoplakia. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1991; 71:670–4.
146. Chiesa F, Tradati N, Sala L, et al. Follow-up of oral leukoplakia after carbon dioxide laser surgery. Arch Otolaryngol Head Neck Surg 1990;116:177–80.
147. Gooris PJ, Roodenburg JL, Vermey A, Nauta JM. Carbon dioxide laser evaporation of leukoplakia of the lower lip: a retrospective evaluation. Oral Oncol 1999;35:490–5.
148. Schoelch ML, Sekandari N, Regezi JA, Silverman S. Laser management of oral leukoplakias: a follow-up study of 70 patients. Laryngoscope 1999;109:949–53.
149. Shapshay SM, Davis RK, Vaughan CW, et al. Palliation of airway obstruction from tracheobronchial malignancy: use of the CO_2 laser bronchoscope. Otolaryngol Head Neck Surg 1983;91:615–9.
150. Laccourreye O, Lawson G, Muscatello L, et al. Carbon dioxide laser debulking for obstructing endolaryngeal carcinoma: a 10-year experience. Ann Otol Rhinol Laryngol 1999;108:490–4.
151. Healy GB, McGill T, Jako GJ, et al. Management of choanal atresia with the carbon dioxide laser. Ann Otol Rhinol Laryngol 1978;87:658–62.
152. Mittelman H. CO_2 laser turbinectomies for chronic obstructive rhinitis. Lasers Surg Med 1982;2:29–36.
153. Simpson GT, Strong MS. Recurrent respiratory papillomatosis: the role of the carbon dioxide laser. Otolaryngol Clin North Am 1983;16:887–94.
154. Strong MS, Vaughan CW, Cooperband SR, et al. Recurrent respiratory papillomatosis: management with the CO_2 laser. Ann Otol Rhinol Laryngol 1976;85:508–16.
155. Siegel MB, Keane WM, Atkins JF, Rosen MR. Control of epistaxis in patients with hereditary hemorrhagic telangiectasia. Otolaryngol Head Neck Surg 1991;105:675–9.

156. McCombe AW, Cook J, Jones AS. A comparison of laser cautery and submucosal diathermy for rhinitis. Clin Otolaryngol 1992;17:297–9.
157. Testa B, Mesolella M, Squeglia C, et al. Carbon dioxide laser turbinate surgery for chronic obstructive rhinitis. Lasers Surg Med 2000;27:49–54.
158. Catalfumo FJ, Golz A, Westerman ST, et al. The epiglottis and obstructive sleep apnoea syndrome. J Laryngol Otol 1988;112:940–3.
159. Papadakis CE, Skoulakis CE, Nikolidakis AA, et al. Swiftlase inferior turbinoplasty. Am J Rhinol 1999;13:479–82.
160. Lippert BM, Werner JA. Comparison of carbon dioxide and neodymium:yttrium-aluminum-garnet lasers in surgery of the inferior turbinate. Ann Otol Rhinol Laryngol 1997;106:1036–42.
161. Takeno S, Osada R, Furukido K, Yajin K. Analysis of local cytokine gene expression in patients with allergic rhinitis treated with CO_2 laser surgery. Laryngoscope 2000;110:1968–74.
162. Tanigawa T, Yashiki T, Hayashi K, Sato T. Carbon dioxide laser vaporization for turbinate: optimal conditions and indications. Auris Nasus Larynx 2000;27:137–40.
163. Wouters B, van Overbeek JJ, Buiter CT, Hoeksema PE. Laser surgery in lingual tonsil hyperplasia. Clin Otolaryngol 1989;14:291–6.
164. Puar RK, Puar HS. Lingual tonsillitis. South Med J 1986;79:1126–8.
165. Shapshay SM, Strong MS, Anastasi GW, Vaughan CW. Removal of rhinophyma with the carbon dioxide laser. Arch Otolaryngol 1980;106:257–9.
166. Gjuric M, Rettinger G. Comparison of carbon dioxide laser and electrosurgery in the treatment of rhinophyma. Rhinology 1993;31:37–9.
167. Har-El G, Shapshay SM, Bohigian RK, et al. The treatment of rhinophyma. 'Cold' vs laser techniques. Arch Otolaryngol Head Neck Surg 1993; 119:628–31.
168. Haas A, Wheeland RG. Treatment of massive rhinophyma with the carbon dioxide laser. J Dermatol Surg Oncol 1990;16:645–9.
169. Kirschner RA. Cutaneous plastic surgery with the CO_2 laser. Surg Clin North Am 1984;64:871–83.
170. Levine H, Bailin P. Carbon dioxide laser treatment of cutaneous hemangiomas and tattoos. Arch Otolaryngol 1982;108:236–8.
171. Fairhurst MV, Roenigk RK, Brodland DG. Carbon dioxide laser surgery for skin disease. Mayo Clin Proc 1992;67:49–58.
172. Dinehart SM, Gross DJ, Davis CM, Herzberg AJ. Granuloma faciale. Comparison of different treatment modalities. Arch Otolaryngol Head Neck Surg 1990;116:849–51.
173. Sunde D, Apfelberg DB, Sergott T. Traumatic tattoo removal: comparison of four treatment methods in an animal model with correlation to clinical experience. Lasers Surg Med 1990;10:158–64.
174. Kamami YV. Laser CO_2 for snoring. Preliminary results. Acta Otorhinolaryngol Belg 1990;44:451–6.
175. Rollheim J, Miljeteig H, Osnes T. Body mass index less than 28 kg/m^2 is a predictor of subjective improvement after laser-assisted uvulopalatoplasty for snoring. Laryngoscope 1999;109:411–4.
176. Brodsky L, Brookhauser P, Chait D, et al. Office-based insertion of pressure equalization tubes: the role of laser-assisted tympanic membrane fenestration. Laryngoscope 1999;109:2009–14.
177. Kattner MN, Clark GD. Recurrent respiratory papillomatosis. Nurs Anesth 1993;4:28–35.
178. Ossoff RH, Werkhaven JA, Dere H. Soft-tissue complications of laser surgery for recurrent respiratory papillomatosis. Laryngoscope 1991;101: 1162–6.
179. Shapshay SM, Rebeiz EE, Bohigian RK, Hybels RL. Benign lesions of the larynx: should the laser be used? Laryngoscope 1990;100:953–7.
180. Kay GA, Lobe TE, Custer MD, Hollabaugh RS. Endoscopic laser ablation of obstructing congenital duodenal webs in the newborn: a case report of limited success with criteria for patient selection. J Pediatr Surg 1992;27:279–81.
181. Werkhaven JA, Weed DT, Ossoff RH. Carbon dioxide laser serial microtrapdoor flap excision of subglottic stenosis. Arch Otolaryngol Head Neck Surg 1993;119:676–9.
182. Ossoff RH, Tucker JA, Werkhaven JA. Neonatal and pediatric microsubglottiscope set. Ann Otol Rhinol Laryngol 1991;100:325–6.
183. Bagwell CE. CO_2 laser excision of pediatric airway lesions. J Pediatr Surg 1990;25:1152–6.
184. McGill TJ, Friedman EM, Healy GB. Laser surgery in the pediatric airway. Otolaryngol Clin North Am 1983;16:865–70.
185. Lomeo P, McDonald J, Finneman J. Adult laryngeal hemangioma: report of four cases. Ear Nose Throat J 2000;79:594–8.
186. Vaughan CW. Transoral laryngeal surgery using the CO_2 laser: laboratory experiments and clinical experience. Laryngoscope 1978;88:1399–420.

187. Vaughan CW. Use of the carbon dioxide laser in the endoscopic management of organic laryngeal disease. Otolaryngol Clin North Am 1983;16:849–64.
188. Danish HM, Meleca RJ, Dworkin JP, Abbarah TR. Laryngeal obstructing saccular cysts: a review of this disease and treatment approach emphasizing complete endoscopic carbon dioxide laser excision. Arch Otolaryngol Head Neck Surg 1998;124: 593–6.
189. Dedo HH, Sooy CD. Endoscopic laser repair of posterior glottic, subglottic and tracheal stenosis by division or micro-trapdoor flap. Laryngoscope 1984;94:445–50.
190. Holinger LD. Treatment of severe subglottic stenosis without tracheotomy: a preliminary report. Ann Otol Rhinol Laryngol 1982;91:407–12.
191. Lebovics RS, Hoffman GS, Leavitt RY, et al. The management of subglottic stenosis in patients with Wegener's granulomatosis. Laryngoscope 1992; 102:1341–5.
192. Whitehead E, Salam MA. Use of the carbon dioxide laser with the Montgomery T-tube in the management of extensive subglottic stenosis. J Laryngol Otol 1992;106:829–31.
193. Ossoff RH, Karlan MS, Sisson GA. Endoscopic laser arytenoidectomy. Lasers Surg Med 1983;2:293–9.
194. Ossoff RH, Sisson GA, Duncavage JA, et al. Endoscopic laser arytenoidectomy for the treatment of bilateral vocal cord paralysis. Laryngoscope 1984; 94:1293–7.
195. Ossoff RH, Duncavage JA, Shapshay SM, et al. Endoscopic laser arytenoidectomy revisited. Ann Otol Rhinol Laryngol 1990;99:764–71.
196. Netterville JL, Aly A, Ossoff RH. Evaluation and treatment of complications of thyroid and parathyroid surgery. Otolaryngol Clin North Am 1990; 23:529–52.
197. Komisar A. Laser laryngoscopic management of internal laryngocele. Laryngoscope 1987;97:368–9.
198. van Overbeek JJ, Hoeksema PE, Edens ET. Microendoscopic surgery of the hypopharyngeal diverticulum using electrocoagulation or carbon dioxide laser. Ann Otol Rhinol Laryngol 1984;93:34–6.
199. Benjamin B, Gallagher R. Microendoscopic laser diverticulotomy for hypopharyngeal diverticulum. Ann Otol Rhinol Laryngol 1993;102:675–9.
200. van Overbeek JJ. Meditation on the pathogenesis of hypopharyngeal (Zenker's) diverticulum and a report of endoscopic treatment in 545 patients. Ann Otol Rhinol Laryngol 1994;103:178–85.
201. Lippert BM, Folz BJ, Rudert HH, Werner JA. Management of Zenker's diverticulum and postlaryngectomy pseudodiverticulum with the CO_2 laser. Otolaryngol Head Neck Surg 1999;121:809–14.
202. Ossoff RH, Duncavage JA, Dere H. Microsubglottoscopy: an expansion of operative microlaryngoscopy. Otolaryngol Head Neck Surg 1991;104: 842–8.
203. Duncavage JA, Piazza LS, Ossoff RH, Toohill RJ. The microtrapdoor technique for the management of laryngeal stenosis. Laryngoscope 1987;97:825–8.
204. Beste DJ, Toohill RJ. Microtrapdoor flap repair of laryngeal and tracheal stenosis. Ann Otol Rhinol Laryngol 1991;100:420–3.
205. Lynch RC. Intrinsic carcinoma of the larynx with a second report of the cases operated on by suspension and dissection. Trans Am Laryngol Assoc 1920;42:119–20.
206. New GB, Dorton HE. Suspension laryngoscopy and the treatment of malignant disease of the hypopharynx and larynx. Mayo Clin Proc 1941;16:411.
207. Lillie JC, DeSanto LW. Transoral surgery of early cordal carcinoma. Trans Am Acad Ophthalmol Otolaryngol 1973;77:92–6.
208. Damm M, Sittel C, Streppel M, Eckel HE. Transoral CO_2 laser for surgical management of glottic carcinoma in situ. Laryngoscope 2000;110:1215–21.
209. Koufman JA. The endoscopic management of early squamous carcinoma of the vocal cord with the carbon dioxide surgical laser: clinical experience and a proposed subclassification. Otolaryngol Head Neck Surg 1986;95:531–7.
210. Wetmore SJ, Key JM, Suen JY. Laser therapy for T1 glottic carcinoma of the larynx. Arch Otolaryngol Head Neck Surg 1986;112:853–5.
211. Rothfield RE, Myers EN, Johnson JT. Carcinoma in situ and microinvasive squamous cell carcinoma of the vocal cords. Ann Otol Rhinol Laryngol 1991; 100:793–6.
212. Rice DH, Wetmore SJ, Singer M. Recurrent squamous cell carcinoma of the true vocal cord. Head Neck 1991;13:549–52.
213. Myers EN, Wagner RL, Johnson JT. Microlaryngoscopic surgery for T1 glottic lesions: a cost-effective option. Ann Otol Rhinol Laryngol 1994;103:28–30.
214. Ossoff RH, Sisson GA, Shapshay SM. Endoscopic management of selected early vocal cord carcinoma. Ann Otol Rhinol Laryngol 1985;94:560–4.
215. Blakeslee D, Vaughan CW, Shapshay SM, et al. Excisional biopsy in the selective management of T1

glottic cancer: a three-year follow-up study. Laryngoscope 1984;94:488–94.
216. Hirano M, Hirade Y, Kawasaki H. Vocal function following carbon dioxide laser surgery for glottic carcinoma. Ann Otol Rhinol Laryngol 1985;94: 232–5.
217. Hormann K, Baker-Schreyer A, Keilmann A, Biermann G. Functional results after CO_2 laser surgery compared with conventional phonosurgery. J Laryngol Otol 1999;113:140–4.
218. Lesinski SG, Newrock R. Carbon dioxide lasers for otosclerosis. Otolaryngol Clin North Am 1993;26: 417–41.
219. Andrews AH, Horowitz SL. Bronchoscopic CO_2 laser surgery. Lasers Surg Med 1980;1:35–45.
220. Katlic MR, Burick AJ, Lucchino DB. Experiences with laser bronchoscopy. Pa Med 1991;94:24–7.
221. Simpson GT, Strong MS, Healy GB, et al. Predictive factors of success or failure in the endoscopic management of laryngeal and tracheal stenosis. Ann Otol Rhinol Laryngol 1982;91:384–8.
222. Duncavage JA, Ossoff RH, Toohill RJ. Carbon dioxide laser management of laryngeal stenosis. Ann Otol Rhinol Laryngol 1985;94:565–9.
223. Stanley PR, Anderson T, Pagliero KM. Laser photoresection in the preoperative assessment of a bronchial adenoma. Thorax 1988;43:741–2.
224. Shapshay SM, Strong MS. Tracheobronchial obstruction from metastatic distant malignancies. Ann Otol Rhinol Laryngol 1982;91:648–51.
225. Sutedja G, Koppenol W, Stam J. Nd-YAG laser under local anaesthesia in obstructive endobronchial tumours. Respiration 1991;58:238–40.
226. Edell ES, Cortese DA, McDougall JC. Ancillary therapies in the management of lung cancer: photodynamic therapy, laser therapy, and endobronchial prosthetic devices. Mayo Clin Proc 1993;68:685–90.
227. Keon TP. Anesthetic management during laser surgery. Int Anesthesiol Clin 1992;30:99–107.
228. Spiess BD, Ivankovich AD. Anesthetic management of laser airway surgery. Semin Surg Oncol 1990;6: 189–93.
229. Burgess GE, LeJeune FE. Endotracheal tube ignition during laser surgery of the larynx. Arch Otolaryngol 1979;105:561–2.
230. Cozine K, Rosenbaum LM, Askanazi J, Rosenbaum SH. Laser-induced endotracheal tube fire. Anesthesiology 1981;55:583–5.
231. Hirshman CA, Smith J. Indirect ignition of the endotracheal tube during carbon dioxide laser surgery. Arch Otolaryngol Head Neck Surg 1980;106:639–41.
232. Ossoff RH, Duncavage JA, Eisenman TS, Karlan MS. Comparison of tracheal damage from laser-ignited endotracheal tube fires. Ann Otol Rhinol Laryngol 1983;92:333–6.
233. Treyve E, Yarington CT, Thompson GE. Incendiary characteristics of endotracheal tubes with the carbon dioxide laser. Ann Otol Rhinol Laryngol 1981;90:328–30.
234. Hawkins DB, Joseph MM. Avoiding a wrapped endotracheal tube in laser laryngeal surgery: experiences with apneic anesthesia and metal laser-flex endotracheal tubes. Laryngoscope 1990;100:1283–7.
235. Fried MP, Mallampati SR, Liu FC, et al. Laser resistant stainless steel endotracheal tube: experimental and clinical evaluation. Lasers Surg Med 1991;11:301–6.
236. Sosis MB, Dillon F. Prevention of CO_2 laser-induced endotracheal tube fires with the laser-guard protective coating. J Clin Anesth 1992;4:25–7.
237. Sosis MB. Which is the safest endotracheal tube for use with the CO_2 laser? A comparative study. J Clin Anesth 1992;4:217–9.
238. Sosis MB, Dillon FX. A comparison of CO_2 laser ignition of the Xomed, plastic, and rubber endotracheal tubes. Anesth Analg 1993;76:391–3.
239. Hayes DM, Gaba DM, Goode RL. Incendiary characteristics of a new laser-resistant endotracheal tube. Otolaryngol Head Neck Surg 1986;95:37–40.
240. Fontenot R, Bailey BJ, Stiernberg CM, Jenicek JA. Endotracheal tube safety during laser surgery. Laryngoscope 1987;97:919–21.
241. Tomita Y, Mihashi S, Nagata K, et al. Mutagenicity of smoke condensates induced by CO_2 laser irradiation and electrocauterization. Mutat Res 1981;89:145–9.
242. Wenig BL, Stenson KM, Wenig BM, Tracey D. Effects of plume produced by the Nd:YAG laser and electrocautery on the respiratory system. Lasers Surg Med 1993;13:242–5.
243. Baggish MS, Poiesz BJ, Joret D, et al. Presence of human immunodeficiency virus DNA in laser smoke. Lasers Surg Med 1991;11:197–203.
244. Abramson AL, DiLorenzo TP, Steinberg BM. Is papillomavirus detectable in the plume of laser-treated laryngeal papilloma? Arch Otolaryngol Head Neck Surg 1990;116:604–7.
245. Ossoff RH. Laser safety in otolaryngology-head and neck surgery: anesthetic and educational considerations for laryngeal surgery. Laryngoscope 1989;99:1–26.

246. Ossoff RH, Karlan MS. Safe instrumentation in laser surgery. Otolaryngol Head Neck Surg 1984; 92: 644–8.
247. Ossoff RH, Hotaling A, Karlan MS, Sisson GA. The CO_2 laser in otolaryngology-head and neck surgery: a retrospective analysis of complications. Laryngoscope 1983;93:1287–9.
248. Alberti PW. The complications of CO_2 laser surgery in otolaryngology. Acta Otolaryngol (Stockh) 1981;91:375–81.
249. Fried MP. A survey of the complications of laser laryngoscopy. Arch Otolaryngol 1984;110:31–4.
250. Healy GB, Strong MS, Shapshay S, et al. Complications of CO_2 laser surgery of the aerodigestive tract: experience of 4416 cases. Otolaryngol Head Neck Surg 1984;92:13–8.
251. Meyers A. Complications of CO_2 laser surgery of the larynx. Ann Otol Rhinol Laryngol 1981;90:132–4.
252. Mohr RM, McDonnell BC, Unger M, Mauer TP. Safety considerations and safety protocol for laser surgery. Surg Clin North Am 1984;64: 851–9.
253. Spilman LS. Nursing precautions for CO_2 laser surgery. Symposium Proceedings of The Laser Institute of America 1983;37:63–5.
254. Crockett DM, McCabe BF, Shive CJ. Complications of laser surgery for recurrent respiratory papillomatosis. Ann Otol Rhinol Laryngol 1987; 96:639–44.
255. Wetmore SJ, Key JM, Suen JY. Complications of laser surgery for laryngeal papillomatosis. Laryngoscope 1985;95:798–801.
256. Cotton RT, Tewfik TL. Laryngeal stenosis following carbon dioxide laser in subglottic hemangioma: report of three cases. Ann Otol Rhinol Laryngol 1985;94:494–7.
257. Durkin GE, Duncavage JA, Toohill RJ, et al. Wound healing of true vocal cord squamous epithelium after CO_2 laser ablation and cup forceps stripping. Otolaryngol Head Neck Surg 1986;95:273–7.

CHAPTER 65

Introduction to Peroral Endoscopy

Ellen S. Deutsch, MD, Jane Y. Yang, MD, James S. Reilly, MD

Aerodigestive endoscopy offers a privileged and exciting view into structures and functions that are vital for the survival of human beings. It provides precise anatomic information about intraluminal anatomy, as well as details about mucosal characteristics and dynamic processes not obtainable by any other method of evaluation. Advances in optics, instruments, and anesthesia, employed in conjunction with equipment and techniques developed by innovative and observant pioneer endoscopists, allow us extraordinary capabilities to evaluate the aerodigestive system.

Although it is important to have appropriate equipment, Chevalier Jackson also stated that

> Endoscopic ability cannot be bought with the instruments. As with all mechanical procedures, facility can be obtained only by educating the eye and the fingers in repeated exercise of a particular series of maneuvers. As with learning to play a musical instrument, a fundamental knowledge of technic, positions, and landmarks is necessary, after which only continued manual practice makes for proficiency.[1]

Aerodigestive endoscopy should be methodical, thorough, and timely. Repetition of a standardized protocol will improve efficiency, but the clinician should also continuously endeavor to improve the quality and ease of the procedure, incorporating improvements in equipment, techniques, or theories of disease pathogenesis.

The topics in this section have been arbitrarily divided, but principles addressed in discussing the selection or use of one instrument or technique may also be applicable to other instruments or techniques. Similarly, dividing diseases into categories defined by anatomic boundaries may provide a framework for methodical evaluation, but the potential for complex relationships between different anatomic areas should also be considered.

HISTORY OF AERODIGESTIVE ENDOSCOPY

Early attempts at aerodigestive endoscopy were hampered by primitive sources of light. In 1806, Bozzini fashioned an angled speculum with a mirror insert for examination of various body cavities, including the larynx, using a single wax candle for illumination.[2,3] In 1829, Babbington described a glottiscope, a three-bladed device including a stainless steel mirror and tongue retractors.[2] Desmoreaux, the "father of endoscopy," redesigned Bozzini's endoscope in 1853 by attaching a gaslight and condensers to project a beam of light down the tube; however, this was cumbersome and not widely accepted.[2] In 1854, Manuel Garcia, a voice teacher preoccupied with the possibility of observing the movements of the vocal cords, noticed sunlight flashing on the window panes of a house and considered the idea of light from one mirror reflected to another. Using a dental mirror, a hand mirror, and sunlight, he was able to visualize his own larynx and vocal cords.[2]

Turck, Czermak, and Morell MacKenzie refined the instruments and techniques for endoscopy,[2] but MacKenzie's reputation suffered following a consultation by German royal physicians to give an opinion about the hoarseness of Crown Prince Frederick William of Prussia, who was married to Queen Victoria's oldest daughter. MacKenzie's biopsy specimen from a tumor of the prince's vocal cord was interpreted by Professor Rudolf Virchow as a benign lesion. It was not until 9 months later, after additional

biopsies, that its malignant nature was identified. Meanwhile, the prince had been crowned emperor but died only 3 months after his coronation.[2]

Killian devised a different approach to the larynx, developing a direct laryngoscope as well as an apparatus for suspending the laryngoscope and using a headlight for illumination. In 1897, he used a rigid endoscope to examine the airways[4] and was the first to remove a foreign body in the bronchial tree using direct upper bronchoscopy without tracheostomy under local anesthesia.[5]

In the early 1900s, Chevalier Jackson refined the rigid bronchoscope,[4] invented distal lighting for endoscopic equipment,[2] and designed a variety of endoscopic instruments. He chronicled his observations with dramatic paintings and drawings.[2] His book, *Bronchoscopy and Esophagoscopy*, first published in 1922, remains a comprehensive treatise on the techniques of peroral endoscopy almost a century later.[1]

The use of magnification to enhance endoscopy was developed by Brunings in Germany and Jackson in the United States in the early 1900s.[2] The addition of the Zeiss operating microscope in the mid-1950s, with appropriate laryngoscopes such as those designed by Jako, allowed binocular magnification.[2]

Advances in illumination and optics have also facilitated the development of rigid esophagoscopy; however, an understanding of optimal patient positioning derives from the insight of physicians working in collaboration with sword swallowers. Similar to the development of laryngoscopy, Bozzini's initial theories of esophagoscopy involved esophagoscopes with a curve for the pharynx where light would be reflected by a mirror.[3] Adolf Kussmaul was the first to use a straight tube, with lighting developed by Desmoreaux using a gasoline lamp; light was reflected by two mirrors passing through a lens for concentration.[3] After studying the techniques of sword swallowers, he was able to look into the stomach using a 47 cm tube but was disappointed as he saw only bubbling,[3] and he never published his technique.

Leiter, an instrument maker, constructed a light source using a platinum wire heated to a white glow by electricity and a cooling system of circulating water. He visited Kussmaul and convinced Johann von Mikulicz-Radecki of the importance of the straight-tube esophagoscope. Von Mikulicz was also inspired by sword swallowers.[3] Huizinga quoted from von Mikulizc's 1881 publication, *Zur Technik der Gastroskopie und Oesophagoskopie*: "The first thing to do was to examine if this masterpiece was the result of an abnormal anatomy or if every normally built man could imitate it with a more innocent straight instrument. So Leiter induced several young men to exercise with rubber sticks and mostly they succeeded in introducing it into the stomach."[3] The two "crucial moments" during sword swallowing occur during passage of the sword at the upper and lower esophageal sphincters.[3] During modern rigid esophagoscopy, the position of the patient's head and neck is changed as insertion of the esophagoscope progresses, but this is not feasible for a sword swallower. The sword swallower must optimize the position of the proximal esophagus by voluntary anterior displacement of the lower jaw, hyoid bone, and tongue base, as well as simultaneously accommodating the forward direction of the distal esophagus.[3]

In the 1960s, flexible fiber-optic bronchoscopes were developed and introduced into clinical use by Shigeto Ikeda[4]; fiber-optic esophagoscopes, shortly thereafter, came into use. Fiber-optic telescopes comprise a bundle of flexible fibers, some carrying light to the object to be examined and others carrying the image back to the viewer.[6] The guidable lens tip allows flexible positioning of the scope, and wide angle and zoom lenses are available.[6] At approximately the same time, rod lens telescopes were invented by Professor H. H. Hopkins, providing improved optics and light transmission via rigid telescopes, with a capability for angled and wide-angle lenses.[7] Modern endoscopy also uses endoscopic cameras and recording equipment, allowing video or digital recording and slow-motion replay, multiple observers, and ease of photographic documentation.[8] Endoscopists continue to refine instruments, the understanding of disease, and patient management.[9]

INDICATIONS FOR LARYNGOSCOPY IN THE OFFICE

Laryngeal examination should be performed when signs and symptoms indicate the possibility of isolated benign or malignant vocal fold lesions, neurologic disease including motoneuron disorders, systemic disease, and hyperfunctional disorders (Table 65–1). Laryngoscopy can contribute to defin-

TABLE 65–1. Indications for Laryngoscopy

Chronic hoarseness
Voice changes
Chronic cough
Choking episodes
Odynophonia
Odynophagia
Chronic throat pain
Globus sensation
Hemoptysis
Referred otalgia
Shortness of breath
Dysarthria
Stridor
Suspicion of laryngeal foreign body
Suspicion of carcinoma
Dyspnea

Adapted from Ferlito A et al[10] and Hayes JT and Houston R.[12]

ing the extent and number of lesions, obtaining a biopsy, and establishing and maintaining an airway.[10] In adults and adolescents, a complete laryngeal examination can usually be accomplished in an office setting. In young children, dynamic or mass lesions of the supraglottis and glottis, such as laryngomalacia, vocal fold paralysis, and vocal fold papillomas, can often be evaluated. Awake laryngoscopy performed in the office is useful if the findings would be likely to obviate the need for endoscopy under general anesthesia. Examples include diagnosing vocal fold nodules, vocal fold paralysis, or laryngomalacia when the history and endoscopic findings are classic. In young patients,[11] patients with significant anxiety or certain anatomic abnormalities, or those who are developmentally or neurologically impaired, examination may be more difficult, and only sporadic glimpses of anatomy and function may be obtained. Visualization of the airway in the operating room may be preferable if there is a need to evaluate the airway below the level of the vocal folds in children, if there is significant airway distress, or if the patient is very ill or uncooperative.

Superimposed on the question of whether to proceed with airway endoscopy is the question of how urgently the endoscopy should be performed. Proceed urgently if there is significant respiratory distress, the patient appears uncomfortable or fatigued, or the patient is newly cyanotic or has a new supplemental oxygen requirement that cannot be explained solely by other organ system problems. Proceed urgently if the stridor is of acute onset, such as from foreign-body aspiration; if certain infections are suspected, such as exudative tracheitis or epiglottitis; or if there has been mechanical or thermal trauma to the airway. Proceed urgently if there is significant dysphonia, which suggests lesions at the level of the vocal folds, such as papilloma, exudative infections, and foreign bodies, which can progress to complete airway obstruction.

TECHNIQUES FOR EXAMINATION OF THE LARYNX IN THE OFFICE

Most adults and older children will tolerate examination of the larynx in an office setting, particularly when the patient is approached with adequate psychological preparation and gentle technique. It may be useful to both explain and then demonstrate that particular instruments are not threatening by touching them to a less vulnerable portion of the patient's body, such as his or her arm. When a camera and monitor are being used, the capability of the telescope can be demonstrated by "examining" a button or detail of the patient's clothing or jewelry, etc. Infants and young toddlers may require restraints.

The larynx should be viewed at rest, during respiration, and during vocalization, noting details of symmetry, motion, surface architecture, evidence of inflammation, and abnormal masses or growths.[12]

Mirror

Advocates of mirror or "indirect" examination of the larynx espouse the glare-free lighting of laryngeal structures, allowing observation of subtle color variations (Figure 65–1). The patient sits erect, facing the endoscopist, in the "sniffing" position, with the neck flexed on the chest and the head extended on the neck. The patient opens his mouth and protrudes his tongue. The endoscopist grasps the tongue, wrapped in gauze, between his or her thumb, on the dorsal surface, and third finger, on the

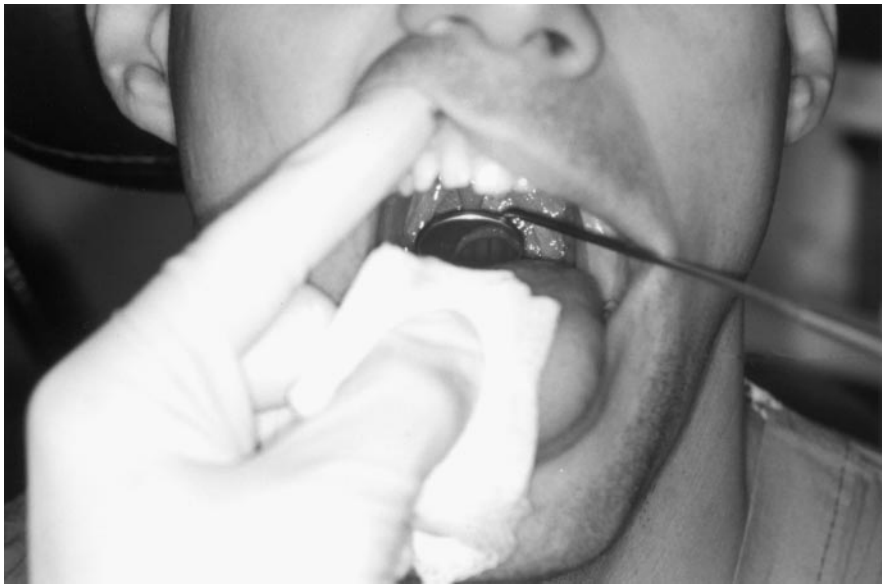


FIGURE 65–1. Mirror examination of the larynx. Note the epiglottis visible within the mirror's image and the vocal folds posterior to the epiglottis.

ventral surface, essentially rolling the tip of the tongue over the endoscopist's finger. The endoscopist's second finger can be used to retract the upper lip superiorly. A mirror on an angled handle is grasped, near the mirror end, in the endoscopist's right hand and placed in the posterior part of the pharynx, lifting the uvula. Encouraging the patient to "breathe through your mouth" or "pant like a puppy" may facilitate placement of the mirror. Warming the mirror or applying soap or a defogging solution will minimize fogging of the mirror as the patient breathes. A concave mirror on a headband, coordinated with an external light source, or a headlight is positioned so that the light is deflected from the mirror onto the pharynx and larynx and the image of the larynx is reflected back to the endoscopist. Having the patient phonate a prolonged, high-pitched /ee/ lengthens the vocal folds and moves the larynx upward vertically[6] to improve visualization of the entire vocal folds, including the anterior commissure.

Mirror examination requires a significant amount of endoscopic skill, as well as patient cooperation. Many patients have a sensitive gag reflex, allowing only brief glimpses into the laryngopharynx that may not be sufficient for proper diagnosis. Limitations include anatomic variations, such as a large tongue, micrognathia, and trismus, that preclude adequate examination and the inability of the patient to speak in a normal manner because of the requisite positioning.[6]

Fiber-optic Telescopes

Fiber-optic telescopes are frequently the preferred tools for laryngeal examinations in the office. The fiber-optic examination is often more comfortable for the patient and allows the otolaryngologist longer intervals to examine the hypopharynx and larynx (Figure 65–2). Fiber-optic nasopharyngoscopes (rather than bronchoscopes) are often of sufficient length but do not contain a suction channel so cannot be used to apply topical anesthesia or to suction excess secretions.[8] Guidable fiber-optic telescopes are available with outer diameters as narrow as 2.2 mm, allowing atraumatic passage through the nares of a normal neonate. Although smaller telescopes are more comfortable for patients, larger lenses provide better quality images. The patient may be examined sitting erect and facing the examiner or supine, with the endoscopist at the patient's head or side. As the fiber-optic telescope is passed through the nose, almost the entire upper aerodigestive tract from the nares to the larynx may be visualized. A limited view of the subglottis may also be obtained but can be deceptive, particularly in children, as the normal narrowing of the cricoid is accentuated by the angle of visualization, and this may give a false impression of subglottic stenosis. The fiber-optic telescope can be passed through the mouth, but a bite block or other device may be necessary to prevent the instrument from being bitten, and the nasal portion of the examination will not be obtained. Folded gauze or preformed plastic bite

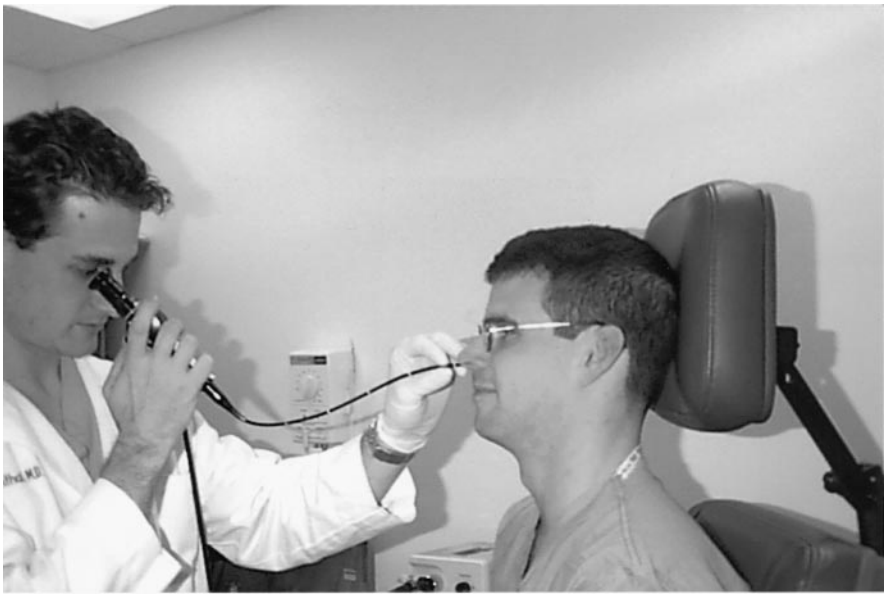


FIGURE 65–2. Fiber-optic laryngoscopy demonstrated by a left-handed endoscopist.

blocks can be used. In an edentulous infant, the endoscopist's finger can be used as a bite block. Among techniques for examination in an office setting, fiber-optic telescopes require the least amount of patient cooperation and cause the least interference with normal speech and voice production. During fiber-optic laryngoscopy, the larynx may be observed while the patient speaks, sings, whistles, plays a wind instrument, or demonstrates a particular voice disturbance[6] as the telescope does not interfere with positioning of the tongue, teeth, lips, etc. The apices of the piriform sinuses may be visualized during a Valsalva's maneuver. Fiber-optic examination is useful for evaluating dynamic lesions, such as laryngomalacia or vocal fold paralysis. Fiber-optic telescopes are also useful in the occasional circumstance when a rigid telescope cannot be safely placed or the larynx is difficult to visualize with a laryngoscope because of anatomic abnormalities such as micrognathia, cervical spine rigidity or instability, or immobility of the temporomandibular joint. In children and some adults, passage of the fiber-optic telescope through the glottis can cause laryngospasm; prophylactic application of local anesthesia may be indicated. Application of an anesthetic may interfere with assessments of sensation.

70- or 90-Degree Telescope

A rigid telescope with an angled lens of 70 or 90 degrees can be passed through the mouth to the oropharynx to view the larynx (Figure 65–3). Twenty centimeters are often a comfortable telescope length for the endoscopist. Patient positioning and the technique for inserting the telescope are similar to that for mirror examination. Placement of the telescope in the mouth will limit the variety of speech sounds that can be elicited. Children as young as 6 or 8 years of age may be able to cooperate for this examination.

Stroboscope

Although stroboscopy for viewing the larynx was first described by Oertel in 1878, early equipment was bulky and expensive, and it has taken another century of technological advances to provide user-friendly equipment.[13] Modern stroboscopes emit rapid pulses of light at a rate that may be set by the examiner or, more commonly, synchronized with the vocal frequency of the patient during phonation. The stroboscope is coupled with a rigid or fiber-optic endoscope and a video recording device, as well as a microphone placed over the lateral aspect of the neck to measure the acoustic characteristics of the patient's voice. The fundamental frequency is the individual's lowest vocal frequency, measured in hertz, and referred to as pitch. If the frequency of the light pulse is the same as the vocal frequency, the resulting image will appear to be static. If the light pulse frequency is slightly longer or shorter than the vocal frequency, the image will

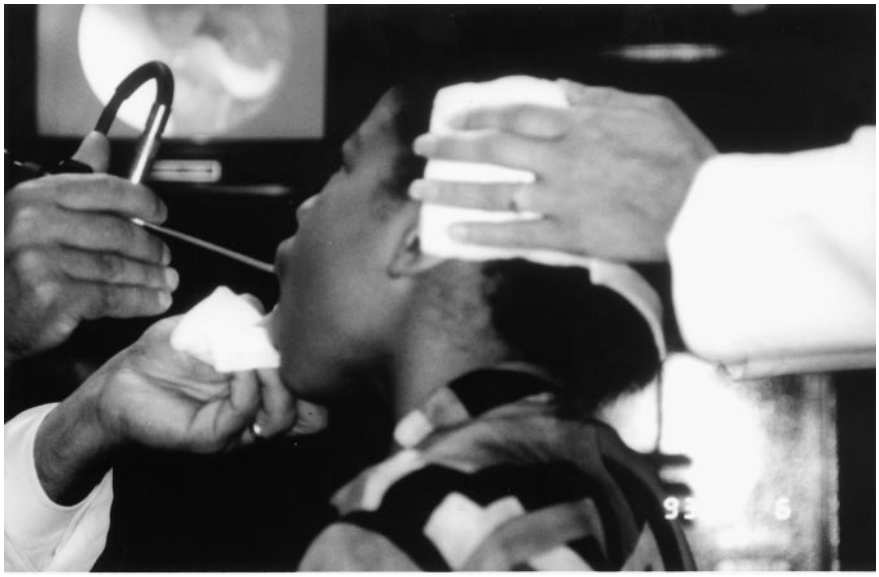


FIGURE 65–3. Laryngeal examination using an angled telescope. Note image of supraglottic structures visible on monitor.

appear to change slowly. Observation of the true vocal fold vibratory pattern may reveal subtle vocal fold pathology that could not otherwise be visualized.[6] Normal vocal folds vibrate symmetrically, that is, as mirror images of each other, during phonation. Successive vocal fold cycles should occur in a regular sequence; aperiodicity is abnormal. Glottal closure refers to the completeness of vocal fold apposition during phonation, and deficits in degree and configuration may occur. Vocal fold stiffness may cause disturbances in the mucosal wave or amplitude (extent of lateral motion). Abnormalities may result from presbyphonia, scarring, masses, or neurogenic lesions.

Fiber-optic Endoscopic Evaluation of Swallowing

Fiber-optic endoscopic evaluation of swallowing (FEES) has become an extremely useful adjunct to videofluoroscopic evaluation for patients with dysphagia or aspiration.[11] Videofluoroscopy provides information about the swallow mechanism during all phases of swallowing; however, FEES allows direct visualization of the velopharynx and larynx with better evaluation regarding sensory and neurologic status. Fiber-optic endoscopic evaluation of swallowing can be performed at bedside without radiation exposure. Recording the examination using a television monitor and videotape recorder allows repeated review, slow-motion review, and patient education. Application of topical anesthetic beyond the nasal cavity is not recommended as it may interfere with assessment of laryngopharyngeal sensation.

The first observation is made in the nasopharynx, where velopharyngeal competence is tested. Palatal closure is observed while the patient performs certain plosive speech tasks such as repeating "k, k, k" or "pick up the cupcake," and a palatal stress test can be accomplished by having the patient phonate a forced "ssss."[14]

The endoscope is passed beyond the nasopharynx to evaluate laryngopharyngeal anatomy. Pooling of secretions in the piriform sinuses or impaired mobility of the tongue base or vocal folds suggests a neurogenic or anatomic abnormality. Adequate pharyngeal squeeze is tested by asking the patient to phonate a high-pitched "/ee/."[14] In patients with pharyngeal paralysis, the inferior constrictors will not contract to close the piriform sinuses. In adults, a gross assessment of laryngeal and hypopharyngeal sensation may be accomplished by touching the tip of the endoscope to the piriform sinuses, supraglottis, and true vocal folds.

A more quantitative analysis may be obtained by positioning an endoscope with a suction channel about 5 mm from the surface of the aryepiglottic fold, delivering discrete pulses of air with pressures ranging from 2 to 10 mm Hg, and ascertaining the threshold pressure necessary to stimulate vocal fold adduction. Aviv et al have demonstrated a correlation between the degree of laryngopharyngeal sensory discrimination deficit and aspiration.[11]

Management of secretions or a food bolus in the hypopharynx and larynx are evaluated before, during, and after swallowing. Three consistencies of food are tested with food coloring added to improve visualization.[11] Pureed foods have the lowest risk of aspiration and are usually tested first. Management of particulate foods is tested using crackers. Water is evaluated last because it provides the greatest neurophysiologic challenge. Premature spillage and postswallow residue in the piriform sinuses or laryngeal vestibule (the area bounded anteriorly by the laryngeal surface of the epiglottis, laterally by the aryepiglottic folds, and posteriorly by the arytenoids) are abnormal. Laryngeal penetration, defined as passage of food below the false vocal folds, or aspiration, passage of food below the true vocal folds, is abnormal.[11] If penetration occurs, note the presence or absence of an appropriate cough reflex and adequate ability to expectorate the aspirated material. Assess compensatory strategies, such as positional alteration, altering the rate of administration or viscosity or texture of the food or liquid, or changes in feeding implements.[11] If the patient has a tracheostomy, the endoscope can be passed through the tracheostomy to further evaluate tracheal penetration.[14]

Local Anesthesia

The use of local anesthesia can contribute significantly to patient comfort and cooperation, as well as ameliorate the reflex response to mucosal stimulation, which includes tachycardia, hypertension, and laryngospasm.[15] Anesthesia can be accomplished by topical application of solutions or gels[16] or by local injection.[17] Local anesthesia can ablate the sensation of touch but may not be sufficient to alleviate discomfort from the traction forces required for suspension laryngoscopy.[15] Local anesthetics block the generation and the conduction of nerve impulses, presumably by increasing the threshold for electrical excitation in the nerve, slowing the propagation of the nerve impulse, and reducing the rate of rise of the action potential. Clinically, the order of loss of nerve function is (1) pain, (2) temperature, (3) touch, (4) proprioception, and (5) skeletal muscle tone.[18]

The internal branch of the superior laryngeal nerve provides sensory innervation to a small portion of the posterior aspect of the base of tongue, the laryngeal surface of the epiglottis, and the mucous membranes of the larynx down to, but not including, the true vocal folds.[17,19,20] The recurrent laryngeal nerve supplies innervation to the glottis and infraglottic portion of the larynx.[17] The remainder of the hypopharynx derives sensory innervation from the glossopharyngeal nerve.

Topical Anesthesia To anesthetize the nose and pharynx, topical anesthetics can be painted, dripped or aerosolized into the nose, gargled,[16] or sprayed onto the palate and posterior pharyngeal wall. A transmucosal superior laryngeal nerve block will anesthetize the hypopharynx and supraglottic larynx. To accomplish this, topical anesthetic agents are applied to the base of tongue and piriform sinuses using a moist cotton swab in a Jackson forceps[15] or are dripped or sprayed via a curved cannula or the side channel of a fiber-optic telescope.

Infiltration of Anesthetic Agents Glossopharyngeal nerve block has been advocated to eliminate the gag reflex and provide sufficient anesthesia for esophagoscopy. As with topical anesthetics, the glossopharyngeal nerve block can ablate sensation from tactile receptors in the mucosa of the lateral and posterior wall of the oropharynx, hypopharynx, and posterior third of the tongue. However, unlike topical anesthetics, it also ablates sensation from pressure receptors in the posterior third of the tongue. An angled 23-gauge tonsil needle is used to slowly inject lidocaine into the lateral oropharyngeal wall at a point 0.5 cm behind the midportion of the posterior tonsillar pillar. The needle is repositioned if blood is aspirated prior to injection, and injection is aborted if the patient complains of a headache. Reported complications include headache from intra-arterial injection and arrhythmias secondary to blocking carotid sinus afferent nerve fibers.[21]

A percutaneous superior laryngeal nerve block can be accomplished by injecting a local anesthetic agent in the region of the internal branch of the superior laryngeal nerve as it pierces the thyrohyoid membrane (Figure 65–4).[17,19,20] The greater cornu of the hyoid on the side to be anesthetized can be made prominent by pressing on the contralateral greater cornu with the forefinger. In adults, the needle is inserted 1 cm below and 2 cm anterior to the greater cornu. The needle is advanced posteriorly and slightly superiorly, as well as deep to the strap muscles. Before injecting, aspirate to ensure that the nee-

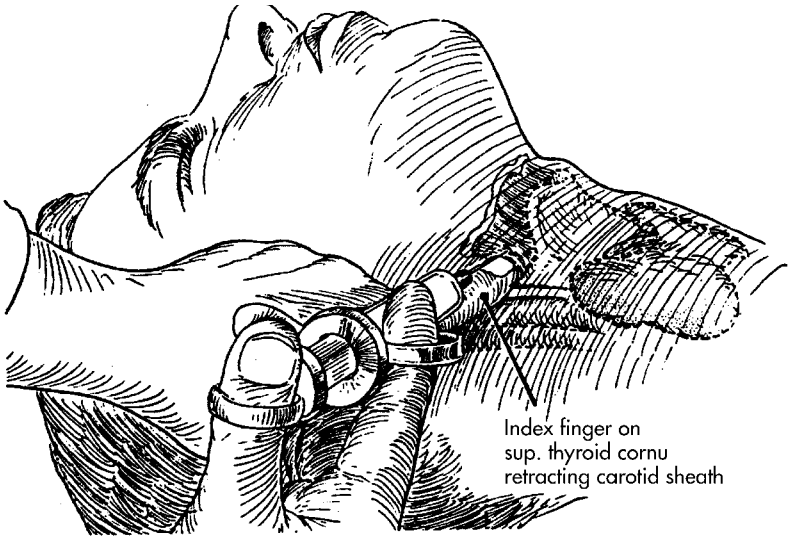


FIGURE 65–4. Technique for blocking the internal branch of the superior laryngeal nerve. Reproduced with permission from Gaskill JR, Gillies DR. Local anesthesia for peroral endoscopy. Using superior laryngeal nerve block with topical application. Arch Otolaryngol 1966;84:654–7. Copyrighted 1966, American Medical Association.

dle is not in an air-containing space or in the superior thyroid artery, which enters the larynx with the superior laryngeal nerve. One to 2 mL of 1% lidocaine or 0.5% bupivacaine should be infused slowly. Transcutaneous superior laryngeal nerve block is relatively contraindicated when the needle might pass through carcinoma or infected tissue.[20] Reported complications of superior laryngeal nerve block include hypotension and bradycardia.[22]

Local Anesthetic Agents *Lidocaine*, with or without epinephrine, is the most widely used topical anesthetic for awake endoscopy and can be used topically or via local infiltration (Table 65–2). *Cocaine* hydrochloride is unique among local anesthetic agents as it provides both topical anesthesia and vasoconstriction. It is rapidly absorbed by mucous membranes; high plasma levels owing to excessive and rapid absorption may lead to toxicity.[23–25] Although 1.2 g is estimated to be a fatal dose, toxicity occurs at variable doses, and severe toxicity from 20 mg has been reported.[23,24] *Tetracaine* is a relatively toxic topical anesthetic agent.[15] *Bupivacaine* is administered by injection; convulsions owing to systemic toxicity leading to cardiac arrest have been reported, presumably following unintentional intravascular injection.[25]

Precautions The lowest effective doses should be used. Appropriate doses should not be exceeded and should be carefully calculated in small patients, geriatric patients, children, and patients with cardiac or liver disease. Appropriate monitoring and resuscitation equipment and supplies and appropriately trained personnel should be available. After application of local anesthesia, patients should not eat or drink until adequate laryngeal sensation has returned.

Topical Decongestants Topical decongestants may be added to decrease nasal mucosal edema, congestion, or secretions during transnasal examinations (Table 65–3). They should be used with caution in children and patients with hyperthyroidism, cardiovascular disease, hypertension, diabetes mellitus, narrow-angle glaucoma, or benign prostatic hypertrophy.[23]

Bacterial Endocarditis Prophylaxis

Bacterial endocarditis usually develops in individuals with underlying structural cardiac defects that develop bacteremia with microorganisms likely to cause endocarditis. In these individuals, endocarditis prophylaxis is recommended for surgical operations that involve respiratory mucosa including bronchoscopy with a rigid telescope. Current criteria and regimens are available from the American Heart Association.[26]

EXAMINATION OF THE LARYNX IN THE OPERATING ROOM

Airway endoscopy entails cooperation among the endoscopist, anesthesiologist, nurses, and other support personnel to coordinate both the procedure and

TABLE 65–2. Local Anesthetic Agents

Anesthetic Agent	*Available Concentrations*	*Maximum Dose*	*Contraindications*	*Signs of Toxicity*
Lidocaine *without* epinephrine		4.5 mg/kg, not to exceed 300 mg	Hypersensitivity to lidocaine/amide-type anesthetics	Seizures, drowsiness, tremors, hypotension, bradycardia, respiratory depression
Lidocaine *with* epinephrine		7 mg/kg, not to exceed 500 mg	Hypersensitivity to lidocaine/amide-type anesthetics	Seizures, drowsiness, tremors, hypotension, bradycardia, respiratory depression
Tetracaine	0.25% or 0.5% topical solution	1 mg/kg	Hypersensitivity to tetracaine/ester-type anesthetics Should not apply to inflamed tissue	Hypo- or hypertension, respiratory arrest, hypersensitivity, nausea and vomiting
	Cetacaine spray™: 2% tetracaine hydrochloride and 14% benzocaine	Cetacaine spray: apply in adults for 1 second or less	Cetacaine spray should not be sprayed in excess of 2 s	
Cocaine	1%, 4%, 10% topical solution	1–3 mg/kg Generally 1 mg/kg sufficient	Hypersensitivity to cocaine products Systemic use contraindicated	Tachycardia, myocardial ischemia, seizures, anxiety, nervousness, restlessness, excitability, respiratory arrest

Adapted from Hunsaker DH,[15] Taketomo CK et al,[25] PDR.net for Physicians,[24] Azarnoff D et al,[18] and MICROMEDEX® Healthcare Series. Cocaine hydrochloride; lidocaine; tetracaine (DrugPoints® System); cocaine (ToxPoints® System) [serial online]. 2000;106:http://hcs.mdx.com/ (accessed Dec 8, 2000) (a password from MICROMEDEX is needed to access the site and its databases).

the relevant equipment. Advice historically attributed to Chevalier Jackson is stated as "2 minutes of preparation, 2 hours of surgery; 2 hours of preparation, 2 minutes of surgery." Establishing a routine, organized approach with other members of the team will facilitate efficient and directed patient management if a component of the procedure becomes urgent. Review the anticipated sequence of events with any new members of the team, including the anesthesiologist, nurses, and residents. Anticipate management options if an adverse event should occur. In time, the process of the procedure may flow without formal discussion. Select, assemble, and test any equipment that is likely to be used and locate any additional equipment the use of which is less likely but still possible.

The operating room should be set up to optimize safe and ergonomic relationships among the patient, anesthesiologist, endoscopist, and nurse. Both the endoscopist and the anesthesiologist should be able to access the patient's airway emergently. The cephalad portion of the operating room table should protrude over the base so that the endoscopist can sit comfortably at the head of the table. The table should include a joint at the level of the patient's neck to

TABLE 65–3. Topical Decongestants

Agent	*Available Solution Concentrations*	*Signs of Toxicity*
Oxymetazoline	0.025% 0.05%	Somnolence, sweating, central nervous system depression with hypertension, bradycardia
Phenylephrine hydrochloride	0.125% 0.16% 0.25% 0.5% 1%	Palpitations, paresthesias, vomiting, arrhythmias, hypertension. Use particular caution in children under 6 y of age

Adapted from MICROMEDEX® Healthcare Series. Oxymetazoline hydrochloride; phenylephrine hydrochloride; (DrugPoints® System) [serial online]. 2000; http://hcs.mdx.com/ (accessed Dec 8, 2000) (a password from MICROMEDEX is needed to access the site and its databases); and Drug facts and comparisons [pamphlet]. St. Louis: Wolters Kluwer Company; 2000.

allow adjustments in flexion or extension as endoscopy proceeds. Ideally, instruments are passed over the right shoulder of a right-handed endoscopist. The assistant guides the distal tip of the instrument to the proximal opening of the laryngoscope.

The patient is placed in a supine position, with the head reaching the proximal edge of the operating room table. The "sniffer" or Boyce-Jackson position provides the best visualization of the larynx, with the neck flexed on the shoulders and the head extended on the neck (Figure 65–5).[27] A gelatin pad, folded towel, or other support can be used to extend the head. A shoulder roll is useful in small children in whom the relatively large occiput may cause excessive neck flexion. Optimal patient positioning may entail the use of an assistant.

Commercial defogging agents, soap, or warm water can be used to minimize fogging of the lens or the telescope can be touched to mucosa to warm the distal tip. Lubricant applied to the upper lip allows instruments to glide gently; otherwise, the resistance of the instrument rubbing across the lip can be misinterpreted as resistance from the airway, giving a false impression of narrowing.

Eye protection is recommended for the endoscopist and may consist of the surgeon's own refractive eyeglasses, protective goggles, or laser-compatible lenses. Endoscopy is generally performed as a "clean" procedure. To prevent nosocomial infections as well as erroneous diagnostic results, instruments should be disassembled after each use and all surfaces and channels wiped or brushed to remove mucus or other contaminants prior to disinfection or sterilization.

MEDICATIONS

Corticosteroids are usually given preoperatively or intraoperatively to minimize edema, which may be an

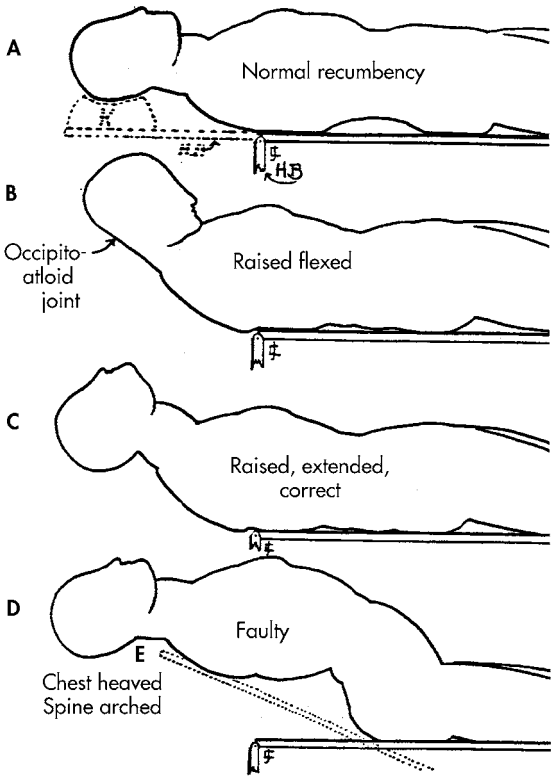


FIGURE 65–5. Schema of optimal patient position for rigid endoscopy. Reproduced with permission from Jackson C. Bronchoscopy and esophagoscopy. A manual of peroral endoscopy and laryngeal surgery. Philadelphia: WB Saunders; 1922.

TABLE 65–4. Potency of Glucocorticoid Agents Relative to Hydrocortisone[25]

Hydrocortisone	1
Prednisolone	4
Prednisone	4
Methylprednisolone	5
Dexamethasone	25–30

intrinsic component of the airway lesion or may occur following airway manipulation (Table 65–4). Edema generally occurs at the expense of the lumen, contributing to airway obstruction. The incidence of harmful effects after short-term administration of even massive doses of corticosteroids is low.[28]

Local anesthetic agents and topical decongestants may also be useful; the anesthesiologist should be made aware of any medications administered preoperatively or intraoperatively. Both topical and submucosal vasoconstricting agents, such as phenylephrine or epinephrine, can cause hypertension and acute, severe, increased left ventricular afterload. Subsequent administration of beta-blocking agents, to treat the hypertension, may further compromise left-sided cardiac output and result in acute pulmonary edema, profound hypotension, and cardiac arrest.[29]

General Anesthesia

In children for whom examination of dynamic laryngeal lesions such as vocal fold paralysis and laryngomalacia is performed in the operating room, limited sedation can be useful to achieve patient comfort, but oversedation will interfere will the endoscopist's ability to discriminate subtle findings. Similarly, evaluation of dynamic laryngeal lesions may be difficult during emergence from general anesthesia.

General anesthesia is necessary for suspension microlaryngoscopy and facilitates rigid bronchoscopy and rigid esophagoscopy. General anesthesia is managed by the anesthesiologist using a variety of inhaled and/or intravenous agents to achieve hypnosis, analgesia, amnesia, and muscle relaxation.[15] Adequate anesthesia should relax the muscle tone of the mandible and pharynx and blunt the laryngeal and bronchial reflexes.[30] Antisialagogues are often given to decrease the amount of secretions[30]; among these, glycopyrrolate produces less tachycardia than atropine.[31] Monitoring heart rate and adequate pulmonary ventilation are mandatory.[30] Adequate lighting should be available to permit assessment of the patient's color and visualization of the chest.[30] Pulse oximetry should be used prior to induction through full recovery.

Endoscopy Equipment

A variety of laryngoscopes have been developed in attempts to facilitate visualization and instrumentation for diverse lesions and anatomic constraints.[27] Most laryngoscopes used for adults are tubular, not round, so that the laryngoscope itself helps to distract tissue that might interfere with visualization. Anterior commissure laryngoscopes, such as the Holinger laryngoscope, are flared at the distal tip to improve anterior exposure of the larynx and are particularly well adapted to patients with an "anterior" larynx, large tongue, or short obese neck.[32,33] Lindholm laryngoscopes are wide enough to permit binocular vision, particularly of supraglottic lesions. Laryngoscopes such as Jako, Ossoff, and Abramson-Dedo have been specially modified for laser surgery; they provide excellent exposure and contain suction channels for the evacuation of smoke produced during laser vaporization.[32] The Jackson sliding laryngoscopes allow removal of a side or inferior panel, opening the side or bottom of the laryngoscope to facilitate instrumentation, particularly intubation. In children, laryngoscope blades are often open along the side or bottom; intubating laryngoscopes, with blades such as the Miller, can also be used. Laryngoscopes may contain side channels for suction, ventilation, or light carriers; light clips can also be attached to the proximal end of the laryngoscope. Specialized laryngoscopes are available, such as subglottiscopes or bivalved scopes used in the management of Zenker's diverticula.

Hopkins Rod Lens Telescopes

Rigid Hopkins rod lens telescopes are the workhorses of the airway evaluation. The optics and images are superior to those of comparably sized fiber-optic telescopes. In adults, standard telescopes measure 5.5 mm in outer diameter. For pediatric endoscopy, a telescope with an outer diameter of 4 mm and a length of 20 cm, such as the 7200 A telescope (Storz™), is relatively sturdy, passes easily through

the vocal folds and subglottis of a normal neonate, and provides a good quality image. It is short enough to be handled easily yet long enough to visualize beyond the carina in most children through school age. Larger-diameter telescopes are sturdier, transmit more light, and have superior optical resolution.[33] Narrower telescopes are useful for evaluating stenotic airways. For pediatric endoscopy, a variety of telescopes are necessary to fit a variety of sizes of ventilating bronchoscopes, optical forceps, etc; rigid telescopes are less flexible than fiber-optic telescopes, but this is not usually a practical limitation. The telescope tip may be defogged with a commercial solution or soap, by warming in warm water, or by touching mucosa.

Optical forceps, comprised of peanut, biopsy, or cup forceps with double-action jaws and containing a telescope for distal illumination and magnification, enable the endoscopist to visualize the airway lesion during attempts at manipulation, improving the safety and precision of the procedure (Figure 65–6). Although optical forceps can be inserted unsheathed, they are very fragile and are generally passed through a ventilating bronchoscope. They allow only a limited peripheral view, so visualization of the sidewall ventilation ports as the optical forceps is passed via the bronchoscope indicates proximity to the distal opening of the bronchoscope. The expandable tip of the instrument should be passed beyond the distal opening of the bronchoscope before the jaws of the instrument are opened.

Suspension Microlaryngoscopy

Supraglottic, glottic, and some subglottic lesions may be visualized using suspension microlaryngoscopy; it may also facilitate bronchoscopic procedures. Suspension devices free the surgeon's hands and provide stable access for visualization. Procedures using lasers, microdébriders, or microlaryngeal instruments can be accomplished using intermittent apnea techniques. Many laryngoscopes are designed or can be retrofitted for attachment to suspension equipment. Visualization, magnification, and illumination can be accomplished with a microscope using a 400 mm focal-length lens or a telescope attached to the laryngoscope. External laryngeal pressure applied to the lower portion of the thyroid cartilage will maximize visualization of the anterior commissure. Occasionally, a proximal light clip is useful.

Additional Equipment

Suctioning is often necessary to remove secretions interfering with visualization. Flexible suction catheters can be useful for routine, atraumatic suctioning. After exposing the larynx with a laryngoscope, the suction catheter, with the suction inactivated, can be insinuated between the vocal folds (as if intubating the patient with the suction catheter) and advanced to the carina. The suction is then activated, and the catheter is withdrawn expeditiously. Rigid suction catheters may be used for precise suc-

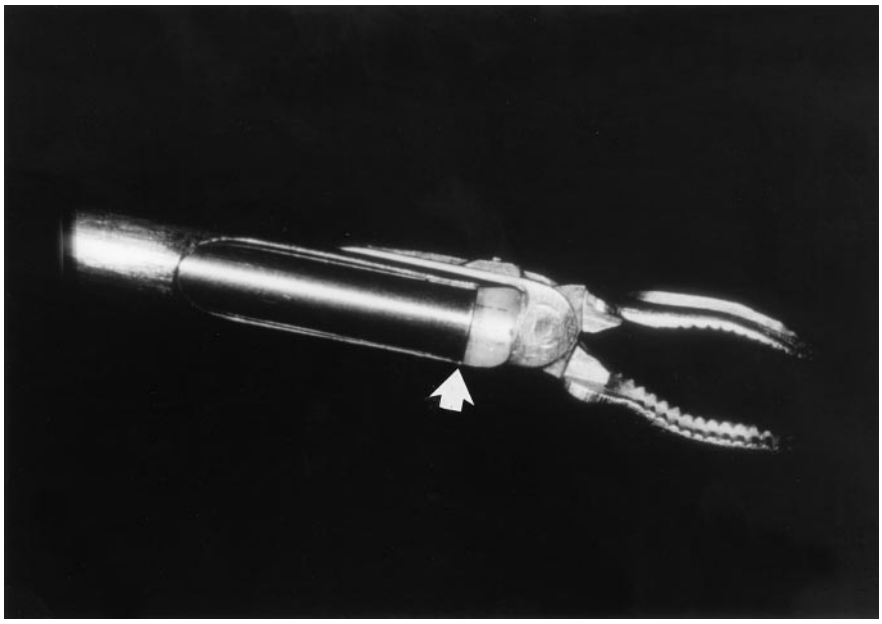


FIGURE 65–6. Distal end of "peanut" optical forceps, demonstrating proximity of telescope tip (*arrow*) to open jaws.

tioning; thumb controls are helpful. Large-bore catheters may be used to suction thick material but can cause trauma if not inserted gently. Placement is guided by feel and by listening for the tip of the catheter to contact the secretions. Small-caliber plastic suction catheters can be inserted via the side channel of ventilating bronchoscopes and allow simultaneous visualization while suctioning. It may be necessary to use a smaller-diameter telescope or larger bronchoscope to allow room for the suction catheter to pass adjacent to the telescope within the lumen of the bronchoscope.

A variety of microlaryngeal instruments are available, including straight and angled cup forceps, alligator forceps, straight and angled scissors, knives, retractors, and reflectors; some also contain aspiration tubes.[32] Older styles of laryngeal and endobronchial forceps are available in a great variety of configurations. Ring rotation forceps and rat-toothed forceps allow a grasped object to swivel in the plane of least resistance. A safety pin closer can be used to sheath the point of the pin within the keeper in situ (Figure 65–7). The lack of distal lighting and limitations of proximal magnification restrict the use of these instruments. In larger children, it is sometimes possible to pass a traditional instrument alongside a Hopkins rod telescope, but this can be very cumbersome and is not feasible in small children.

Although not widely used, contact endoscopy during microlaryngoscopy is a new tool that allows intraoperative examination of epithelial cells and the microvascular network of vocal fold mucosa that may provide in vivo and in situ diagnosis.[34] Magnification of 60 to 100 times is used in conjunction with methylene blue epithelial staining.[34]

Laryngoscopy Technique

Intubation prior to endoscopy is often avoided during evaluation for foreign-body aspiration or trauma as it may convert partial to complete airway obstruction, exacerbate trauma, or affect the endoscopic findings. However, intubation prior to endoscopy may be indicated in selected patients for safe airway management.

The introduction of the direct laryngoscope and exposure of the larynx comprises two stages: exposure and identification of the epiglottis and then elevation of the epiglottis and all of the tissues attached to the hyoid bone to expose the larynx to direct view. The laryngoscope is held in the left hand, and the spatular end is introduced in the right side of the patient's mouth, along the right side of the tongue, so the force necessary to elevate the powerful muscular tongue may be avoided by sweeping the tongue to the left. When the posterior third of the tongue is reached, the tip of the laryngoscope is directed toward the midline and the dorsum of the tongue is elevated by a lifting motion imparted to the laryngoscope. In adults,

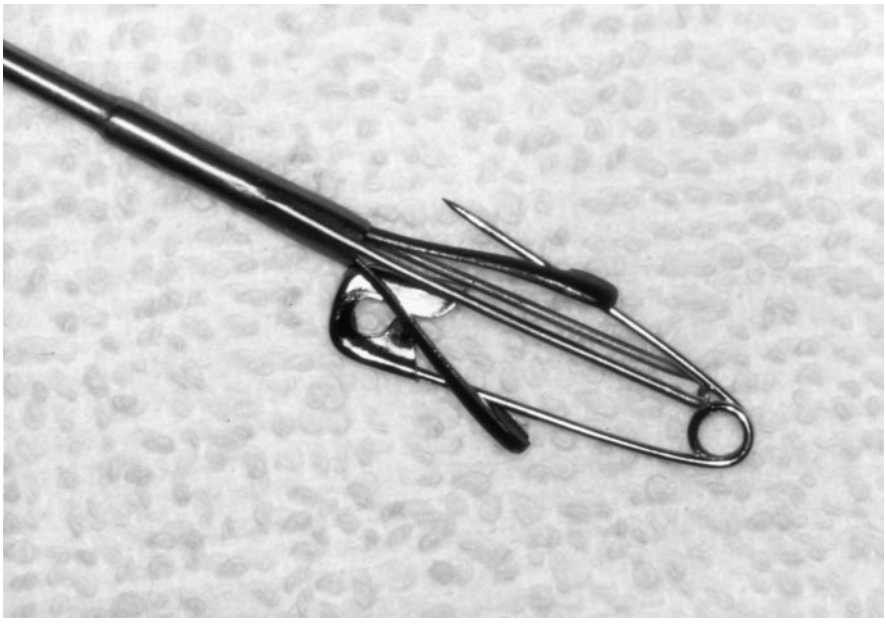


FIGURE 65–7. Safety pin closer. Note central rigid wire that abuts the ring of the spring, while two lateral components of the forceps engage the pointed and keeper branches of the safety pin. Withdrawing the lateral forcep components while maintaining the position of the ring will adduct the branches, closing the pin.

the spatular end of the laryngoscope is tipped back under the epiglottis and advanced about 1 cm. The larynx is now exposed by a motion that is best described as a suspension of the head and all of the structures attached to the hyoid bone on the tip of the spatular end of the laryngoscope. Particular care must be taken at this stage not to use the upper teeth as a fulcrum but rather to impart a lifting motion with the tip of the speculum.[1] The tip of the laryngoscope is lifted forward and upward ("toward the corner of the room, where the wall in front of you meets the ceiling"), exposing the vocal folds. If the anterior commissure of the larynx is not readily seen, the lifting motion and elevation of the head should be increased, or external pressure can be applied, depressing the thyroid cartilage.[1] Generally, the vocal folds are well seen before proceeding to insert a telescope or bronchoscope. A 0-degree telescope provides an excellent view of laryngeal structures; occasionally, an angled telescope may be used to visualize the ventricles, anterior commissure, or undersurface of the vocal folds.

In the pediatric patient, elevation of the occiput is not always necessary because the child has a proportionately large head and consequently already has the elevation of a "pillow" when the head, neck, and shoulders are in a neutral position on a flat surface.[35] The larynx is exposed using a rigid laryngoscope or an intubation laryngoscope with a straight blade, such as a Miller. The tip of the blade is passed along the right side of the tongue, sweeping the tongue to the left, and, instead of lifting the epiglottis, the tip of the blade is insinuated into the vallecula and then elevated. Elevating the base of the tongue and vallecula anteriorly provides an excellent view of the larynx that is easy to maintain for sustained periods of time.[35]

Despite appropriate technique and adequate anesthesia, rigid endoscopy may be difficult in patients with abnormal anatomy, such as morbid obesity, trismus or temporomandibular joint ankylosis, or inability to adequately flex or extend the neck, and fiber-optic endoscopy may be necessary.

Complications

Airway structures should be handled in a gentle, atraumatic manner to minimize the possibility of postoperative edema or mucosal abrasion, potentially progressing to scarring and subsequent stenosis. Edema may already be present or may be caused or exacerbated by manipulation of an inflamed or narrowed portion of the airway. In infants and small children, the subglottis is particularly at risk for injury and complications as it is the narrowest portion of the airway and is nondistensible. Whereas most of the larynx and trachea contains a membranous posterior component, the cricoid ring within the subglottis is a complete ring, so it cannot expand to accommodate the passage of instruments.

Bleeding, although generally minor, can interfere with visualization as well as ventilation and may be controlled by irrigation, application of neuropatties saturated with epinephrine, or tamponade.

Ventilation

Otolaryngologists and anesthesiologists have a unique relationship as they share responsibility for the airway during aerodigestive endoscopic procedures performed using general anesthesia. Selecting a safe and effective method of ventilation and oxygenation, while allowing adequate surgical exposure and access, requires planning and cooperation.

Standard endotracheal tubes, passed nasally or orally, provide good control of the airway and a closed ventilation system and are the most familiar to the anesthesiologist. They allow delivery of anesthetic gases and oxygen as well as removal of carbon dioxide. However, the tube will interfere with the surgeon's ability to see or work in the posterior part of the larynx[15] or subglottis. In addition, standard polyvinyl chloride endotracheal tubes are readily ignited by laser, producing an intense flame, scattering carbonaceous material through the endobronchial tree and releasing hydrogen chloride gas, which is a potent pulmonary toxin.[32]

With the increasing use of lasers in surgery, attempts have been made to develop tubes resistant to laser ignition. Silicone or red rubber tubes are fairly resistant to laser penetration; this resistance is increased if they are completely wrapped in an adhesive aluminum foil tape. Some tubes, including all-metal tubes, are "laser safe" but are fitted with separate rubber cuffs, which are not laser resistant; instilling saline into the cuff adds heat-sink protection. With any type of endotracheal tube, protection of the cuff and subglottic space is provided by saline-soaked cottonoids.[32]

Low-frequency subglottic jet ventilation can be accomplished using a narrow, self-centering tube designed to minimize interference with visualization of laryngeal structures. Potential complications include carbon dioxide retention,[36] pneumothorax, and pneumomediastinum, particularly if vocal fold adduction or mass lesions preclude adequate expiration of gases. The Hunsaker Mon-Jet subglottic tube (Tuta Labs, Sydney, Australia) is made of combustion-resistant fluoroplastic and allows periodic measurement of end-tidal carbon dioxide.[36] To prevent pneumothorax, a port allows continuous monitoring of end-expiratory and peak airway pressures, which will automatically stop insufflation if these pressures rise above a preset level.

Supraglottic jet ventilation is accomplished using a fireproof metal tube placed above the glottis. Adequate tidal volumes are accomplished by entrainment of room air ahead of the jetted oxygen, air, or anesthetic gas.[15] Precise positioning of tube and alignment along the central axis of the airway are critical.[15] Blood and debris can be blown into the trachea.[15] If the tube is misaligned or dislodged, submucosal emphysema, tension pneumothorax, or inadequate ventilation may occur.[15]

The *laryngeal mask airway* is an inflatable mask that is blindly inserted into the hypopharynx where it forms a low pressure, end-to-end seal with the glottis; a tube connects the mask to the breathing circuit.[37] A laryngeal mask airway may be useful in circumstances when cervical spine range of motion is limited. Laryngeal mask airways have been used to provide a conduit for blind intubation of a difficult airway or to facilitate passage of a fiber-optic bronchoscope.[37] The disadvantages include lack of protection against aspiration and the potential for airway obstruction owing to malposition, abnormal position of the epiglottis, or laryngospasm during induction or emergence. The laryngeal mask airway is contraindicated in patients at increased risk for regurgitation and aspiration. Contraindications relevant to otolaryngologists include pharyngeal pathology such as abscess, hematoma, tumor, or edema; airway obstruction at or below the larynx; inability to extend the neck or open the mouth more than 1.5 cm; and high airway resistance.[37]

Using intermittent apnea, the surgeon can work in the airway without the presence of tubes obstructing surgical access or distorting laryngeal anatomy. The risk of ignition by laser pulses is eliminated. Airway control and monitoring capabilities are limited,[15] and surgical procedures may be interrupted and rushed.[15] A facemask, endotracheal tube,[15] or jet ventilation can be used to provide oxygenation between episodes of apnea.

Spontaneous respiration requires a carefully balanced gas mixture to anesthetize the patient without significantly suppressing spontaneous respiration. Topical anesthesia may be necessary to suppress the laryngeal reflexes, that is, bradycardia, hypertension, and arrhythmias. This open system exposes operating room personnel to the anesthetic gases released into the room. Arrhythmias and inadequate ventilation can occur, depending on the anesthesiologist's success in maintaining adequate general anesthesia while allowing sufficient spontaneous respiration.[15]

Difficult Intubation

The otolaryngologist may be called on to secure the airway in emergent or nonemergent situations when anatomic abnormalities make intubation difficult. In these circumstances, the otolaryngologist's knowledge of abnormal anatomy and variations of normal anatomy is invaluable. If an elective intubation is expected to be difficult, appropriate equipment for endoscopic and surgical control of the airway should be available. Patients with morbid obesity, trismus or temporomandibular ankylosis, inability to flex or extend the neck adequately, suspicion of cervical vertebrae instability, possible traumatic laryngotracheal separation, or congenital or acquired oral, pharyngeal, laryngeal, or tracheal abnormalities may be difficult or even impossible to intubate. Patients who have undergone surgical laryngotracheal separation or diversion or laryngectomy cannot be intubated perorally or nasally; a tracheal stoma will be present but may be stenotic.

Several options are available to access the airway when standard methods are unsuccessful. After visualizing the larynx with a sliding laryngoscope, the side or bottom of the laryngoscope is removed so that an endotracheal tube can be passed adjacent to the laryngoscope lumen. Alternately, an endotracheal tube can be passed directly through the lumen of a Dedo-style laryngoscope. A stylet may be used to stiffen the endotracheal tube sufficiently to allow passage through a partially obstructing lesion. Intubation can also be accomplished during spontaneous ventilation using a fiber-optic bronchosope. An endotracheal tube, with the adaptor removed, is

placed proximally over the bronchoscope shaft, which is passed through the nose or mouth and advanced through the glottis to the carina. Application of lidocaine may be indicated. The endotracheal tube is then advanced over the bronchoscope; a rotating motion is used as the tube traverses the nose and larynx.[8] A 4.0 mm fiber-optic bronchoscope (standard adult size) can be used in endotracheal tubes larger than 6 mm, a 3.5 mm bronchoscope should be used in tubes from 4.5 to 6 mm, and the smallest of the ultrathin instruments with distal angulation (2.2 mm, Olympus PF22) will pass through 2.5 to 4 mm endotracheal tubes.[8]

If intubation attempts are unsuccessful in emergent situations, cricothyroidotomy or tracheostomy may be necessary. Mask ventilation can sometimes be used to temporize. If a tracheostomy stoma is present, cannulation may be accomplished with an endotracheal tube or a tracheostomy tube advanced with an obturator or over a suction catheter; serial dilatation may be required.

Patients who struggle to breathe spontaneously against a partially or totally obstructed airway may develop postobstructive (noncardiogenic) pulmonary edema.[38] Although the most common cause is postextubation laryngospasm,[38] postobstructive pulmonary edema can occur following acute airway obstruction of any cause and is estimated to occur in more than 10% of patients requiring active airway intervention for acute upper airway obstruction.[38] If reintubation is necessary, pink, frothy sputum may be appreciated; mechanical ventilation with positive end-expiratory pressures (PEEPs) may be indicated.[38]

Emergence and Postoperative Care

Emergence from general anesthesia can be accomplished using mask ventilation or intubation; the decision will depend on both the nature of the airway lesion and the preference of the anesthesiologist; however, minimizing airway manipulation is preferred when feasible. The level of care required after airway endoscopy is related to the severity of the airway lesion and the type of anesthesia used. Additional systemic corticosteroids and racemic epinephrine, with or without nebulized dexamethasone, may be indicated. Postoperative discomfort from the endoscopy itself is usually minor.

BRONCHOSCOPY

Bronchoscopy can be diagnostic or therapeutic, and interval examinations may be indicated to follow disease progression or response to treatment; additional indications are discussed in Chapter 67. During bronchoscopy, "all bronchial orifices must be identified seriatim; because this is the only way by which the bronchoscopist can know what part of the tree he is examining. Appearances alone are not enough. It is the order in which they are exposed that enables the inexperienced operator to know the orifices."[1]

Bronchoscopy can be accomplished using rigid or fiber-optic equipment. "Clearly rigid and flexible bronchoscopes neither duplicate or replace each other, rather they are complementary and should be used when appropriate for the particular problem."[39] Patient monitoring should be appropriate relative to the endoscopy technique, the lesion, and the patient's underlying medical status.

Rigid Bronchoscopy

Rigid ventilating bronchoscopes should be available during rigid airway endoscopy. They allow control of the airway, capability for ventilation and oxygenation, and a stable, protective conduit for manipulation and instrumentation. They are the instruments of choice for controlling massive pulmonary hemorrhage, removing tracheobronchial foreign bodies, dilating stenoses, excising granulation tissue, debulking large tumors, evaluating tracheoesophageal fistulae or laryngeal clefts, and performing laser therapy or inserting an airway stent[4,8,39] because of their rigidity and large caliber.

When Hopkins rod telescopes are used with rigid bronchoscopes, they provide better optical quality than comparably sized fiber-optic telescopes. The working channel of a rigid bronchoscope is larger than those of fiber-optic bronchoscopes, allowing passage of a greater variety and larger caliber of instruments, including large-bore suction catheters, various grasping or biopsy forceps, and devices to install stents. Larger biopsies and tumor debulking can be accomplished more efficiently. They can be used to shield, protect, and guide fragile instruments such as thin telescopes and optical forceps. The tactile sensation of fixation of structures is readily apparent with the rigid open-tube bronchoscope and seldom appreciated with a fiber-

optic instrument.[40] Rigid bronchoscopes are compatible with both carbon dioxide and flexible laser applications and are nonflammable. Pediatric bronchoscopes contain side-channel ports allowing passage of narrow-bore suction catheters, Dormia urologic stone baskets, or side-channel alligator or cup forceps while maintaining ventilation, visualization, and illumination during procedures such as biopsies or retrieval of foreign bodies.

The outer diameters of pediatric bronchoscope range from 4.2 to 8.2 mm and of adult bronchoscopes from 9.2 to 14 mm.[7] The number designated as the size of the bronchoscope does not necessarily directly correlate with the inner or outer diameter measurement of that instrument.

Although early rigid bronchoscopy was performed in conscious, awake patients, modern rigid bronchoscopy is usually performed using general anesthesia. Direct laryngoscopy is performed first to avoid missing a laryngeal or hypopharyngeal lesion. Rigid bronchoscopy can be performed using an open-tube bronchoscope containing a Hopkins rod telescope to provide magnification and distal illumination or a side-channel light carrier and a proximal magnification lens. Alternatively, an unsheathed Hopkins rod telescope can be used, providing excellent magnification and distal illumination. An unsheathed telescope is more fragile and does not provide access for ventilation. However, this optimizes the size of the telescope that can be inserted and, therefore, image quality, while limiting the outer diameter of instruments that may cause trauma in small children or narrowed airways.

The patient's head and neck are initially placed in a neutral position. A tooth guard is used in adults to prevent injury to the maxillary teeth. The bronchoscope is oriented so that the bevel is oriented vertically and the ventilation port is accessible to the anesthesiologist. The bronchoscope can be introduced directly into the oral cavity, or it can be passed through a Jackson slide laryngoscope. The bronchoscope is inserted into the oral cavity from the right side of the mouth; the endoscopist uses his or her left hand to hold the bronchoscope and his or her right hand to steady and maneuver the maxilla and retract the lips. The endoscopist's right hand will eventually be used for instrumentation while his or her left hand maintains stability and controls advancement of the bronchoscope. The bronchoscope is used to sweep the tongue to the left and is advanced until the tip of the epiglottis is visualized. At this point, the patient's head is extended relative to his neck. The tip of the bronchoscope is advanced in the midline until it passes through the laryngeal inlet and the vocal folds are visualized. The bronchoscope is positioned with the long end of the bevel parallel to the longitudinal axis of the vocal folds so that the long end of the bevel enters the glottis in the midline. The bronchoscope should pass gently through abducted vocal folds. Once beyond the glottis, the bronchoscope can be rotated in any desired direction, and the camera is rotated in a compensatory fashion to maintain an upright view on the monitor.

An alternate method, particularly useful in infants and children, is to first expose the vocal folds using a sliding laryngoscope or an intubating laryngoscope blade with an open back, such as a Miller. In adults, the laryngoscope is inserted under the epiglottis, and in children into the vallecula. Once the bronchoscope is passed through the vocal folds, the laryngoscope is gently removed. At this point,

> the bronchoscope is to be held by the left hand like a billiard cue, the terminal phalanges of the left middle and ring fingers hooking over the upper teeth, while the thumb and index finger hold the bronchoscope, clamping it to the teeth tightly or loosely as required. Thus the tube may be anchored in any position, or at any depth, and the right hand which was directing the tube may be used for the manipulation of instruments.[1]

Once the bronchoscope is within the trachea, the ventilation port can be used. As the bronchoscope is further inserted, the head must be positioned so that the tube corresponds to the axis of the lumen of the passage to be examined.[1] Examination of the trachea and bronchi should be performed in a routine sequence; the potential for bleeding, airway obstruction, or multiple lesions may affect the sequence in an individual patient. Mucosal characteristics are continuously examined as the bronchoscope is advanced and withdrawn. Mucosal abnormalities can be subtle because of small size, minimal surface effects, or obscure location. Slight surface irregularity or friability may be the only visual clue to a neoplasm. Often the best site for biopsy can be determined with telescopic magnification, but in vivo tissue staining has yielded diagnostic information in special circumstances. The

trachea and bronchi are also inspected for patency, deviation, configuration and intrinsic lesions, intraluminal lesions, or evidence of external compression. Subcarinal lymphadenopathy or other lesions may cause the carina or subcarinas to be blunted, broadened, displaced, or abnormally tethered during respiration and cardiac contractions.

As the bronchoscope approaches the carina, the patient's head is turned to the left to allow the bronchoscope to enter the right main bronchus. To enter the bronchi, the bevel of the bronchoscope should be turned in the direction of the bronchus to be explored.[1] The first orifice within the right main bronchus, on the lateral wall, is the opening to the right upper lobe bronchus (Figure 65–8). The bronchoscope can be stabilized just proximal to this subcarina and a 90-degree telescope passed through the bronchoscope, with its lens pointed laterally, to visualize the apical, posterior, and anterior segmental bronchi. As the bronchoscope is advanced through the bronchus intermedius, the head requires some left lateral deflection and a considerable degree of lowering[1] as the right middle lobe bronchus is identified anteriorly with its lateral and medial segments. The inferior lobe bronchus, located just posterior to the middle lobe bronchus, contains the superior, anterior basal, medial basal, lateral basal, and posterior basal segmental bronchi. A 0-degree Hopkins rod telescope may be passed through the bronchoscope to view the middle and lower lobe segments.

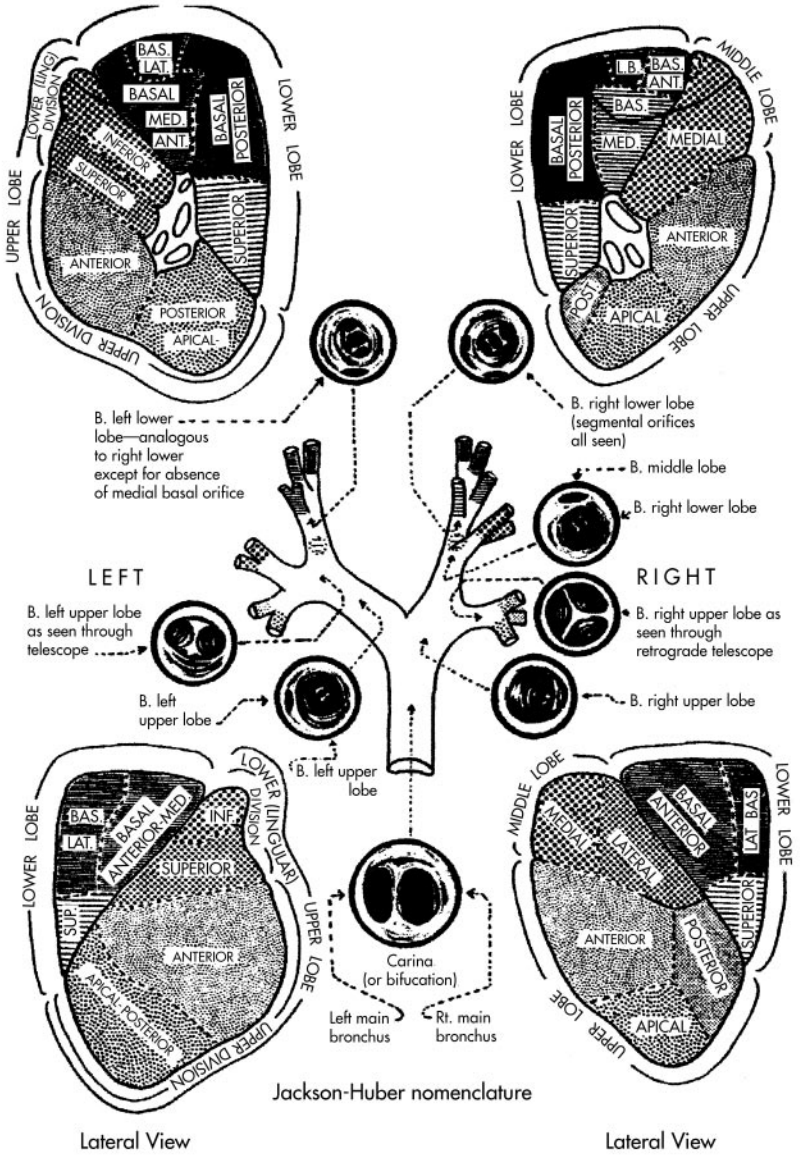


FIGURE 65–8. Schema showing the tracheobronchial tree, bronchopulmonary segments, and endoscopic landmarks.

The bronchoscope is then withdrawn just proximal to the carina, the patient's head is rotated to the right, and the bronchoscope is advanced into the left main bronchus. The left upper lobe bronchus, on the lateral wall, contains upper and lingular divisions, which can be visualized with a 90- or 120-degree telescope. The upper bronchus contains the apical, posterior, and anterior segmental bronchi; the lingular bronchus contains the superior and inferior segmental bronchi. Continuing distally, the lower lobe bronchus contains the superior segmental bronchus along its posterior wall and further distally the anterior, medial basal, lateral basal, and posterior basal segmental bronchi. As the bronchoscope is being withdrawn at the completion of bronchoscopy, careful retrograde examination should be performed.[41,42] During pediatric bronchoscopy, the telescope is connected to the ventilating bronchoscope by a locking bridge; the flexibility of the patient's neck and airway facilitate thorough examination using a 0-degree lens.

Mass lesions, including papillomas and granulation tissue, may be prudently biopsied or completely removed, particularly if obstructing and pedunculated. Removal of large sessile lesions, particularly if present in a circumferential distribution, may contribute to subsequent stenosis.

Limitations of rigid bronchoscopy include the requirement of adequate neck mobility and stability and the ability to open the patient's mouth; damage to teeth can occur.

Fiber-optic Bronchoscopy

Fiber-optic bronchoscopy is useful for examination of distal airway lesions, as a wedge for broncheoalveolar lavage, or as a conduit for fiber-optic lasers. Fiber-optic bronchoscopes can generally be inserted further toward the periphery of the lungs than rigid telescopes, particularly in the upper lobes. Fiber-optic endoscopy may be preferable in patients with cervical spine instability or limited mobility or trismus, for bedside examination, or when attempting to avoid the use of general anesthesia. Fiber-optic bronchoscopy can be performed in patients using topical or general anesthesia; however, it is more difficult to pass the bronchoscope through and beyond the glottis unless the patient is spontaneously ventilating. It may be performed using a conduit, such as an endotracheal or tracheostomy tube.

Fiber-optic bronchoscopy during spontaneous ventilation, with or without general anesthesia, is useful in evaluating dynamic lesions such as tracheomalacia or bronchomalacia.[8] Fiber-optic bronchoscopy may reveal significant airway collapse not evident during rigid bronchoscopy during general anesthesia; however, care must be taken to avoid overdiagnosis of airway dynamics. The presence of fiber-optic bronchoscope may increase airway resistance to the point that normal dynamics are exaggerated, particularly in infants whose airway structures are more compliant than those of adults.[8] Use of simultaneous fluoroscopic guidance[43] may facilitate intralesional placement of brushes, biopsy forceps, or catheters.

Equipment The outer diameter of adult fiber-optic bronchoscopes ranges from 3.4 to 6.2 mm, with side-channel diameters ranging from 1.2 to 3.2 mm; standard pediatric fiber-optic scopes have a 3.5 mm outer diameter and an operating channel of 1.2 mm.[39] Side channels can be used for suctioning, application of topical anesthetics or liquid radiographic contrast material, or passage of instruments such as biopsy forceps, balloon catheters, laser fibers, and wire brushes.[4] Fiber-optic bronchoscopes with tip angulation capability but without side channels are available with outer diameters as small as 2.2 mm.

Technique A sufficiently alert and cooperative patient can undergo fiber-optic bronchoscopy in a sitting position; the examination can also be performed supine. If the procedure is performed using local anesthesia, sedation and premedication with atropine may be used to reduce secretions and to prevent the vagovagal reflex. Anesthetizing the supraglottis and glottis prior to awake bronchoscopy, as described in the section on laryngoscopy, is helpful in preventing coughing and laryngospasm. Topical anesthesia of the trachea and subglottis, supplied by the recurrent laryngeal nerve,[17,20] can be accomplished by instilling anesthetic agents via the bronchoscope's instrument channel or by spraying the vocal folds and trachea at the beginning of each inspiration using a long, curved nozzle inserted perorally. Transtracheal injection of lidocaine can be performed via cricothyroid membrane puncture at the end of maximal expiration; subsequent coughing will spread anesthesia droplets from the carina to the inferior surface of the vocal folds.[17] Supplemental oxygen through the free

naris may be indicated in infants less than 5 kg or patients with diffuse lung disease or abnormal blood gas measurements.[39]

The bronchoscope's control unit and proximal lens are held with the endoscopist's left hand, leaving the right hand free to manipulate the bronchoscope or insert instruments. The endoscopist's thumb operates the angulation lever, which controls the angle of the distal tip, while the index finger is used to occlude the suction port as needed. Passing the bronchoscope through the nose is generally well tolerated and provides a stable, direct route to the vocal folds. A bite guard should be used to protect the bronchoscope if it is passed perorally. If the bronchoscope is inserted via the pharynx, it is advanced through the true vocal folds during inspiration; bronchoscopes may also be inserted through a facemask, laryngeal mask airway, or an endotracheal or tracheostomy tube. For patients receiving mechanical ventilation, various adapters can be used to minimize air leak and maintain mechanical ventilation and oxygen delivery.[43] When feasible in adults, the endotracheal tube size, based on the inner diameter, should be at least 1.5 mm larger than the external diameter of the fiber-optic bronchoscope.[44]

Pharyngeal hypotonia, from muscle or central nervous system disease, may make exposure of the larynx difficult. Techniques that may improve visualization include occluding the tracheostomy tube, if present; insufflating oxygen via the bronchoscope side cannel or a suction catheter or nasopharyngeal airway in the contralateral nasal cavity; or using a rigid laryngoscope.[8]

To obtain a biopsy, grasp the specimen with a cup forceps, then abruptly twist and withdraw the forceps. Aspirates or biopsy specimens obtained for culture should be placed in nonbacteriostatic saline; carinal biopsy specimens obtained in evaluation of ciliary motility abnormalities are placed in glutaraldehyde. If biopsy is performed for diffuse pulmonary disease, the specimen is obtained from the most gravity-dependent portion of lung to facilitate bleeding control.[16] After the biopsy is obtained, the fiber-optic bronchoscope is "wedged" to tamponade the bleeding for 3 to 5 minutes and then withdrawn without suctioning to avoid disturbing the clot.[16] In immunocompromised, thrombocytopenic patients undergoing evaluation for infections, broncheoalveolar lavage should be considered in lieu of biopsy.[16]

Although small forceps and baskets that will pass through the pediatric (or adult) flexible scope are available, foreign-body extraction is sufficiently difficult and risky that it should rarely, if ever, be attempted with a flexible bronchoscope unless attempts with a rigid instrument have failed and the flexible instrument offers some distinct advantage in the specific situation.[8]

Fiber-optic bronchoscopy is relatively contraindicated in patients who are hemodynamically unstable or who have unstable arrhythmias, increased intracranial pressure, severe hypoxemia, hypercapnia, systemic coagulopathy, or thrombocytopenia.[16]

Meduri and Chastre have defined factors affecting the risk of complications in mechanically ventilated adults undergoing fiber-optic bronchoscopy.[44] Criteria for high risk include active bronchospasm, PaO_2 of 70 mm Hg despite $FIO_2 > 70\%$, PEEPs of ≥ 15 cm H_2O, recent acute myocardial infarction (≤ 48 hours), unstable arrhythmia, mean arterial pressure < 65 mm Hg on vasopressor therapy, or platelet count < 20,000/mm^3.[44] Criteria for relatively high risk include PEEP > 10 cm H_2O, significant auto-PEEP (≥ 15 cm H_2O), prothrombin time or partial thromboplastin time greater than 1.5 times control, or increased intracranial pressure.[44]

Fiber-optic bronchoscopy can cause airway obstruction as the portion of the airway lumen occupied by the scope is not available for ventilation. The degree of obstruction is related to the outer diameter of the bronchoscope relative to the size of the narrowest portion of the airway; a 3.5 mm fiber-optic bronchoscope will nearly totally obstruct the airway of an infant less than 3,000 g in weight.[39] Other potential complications of rigid or fiber-optic bronchoscopy include drop in PaO_2, vagovagal reaction, fever, cardiac arrhythmias, bronchospasm, pneumonia, transient worsening in pulmonary infiltrates, and pneumothorax.[16,44] Suctioning and bronchial lavage may decrease effective tidal volume and decrease arterial oxygenation[43]; patients receiving mechanical ventilation may require compensatory changes in ventilator settings.[44]

Bleeding may occur at any site of manipulation, including the nose[45]; most resolve spontaneously.[16] The risk of bleeding is increased in patients with chest malignancies, pneumothorax,[44] immunosuppression, or coagulopathies.[16] Profuse bleeding, consisting of more than 100 mL of blood and lavage fluid in adults, can occur because of

trauma or underlying disease or following transbronchial, endobronchial, or brush biopsy.[16] Techniques helpful in controlling hemorrhage include intrabronchial instillation of 2 mL aliquots of 1:10,000 epinephrine; lateral decubitus positioning of the bleeding hemithorax; tamponade using the bronchoscope, Fogarty catheter, or a specially designed 200 cm long balloon catheter; bronchial artery embolization; open surgical procedures; or coagulation of endoscopically visible bleeding sites with a neodymium:yttrium-aluminum-garnet laser.[16] Compensatory strategies include selective intubation of the nonbleeding lung.[16] Fatal hemorrhage is uncommon.[16]

Limitations of Bronchoscopy

During rigid or fiber-optic bronchoscopy, care must be taken to minimize edema in distal airways as this may interfere with ventilation and oxygenation. Before manipulating a distal lesion or structure, it may be useful to review with the anesthesiologist how ventilation will be accomplished if the lesion is displaced or if edema occurs. Sometimes smaller telescopes can be passed beyond a lesion to view distal structures. Consider the risks of proceeding and the risk and availability of other avenues of investigation, such as medical imaging, to accomplish the intended objective.

ESOPHAGOSCOPY

Esophagoscopy provides details about esophageal anatomy and mucosal characteristics not obtainable by other methods. Specific indications for esophagoscopy are discussed in Chapter 67. As with bronchoscopes, rigid and fiber-optic esophagoscopes are complementary rather than rival instruments. Flexible instruments can be used in conscious patients, permitting assessment of function, as well as in patients with kyphoscoliosis, trismus, or cervical spine instability or rigidity, and for retrograde evaluation of the gastroesophageal junction. Endoscopy of the esophagus, stomach, and duodenum can be accomplished with a single instrument, and the use of general anesthesia may be avoided. Rigid instruments are superior for evaluation, biopsy, or management of the pharyngoesophageal segment and for particular lesions such as diverticula.[5]

Rigid Esophagoscopy

During rigid esophagoscopy, telescopes, light carriers, metal suction catheters, optical forceps, or other grasping instruments can be placed in the lumen of an open-tube esophagoscope. Pediatric esophagoscopes also contain a side channel oriented such that stiff plastic suction catheters, Dormia urologic stone baskets,[46] or narrow, flexible alligator or biopsy forceps can be inserted into the lumen alongside a Hopkins rod telescope, allowing simultaneous magnification, visualization, and distal illumination. Assembling the telescope within a pediatric esophagoscope with a locking bridge will result in the tip of the telescope protruding beyond the opening of the esophagoscope, limiting visualization. Alternatively, the telescope can be secured within the esophagoscope by inserting it through a cap with a rubber dam. This is less stable than a locking bridge but allows the tip of the esophagoscope to distend the esophageal lumen in advance of the telescope, improving visualization. Using this technique, the grip in the right hand must remain relaxed; clenching the fingers will cause the tip of the telescope to protrude beyond the tip of the esophagoscope. In a normal esophagus, straight and rigid tubes of 7 mm diameter should pass freely in infants and tubes of 10 mm in adults.[1]

The operating room table should have a joint under the patient's neck so the patient's neck and head position can be adjusted as the procedure progresses. The end of the table should protrude so that the endoscopist can sit comfortably at the head of the table as needed during the procedure.

Technique An experienced endoscopist may be able to retrieve a proximal esophageal foreign body during episodes of apnea, but generally patients undergoing rigid esophagoscopy are intubated, allowing control of the airway and minimizing time constraints. Following orotracheal intubation, taping the endotracheal tube to the left oral commissure facilitates passage of endoscopic instruments without inadvertent extubation.

The esophagus is a tubular structure that begins at the cricopharyngeus muscle, located about 20 cm from the teeth, and extends to the gastroesophageal junction at the diaphragmatic hiatus, about 40 cm from the teeth in adults. The cricopharyngeus muscle forms the upper esophageal sphincter, and thickening of the smooth muscle distally at the gastroesophageal junction creates the lower

esophageal sphincter. Normally, the lumen of the esophagus is collapsed in a flattened or stellate pattern, but it can be distended to at least 2 cm. The longitudinal mucosal folds flatten with distention. The wall of the esophagus is thin and lined by a shiny, pale pink mucosa.[1,41]

The diameter of the esophagus is reduced at four points: the cricopharyngeus, the crossing of the aorta at 25 to 30 cm from the incisors, the crossing of the left bronchus, and the hiatus for the esophagus at the diaphragm. The average distance to these points from the upper teeth, according to age, is shown in Figure 65–9. There is a definite fifth narrowing of the esophageal lumen not easily demonstrated endoscopically and not seen during dissection but readily shown functionally by the fact that almost all foreign bodies lodge at this point. This narrowing occurs just below the cricopharyngeus, at the superior aperture of the thorax, and is probably produced by the crowding of the numerous organs that enter or leave the thorax through this orifice.[1] In the spontaneously breathing patient, periodic relaxation and contraction of the cricopharyngeus muscle can be viewed via an esophagoscope placed posterior to the cricoid cartilage.

In the "high-low" technique described by Jackson, the position of both the endoscopist and the patient's head and neck relative to his body change as the endoscopy proceeds.[1] Jackson divides the procedure into four stages: entering the right piriform sinus, passing the cricopharyngeus, passing through the thoracic esophagus, and passing through the hiatus. Initially, to bring the cervical spine, and therefore the esophagus, into a straight line with the upper portion of the dorsal spine, the head is elevated into the "high" position.[1] The endoscopist stands cephalad to the patient. Jackson asserted that the services of a trained assistant to place the head in the proper sequential "high-low" positions are indispensable[1]; head position can also be maintained by flexing the neck joint of the operating room table or placing a firm support under the head.

Initially, the esophagoscope is nearly vertical, with the long end of the bevel positioned anteriorly. As the esophagoscope is advanced along the right side of the tongue and down the posterior pharyngeal wall, the long bevel is used to lift the tongue base, endotracheal tube, and arytenoid cartilages in advance of the esophagoscope, and the position of the esophagoscope will approach horizontal. The

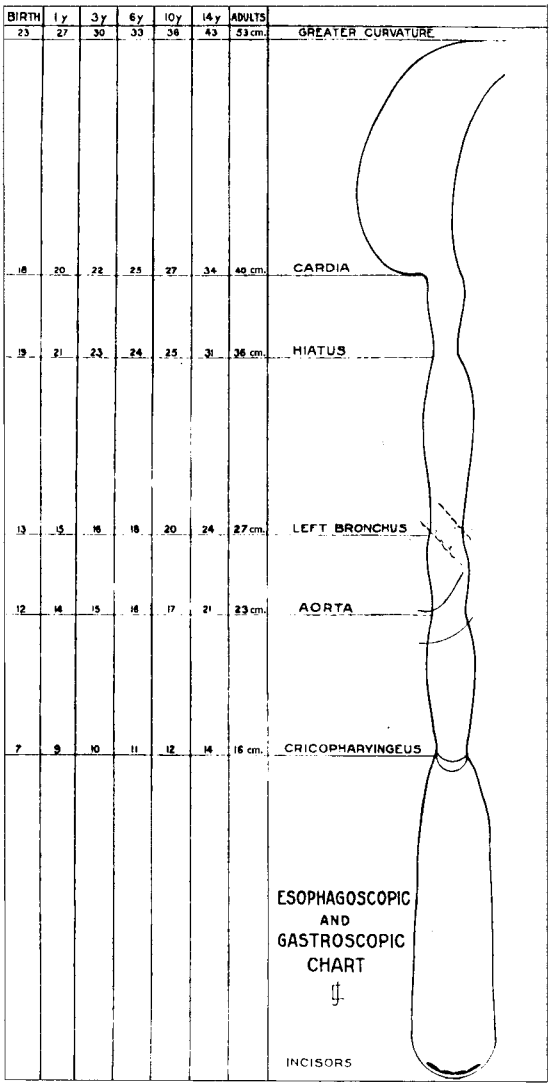


FIGURE 65–9. Jackson's chart of approximate distances of esophageal narrowings from the upper incisors at various ages. Reproduced with permission from Jackson C. Bronchoscopy and esophagoscopy. A manual of peroral endoscopy and laryngeal surgery. Philadelphia: WB Saunders; 1922.

esophagoscope follows the natural food passage; just as the epiglottis splits a bolus of food and the two portions drift laterally into the piriform sinuses, the esophagoscope is introduced via the right piriform sinus. "To insert the esophagoscope in the midline, posterior to the arytenoids, required a degree of force dangerous to exert and almost certain to produce damage to the cricoarytenoid joint or to the pharyngeal wall, or to both."[1] A lifting motion imparted to the tip of the esophagoscope by the left thumb will expose the rounded right arytenoid eminence. The tip

of the esophagoscope should now be directed somewhat toward the midline, following the funnel shape of the hypopharynx. The endoscopist's right hand remains at the proximal end of the esophagoscope to control the angle of insertion and progression of the esophagoscope, as well as the position of the telescope within the esophagoscope, and to manipulate side-channel suctions or other equipment. The endoscopist's left hand is positioned at the alveolar ridge or maxillary teeth. The fingers of the left hand advance the esophagoscope and maintain an awareness of a space between the esophagoscope and the patient's teeth, and the left thumb lifts the tip of the esophagoscope. The cricopharyngeus is entered without force.[1] "To prevent perforation the esophagoscope is to be passed only with ocular guidance."[1] Once the esophagoscope is introduced through the esophageal introitus and cricopharyngeus, the endoscopist sits and the head of the operating table is lowered below horizontal. At this point, the esophagoscope can be rotated so that the side-channel port is on the right, facilitating passage of suction catheters or instruments with the endoscopist's right hand. While instruments are manipulated with the right hand, the left hand maintains the appropriate position of the esophagoscope. The esophagoscope is gently advanced with the left hand, following the lumen.

If the patient's position is correct, the esophagoscope should glide easily through the thoracic esophagus. As the esophagoscope is advanced, the impression of the arch of the aorta and the left bronchus are usually seen only faintly, but pulsations can be felt readily through the endoscope.[1] After the levels of the aorta and the left bronchus are passed, the lumen of the esophagus seems to disappear anteriorly; the patient's head is lowered as required to maintain visualization.[1] As the esophagoscope passes through the hiatus for the esophagus of the diagphragm, the head is lowered and moved horizontally to the right so that the axis of the tube corresponds to the axis of the lower third of the esophagus. As the esophagus deviates to the left and turns anteriorly, the hiatus is usually apparent as an anteriorly placed narrowing. The head and shoulders at this time will be found to be considerably below the plane of the tabletop.[1] In addition to its anterior curvature, the lower third of the esophagus turns strongly to the left, so an esophagoscope inserted from the right angle of the mouth, when introduced into the stomach, points in the direction of the anterior superior spine of the left ileum.[1] As the hiatus is passed, the stellate pattern of the esophagus is replaced by the deeper red, velvety gastric mucosa with prominent rugal folds.

A side-channel suction can be used to aspirate secretions that interfere with visualization. The tip of the suction catheter should not be advanced beyond where it can be visualized so that the tip does not cause a perforation. During the procedure, the suction catheter can be left sheathed within the esophagoscope when not in active use. If the telescope is held in position by a rubber cap, then the length to which the telescope is inserted into the esophagoscope can by varied by the endoscopist. If the esophagoscope becomes flooded with secretions, the telescope can be partially withdrawn proximally and the suction catheter can be positioned with good visualization. Withdrawing the telescope a few centimeters also allows visualization of side-channel instruments as they are advanced before they project into the lumen of the esophagus.

Specular esophagoscopy comprises inspection of the hypopharynx and upper esophagus with an esophageal speculum or a long laryngoscope.[1] The larynx is exposed as in direct laryngoscopy, the right piriform sinus is identified, and the tip of the speculum is inserted therein and gently insinuated to the cricopharyngeal constriction. Too great extension of the head is to be avoided; slight flexion at the occipitoatloid joint is useful at times. Moderate anterior or upward traction pulls the cricoid away from the posterior pharyngeal wall and opens the lumen of the esophagus. The speculum has the disadvantage of limited length so that should the foreign body move downward, it cannot be followed with this instrument.[1]

Fiber-optic Esophagoscopy

Fiber-optic esophagoscopy can be performed under general or local anesthesia. For conscious patients, topical anesthetic agents are sprayed into the oral cavity and oropharynx, a bite guard is placed, and intravenous sedation may be useful. Because the primary complication of endoscopy is cardiopulmonary compromise, vital signs and oxygen saturation should be monitored during the procedure.

As with bronchoscopes, the flexible fiber-optic esophagoscopes have multiple accessories for diagnostic and therapeutic management of esophageal

pathology. Probes for Doppler, ultrasonography, and pH monitors as well as biopsy forceps, baskets, and cytologic brushes and needles are available.

The endoscopist controls the air, water, and suction valves and tip direction wheels. The right hand advances the instrument; rotating the instrument control handles applies torque to the esophagoscope, rotating the shaft and tip.

The esophagoscope can be inserted blindly, by feel, through the esophageal inlet. It can also be inserted under direct visualization; the tip of the endoscope is gently advanced into the hypopharynx to a distance approximately 18 cm from the incisors, reaching the level of the upper esophageal sphincter or muscle. The cricopharyngeus is closed at rest and opens during swallowing. The esophageal inlet is visualized, and the endoscope is gently advanced into the esophagus as the patient swallows. Difficulty passing the esophagoscope past the cricopharyngeus indicates obstruction, which may result from cervical spine spurs, tumors, postsurgical stenosis, or abnormal neural function. Examination is performed to a distance of 40 cm, where the ora serrata is located. This is also referred to as the squamocolumnar mucosal junction or Z line, demarcating the junction between pearly, esophageal, stratified squamous mucosa and the redder gastric columnar epithelium. Fiber-optic esophagoscopes are limited in their ability to visualize the region of the introitus and cricopharyngeus, which are common places for foreign-body lodgment.

Complications of Esophagoscopy

Potential complications of esophagoscopy include cricoarytenoid joint dislocation, bleeding, and perforation[1]; in addition, insufflation used during fiber-optic esophagoscopy can cause abdominal distention, contributing to respiratory compromise. Esophageal perforation may occur during or following esophagoscopy or surgical manipulation such as diagnostic biopsy, dilatation, and repair of tracheoesophageal fistula or hiatal hernia or as a consequence of ulceration or erosion of the esophageal wall from cancer, ectopic gastric mucosa, esophageal peptic ulcer, or foreign body.[47] Judicious choice of instrument diameter and length, maintaining visualization of the esophageal lumen, gentleness of touch, judicious biopsy of abnormal areas, avoidance of persistent attempts to pass the esophagoscope through a neoplastic area, and gentle dilation[47] will help prevent complications.

Perforations most commonly occur at the weakest points of the esophageal wall, such as the area just below the cricopharyngeus, the distal one-fifth, or the site of any local lesion that weakens the wall. The diagnosis is usually suggested by retrosternal or back pain, fever, or subcutaneous emphysema.[1,47] Following rigid or fiber-optic esophagoscopy, patients may be observed and not fed for varying lengths of time; a chest radiograph may be obtained to evaluate the possibility of free mediastinal air or pleural effusion, infiltrate, or pneumothorax. Either or both hemithoraces may be affected. Opaque media given by mouth can demonstrate the site of leakage.[47] Because the retrovisceral space and the superior mediastinum have practically no tissue barriers, infection following rupture spreads rapidly and extensively[47] and may result in septic mediastinitis.[1] Perforation caused by esophagoscopy or following esophageal surgery can often be managed by prohibiting oral intake and administering antibiotics and intravenous fluids; a feeding jejunostomy may be indicated. If the tear is large or there is evidence of infection in the pleural cavities or mediastinum, appropriate drainage should be established; fluid in the chest necessitates drainage.[47] Perforation resulting from cancer or ulceration requires repair of the rent, drainage, and antibiotics.[47] In a study of 24 cases of esophageal perforation, mortality was 25% if the perforation was recognized within 4 hours and 86% when recognition was delayed 24 hours or longer. Mortality was also higher in patients under 10 years and older than 60 years of age.[47]

ACKNOWLEDGMENTS

The authors would like to thank Marc Rosenthal, MD, and Gregory Boger, MD, for their assistance in demonstrating endoscopy techniques, and Joyce A. Schild, MD, and James B. Snow Jr, MD, for their generosity in allowing us to use portions of their material from previous editions of this book.

REFERENCES

1. Jackson C. Bronchoscopy and esophagoscopy. A manual of peroral endoscopy and laryngeal surgery. Philadelphia: WB Saunders; 1922.
2. Karmody CS. Part I: Historical perspectives, the history of laryngology. In: Fried MP, editor. The larynx:

a multidisciplinary approach. 2nd ed. St. Louis: CV Mosby; 1996. p. 3–11.
3. Huizinga E. On esophagoscopy and sword-swallowing. Ann Otol Rhinol Laryngol 1969;78:32–9.
4. Prakash UB. Advances in bronchoscopic procedures. Chest 1999;116:1403–8.
5. Steiner W. Techniques of diagnostic and operative endoscopy of the head and neck. Tracheoscopy, bronchoscopy, esophagoscopy, mediastinoscopy, interdisciplinary panendoscopy. Endoscopy 1979;11:151–7.
6. Colton RH, Casper JK, Minoru H, editors. Understanding voice problems. Baltimore: Williams & Wilkins; 1990.
7. Storz K. The world of endoscopy: extract from catalog endoscopes and instruments for ENT. 5th ed. Culver City (CA): Karl Storz Endoscopy-America; 1997.
8. Wood RE. Pitfalls in the use of the flexible bronchoscope in pediatric patients. Chest 1990;97:199–203.
9. Hochman II, Zeitels SM, Heaton JT. Analysis of the forces and position required for direct laryngoscopic exposure of the anterior vocal folds. Ann Otol Rhinol Laryngol 1999;108:715–24.
10. Ferlito A, Rinaldo A, Marioni G. Laryngeal malignant neoplasms in children and adolescents. Int J Pediatr Otorhinolaryngol 1999;49:1–14.
11. Hartnick CJ, Hartley BE, Miller C, Willging JP. Pediatric fiberoptic endoscopic evaluation of swallowing. Ann Otol Rhinol Laryngol 2000;109:996–9.
12. Hayes JT, Houston R. Flexible nasolaryngoscopy. A low-risk, high-yield procedure. Postgrad Med 1999; 106:107–14.
13. Bless DM, Swift E. Stroboscopy: new diagnostic techniques and applied physiology. In: Fried MP, editor. The larynx: a multidisciplinary approach. 2nd ed. St. Louis: CV Mosby; 1996. p. 81–100.
14. Bastian RW. Contemporary diagnosis of the dysphagic patient. Otolaryngol Clin North Am 1998;31:489–506.
15. Hunsaker DH. Anesthesia for microlaryngeal surgery: the case for subglottic jet ventilation. Laryngoscope 1994;104:1–30.
16. Cordasco EM, Mehta AC, Ahmad M. Bronchoscopically induced bleeding. A summary of nine years' Cleveland Clinic experience and review of the literature. Chest 1991;100:1141–7.
17. Gotta AW, Sullivan CA. Anaesthesia of the upper airway using topical anaesthetic and superior laryngeal nerve block. Br J Anaesth 1981;53:1055–8.
18. Azarnoff D, Belknap S, Branch RA, et al, editors. Physicians GenRx; the complete drug reference. St. Louis: CV Mosby; 1996.
19. Adriani J, editor. Regional anesthesia, technical and clinical applications. 4th ed. St. Louis: Warren H. Green; 1985.
20. Gaskill JR, Gillies DR. Local anesthesia for peroral endoscopy. Using superior laryngeal nerve block with topical application. Arch Otolaryngol 1966;84: 654–7.
21. DeMeester TR, Skinner DB, Evans RH, Benson DW. Local nerve block anesthesia for peroral endoscopy. Ann Thorac Surg 1977;24:278–83.
22. Wiles JR, Kelly J, Mostafa SM. Hypotension and bradycardia following superior laryngeal nerve block. Br J Anaesth 1989;63:125–7.
23. MICROMEDEX® Healthcare Series. Cocaine hydrochloride; lidocaine; oxymetazoline hydrochloride; phenylephrine hydrochloride; tetracaine (DrugPoints® System); cocaine (ToxPoints® System) [Serial online]. 2000;106. http://hcs.mdx.com/ (accessed Dec 8, 2000) (a password from MICROMEDEX is needed to access the site and its databases).
24. PDR.net for Physicians. PDR® entries for Cetacaine topical anesthetic (Cetylite) and cocaine hydrochloride topical solution (AstraZeneca LP) [Serial online]. 2000. http://physician.pdr.net/physician/index.htm (accessed Dec 12, 2000) (a password from PDR.net for Physicians is needed to access the site and its databases).
25. Taketomo CK, Hurlburt Hodding J, Kraus DM. Pediatric dosage handbook. 7th ed. Hudson (OH): Lexi-Comp; 2000.
26. Dajani AS, Taubert KA, Wilson W, et al. Prevention of bacterial endocarditis: Recommendations by the American Heart Association. AHA Medical/Scientific Statement [Serial online]. 1997. http://www.americanheart.org/Scientific/statements/1997/079701.html (accessed June 21, 2001).
27. Zeitels SM. Universal modular glottiscope system: the evolution of a century of design and technique for direct laryngoscopy. Ann Otol Rhinol Laryngol Suppl 1999;179:2–24.
28. Stewart CM, Stewart LB. Pediatric medications: an emergency and critical care reference. Rockville (MD): Aspen Systems Corp; 1984.
29. Kalyanaraman M, Carpenter RL, McGlew MJ, Guertin SR. Cardiopulmonary compromise after use of topical and submucosal alpha-agonists: possible added complication by the use of beta-blocker therapy. Otolaryngol Head Neck Surg 1997;117: 56–61.
30. Maze A, Bloch E. Stridor in pediatric patients. Anesthesiology 1979;50:132–45.

31. Rai NS, Arroliga AC. Indications, contraindications, and medications. In: Wang K-P, Mehta AC, editors. Flexible bronchoscopy. Cambridge (MA): Blackwell Science; 1995. p. 34–59.
32. Keane WM, Atkins JP. CO_2 laser surgery of the upper airway. Surg Clin North Am 1984;64:955–71.
33. Nuss RC, Hillman RE, Eavey RD. Office and operative diagnostic techniques. In: Fried MP, editor. The larynx: a multidisciplinary approach. 2nd ed. St. Louis: CV Mosby; 1996. p. 65–74.
34. Andrea M, Dias O. Rigid and contact endoscopy associated with microlaryngeal surgery. In: Fried MP, editor. The larynx: a multidisciplinary approach. 2nd ed. St. Louis: CV Mosby; 1996. p. 75–9.
35. Eavey RD. The pediatric larynx. In: Fried MP, editor. The larynx: a multidisciplinary approach. 2nd ed. St. Louis: CV Mosby; 1996. p. 25–31.
36. Brooker CR, Hunsaker DH, Zimmerman AA. A new anesthetic system for microlaryngeal surgery. Otolaryngol Head Neck Surg 1998;118:55–60.
37. Griner RL. AANA journal course: update for nurse anesthetists—the laryngeal mask airway: attributes and inadequacies. AANA J 1996;64:485–96.
38. Baltimore JJ. Postlaryngospasm pulmonary edema in adults. AORN J 1999;70:468–79.
39. de Blic J, Scheinmann P. Fibreoptic bronchoscopy in infants. Arch Dis Child 1992;67:159–61.
40. Proctor DF. The upper airways II. The larynx and trachea. Am Rev Respir Dis 1977;115:315–42.
41. Lore JM Jr. An atlas of head and neck surgery. 3rd ed. Philadelphia: WB Saunders; 1988.
42. Agur AMR, editor. Grant's atlas of anatomy. 9th ed. Baltimore: Williams & Wilkins; 1991.
43. Atkins JP, Rowe LD, Jafek BW. Newer techniques in fiberoptic bronchoscopy. Ann Otol Rhinol Laryngol 1976;85:646–51.
44. Meduri GU, Chastre J. The standardization of bronchoscopic techniques for ventilator-associated pneumonia. Chest 1992;102:557S–64S.
45. Slonim AD, Ognibene FP. Amnestic agents in pediatric bronchoscopy. Chest 1999;116:1802–8.
46. Schiratzki H. Removal of foreign bodies in the esophagus. An old device—a new procedure. Arch Otolaryngol 1976;102:238–40.
47. Rabinovich S, Smith IM, McCabe BF. Rupture of the esophagus. Arch Otolaryngol 1967;85:410–5.

CHAPTER 66

Laryngoscopy

Ellen S. Deutsch, MD, Jane Y. Yang, MD, James S. Reilly, MD

The larynx comprises the supraglottis, consisting of the epiglottis, aryepiglottic folds, false vocal folds, and laryngeal ventricle; the glottis or true vocal folds; and the subglottis, including the cricoid cartilage. Proctor described the laryngotracheal airway as "the bottleneck of the respiratory system. Whereas disease in the nose or in the bronchial airways will affect respiratory airflow in a limited way, laryngotracheal disease must always be considered potentially life threatening."[1]

LARYNGEAL DISEASE

Laryngeal disease generally presents with pain, stridor, or hoarseness. Hoarseness is a precise localizing feature, indicating that the abnormality involves, but is not necessarily limited to, the free margin of the vocal folds.

Benign Masses

Vocal fold nodules, polyps, and cysts generally occur at the junction of the anterior and middle third of the vocal fold and involve the epithelial and superficial layers of the lamina propria. Vocal fold nodules (Figure 66–1), generally bilateral and caused by voice abuse, are small, white, firm, sessile masses containing collagenous fibers and edema.[2] A nodule as small as 1 mm may impair the voice.[1]

Vocal fold cysts, usually unilateral, comprise epidermoid cysts and mucous retention cysts. Epidermoid cysts, caused by vocal trauma, arise from an infolding of the mucosa and are filled with caseous material. Mucous retention cysts arise from obstructed excretory ducts and are filled with mucoid material.[2] Posthemorrhagic cysts may contain blood products.

Polyps, caused by vocal trauma, are usually unilateral, pedunculated, or sessile and may have evidence of preceding hemorrhage. A prominent feeding vessel may be observed on the superior aspect of the vocal fold. Histologic analysis may reveal hyalin degeneration, thrombosis, edema, collagenous fibrous proliferation, and cellular infiltration.[2] Polyps are removed by grasping them with cup forceps and retracting them toward the midline. An incision is made through the mucosa into Reinke's space, and the polyp is separated from the underlying vocal ligament. Small mucosal tags can be vaporized or "spot welded" with a laser to provide a smooth surface.[3]

Reinke's edema, also known as polypoid corditis or polypoid degeneration of the vocal folds, is most commonly associated with tobacco use, although vocal misuse and extraesophageal reflux disease may contribute to this condition. The vocal folds have a sausage-shaped edematous appearance. Histologically, the superficial layer of the lamina propria is replaced by a myxomatous-appearing stroma.

Granulomas commonly occur on the vocal process or inner surface of the body of the arytenoid cartilage and can be large enough to obstruct the airway.[4] Primary etiologic factors include traumatic or prolonged intubation, vocal abuse or other trauma, or extraesophageal reflux disease. Allergy, infection, postnasal drainage, smoking, and psychosomatic factors may contribute.[5–7] Small lesions may be managed with voice and antireflux therapy,[5] but recurrence is common.[4] Adjunctive therapies include antibiotics, topical or systemic corticosteroids, and elimination of sources of chronic irritation.[4] Larger or persistent lesions may also require laser or other surgical resection.[3,4] Injection of botulinum toxin or intralesional corticosteroids may be successful for recalcitrant lesions.[6]

Benign neoplasms are less frequent causes of hoarseness. Laryngeal rhabdomyomas are benign but may recur if excision is incomplete. Despite the histologic resemblance to cardiac rhabdomyomas,

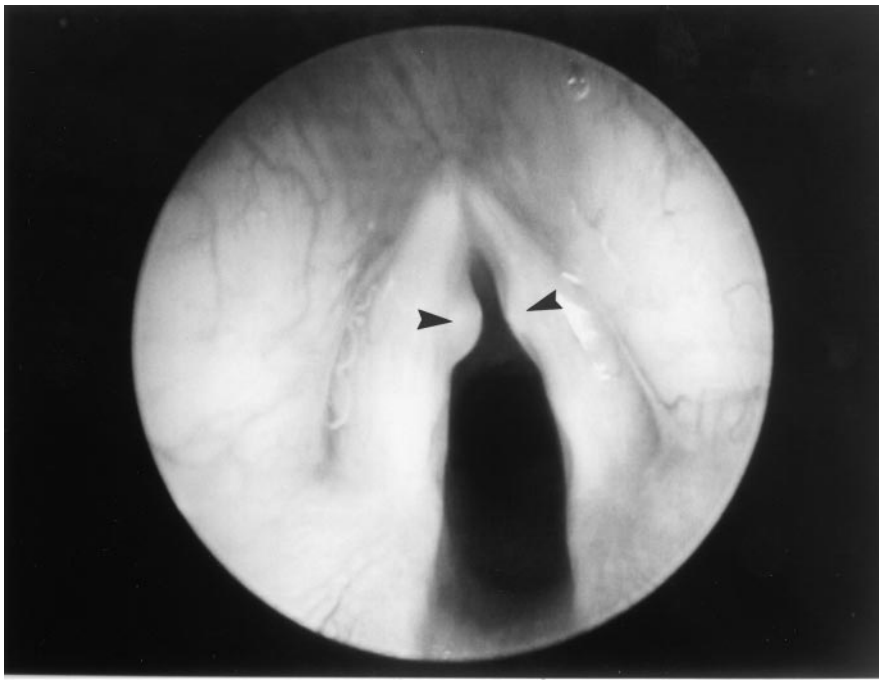


FIGURE 66–1. Vocal fold nodules (*arrows*).

extracardiac rhabdomyomas do not have the same association with tuberous sclerosis.[8]

Cartilaginous laryngeal neoplasms, such as chondromas, have a 5 to 1 male predominance; most occur in patients between the ages of 40 and 60 years. If located within the cricoid cartilage, dyspnea may be a more prominent symptom than hoarseness. Endoscopically, the lesions have a smooth, encapsulated appearance. Biopsy of hard masses, such as chondromas, may be difficult because the hardness of the lesion may permit removal only of overlying mucosa.[8]

Granular cell tumor, previously referred to as granular cell myoblastoma, is a benign, slow-growing neoplasm that can often be removed endoscopically.[8] The characteristic, but not pathognomonic, histologic feature is pseudoepitheliomatous hyperplasia. Immunohistochemical staining is positive for S-100, documenting neurogenic origin.

Neurofibromas and neurilemomas or schwannomas are histologically similar, benign, neurogenic tumors of the larynx originating from the aryepiglottic fold or from the false vocal folds. Reported age at presentation ranges from 3 months to 75 years of age; the lesions are smooth, with intact overlying mucosa. Biopsy may be difficult because the firmness and encapsulation of the mass resist the biopsy forceps; attempts to obtain adequate specimens may cause excessive bleeding. Extralaryngeal approaches may be required for resection of large lesions.[8]

Paragangliomas and atypical carcinoids may have similar clinical appearances. Paragangliomas present as red or blue submucosal masses most commonly located in the supraglottis. Atypical carcinoids may be similar in color but can also be polypoid, ulcerated, nodular, or pedunculated. Paragangliomas are the only laryngeal neoplasm with a female predominance; atypical carcinoid most commonly occurs in males. Both are of neuroendocrine origin, contain neurosecretory granules, and are distinguished by careful immunocytochemical and/or ultrastructural investigation. In contrast to other head and neck paragangliomas, such as chemodectomas, laryngeal paragangliomas do not secrete catecholamines. Treatment for atypical carcinoid is surgical. Regional metastases are common at the time of diagnosis; death usually occurs from distant metastases rather than local recurrence. Paragangliomas have an excellent prognosis and are managed with conservative surgery.[8–10]

Hemangiomas occur in both adult and infantile types. The infantile type typically presents as a compressible, sessile mass in the subglottis. They are usually unilateral but may be circumferential; histologically, they may be cavernous, simple, or capillary in appearance. Infantile subglottic hemangiomas generally increase in size and symp-

tom severity until 6 months of age, after which spontaneous regression occurs. The adult form occurs on or above the vocal folds and is often pedunculated, well demarcated, and reddish blue in color. Biopsy is generally considered both dangerous and unnecessary; diagnosis is based on clinical history and endoscopic appearances[8] and may be supported by radiologic studies. Kaposiform hemangioendothelioma is a rare, aggressive, vascular neoplasm, often associated with Kasabach-Merritt syndrome (platelet trapping associated with thrombocytopenia purpura), which is clinically and histologically distinct from the classic hemangioma of infancy (Figure 66–2).

Lipomas are usually solitary lesions found in men during the seventh decade of life; most can be removed endscopically.[8]

Extralaryngeal lesions may compress the airway and cause stridor by exerting a mass effect. Examples include retropharyngeal abscesses, adjacent tumors, or postcricoid foreign bodies (Figure 66–3).

Autoimmune, immune, and idiopathic diseases may have laryngeal manifestations. Sarcoidosis, discussed further in Chapter 67, is an idiopathic, chronic, noncaseating granulomatous disease that presents with dysphonia, dysphagia, globus sensation, and dyspnea.[11] Laryngeal sarcoidosis most commonly involves the supraglottis, particularly the epiglottis, which appears diffusely enlarged with occasional nodularity.

Scleroma (rhinoscleroma) is a chronic granulomatous infectious disease caused by *Klebsiella rhinoscleromatis*, which primarily occurs in the nose and nasopharynx but may involve the larynx.[12] Treatment includes antibiotics; surgical resection may be necessary for life-threatening airway obstruction.[12]

Wegener's granulomatosis is a necrotizing vasculitis typically involving a triad of organs: the sinuses, lungs, and kidneys. Nasal manifestations may include severe nasal crusting, epistaxis, ulcerations, and septal perforation. Glomerulonephritis may present with microhematuria, proteinuria, and red cell casts. Laryngeal involvement generally causes subglottic stenosis with dyspnea and stridor (Figure 66–4); progression may be followed endoscopically or radiographically.[13,14] Elevated antineutrophilic cytoplasmic antibodies are highly specific for Wegener's granulomatosis.[14] Diagnosis is supported by histology revealing necrotizing vasculitis involving small to medium-sized arteries and multinucleated giant cells or non-necrotizing granulomas.[12] Treatment consists primarily of corticosteroids and oral cyclophosphamide.[12]

Amyloidosis is a disease of unknown etiology characterized by deposition of extracellular, fibrillar protein. It is classified by the name of the fibrillar

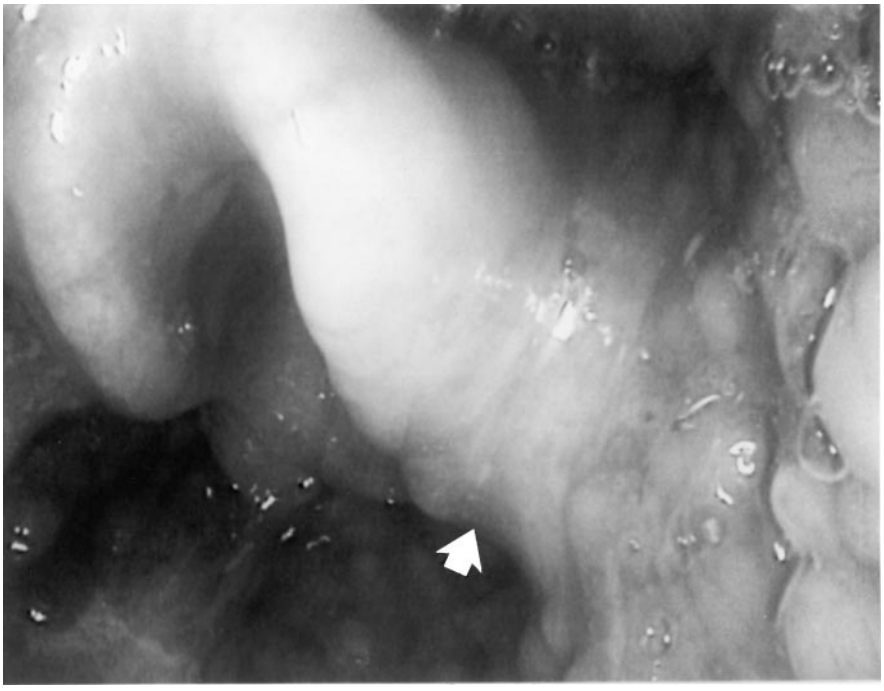


FIGURE 66–2. Kaposiform hemangioendothelioma involving the epiglottis, right pharyngoepiglottic fold (*arrow*), and hypopharynx in an infant. The subglottis and trachea were also affected, and radiologic study documented extensive mediastinal disease.

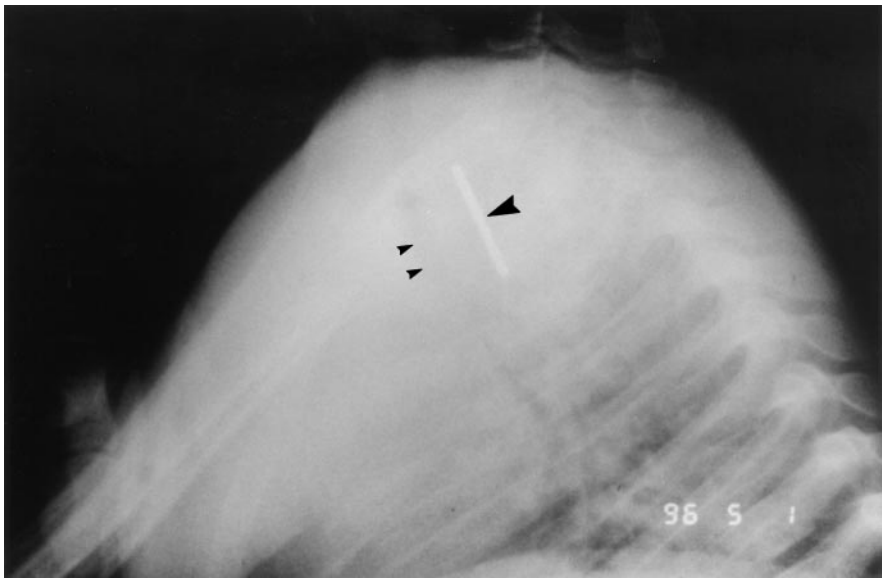


FIGURE 66–3. Lateral radiograph demonstrating a foreign body (*arrow*) in the esophagus in a toddler who had been stridorous for several months. The walls of the esophagus and trachea (*double arrow*) are markedly thickened.

protein (AL, AA, ATTR, or AB2M), the precursor protein (kappa or lambda light chain, apoSSA, transthyretin, or β_2-microglobulin), and the clinical presentation (primary, secondary multiple myeloma associated, familial, or hemodialysis associated).[15] The larynx is the most common site of involvement in the upper aerodigestive tract. In the larynx, amyloidosis is usually localized, idiopathic, and primary,[15,16] presenting with insidious onset of hoarseness. Examination reveals a diffuse submucosal mass, which may cause vocal fold motion to appear sluggish. It is important to determine whether the disease is localized or systemic; systemic disease can involve any organ system and has a poor prognosis. Previously, rectal, lip, and gingival biopsies were obtained, but abdominal fat aspiration is a safe and better-tolerated procedure for identifying systemic disease.[17] Bone marrow biopsy may be necessary to determine whether the patient has multiple myeloma. Pediatric laryngeal amyloidosis is rare.[15]

Rheumatoid arthritis is an autoimmune disorder affecting 2 to 3% of the adult population.[18] Laryngeal involvement is believed to affect up to half of patients with rheumatoid arthritis.[19] Both adult and juvenile rheumatoid arthritis may involve the cricoarytenoid joint. During the acute inflammmatory stages, laryngeal manifestations may include throat pain, dysphagia, dysphonia, odynophagia, odynophonia, globus sensation, and referred otalgia. Examination of the larynx may reveal erythema of the mucosa overlying the joint, edema, and ulcera-

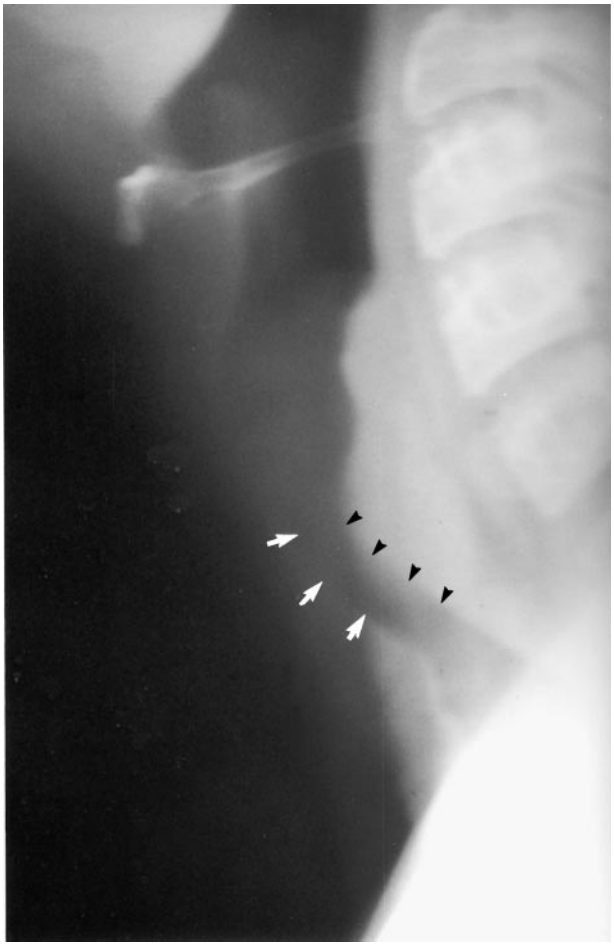


FIGURE 66–4. Lateral neck radiograph demonstrating subglottic narrowing (*arrows*) in an adolescent with Wegener's granulomatosis. (Courtesy of Eric Faerber, MD.)

tions.[1,18,20] Chronic rheumatoid arthritis may cause cricoarytenoid joint stenosis or fixation causing impaired vocal fold mobility; tracheostomy may be required for severe airway obstruction.[1] Rarely, rheumatoid arthritis causes submucosal nodules involving the true vocal folds, aryepiglottic folds, and postcricoid region.[18,20,21]

Relapsing polychondritis is a multisystemic autoimmune inflammatory disorder of obscure etiology characterized by progressive inflammation and degeneration of cartilaginous structures and connective tissues. One of the most prominent features is acute perichonditis of the pinna, sparing the noncartilaginous lobule. The disease is generally first characterized by perichondritis and cartilage dissolution in the nose and ears, but respiratory tract involvement may be the presenting symptom in as many as 26% of patients, and up to 50% of patients will eventually develop respiratory complications.[22] Symptoms of respiratory tract involvement include cough, hoarseness, aphonia, dyspnea, and tenderness over the thyroid cartilage and anterior trachea. Involvement of laryngeal cartilages can cause subglottic stenosis (Figure 66–5) or glottic aperture collapse. Laryngotracheobronchial involvement may cause life-threatening airway obstruction from acute inflammation, scarring, or airway collapse from cartilage dissolution.[1,23] McAdam et al's criteria for diagnosis require three or more of the following manifestations: bilateral auricular chrondritis, nonerosive seronegative inflammatory polyarthritis, nasal chondritis, ocular inflammation (uveitis, conjunctivitis, keratitis, episcleritis), respiratory tract (larynx, trachea, bronchi) chondritis, and audiovestibular damage (sensorineural hearing loss, vertigo, tinnitus); the ribs and heart can also be affected.[24] Laboratory studies may reveal normocytic, normochromic anemia, mild leukocytosis, and an elevated erythrocyte sedimentation rate. Medical management in the acute phases includes corticosteroids; dapsone and immunosuppresant agents, such as cyclophosphamide, may be helpful.[25]

Examination of supraglottic allergic edema reveals pale, watery swelling of the epiglottis and aryepiglottic folds.[26] The edema usually responds quickly to subcutaneous administration of 0.2 mL of epinephrine 1:1,000 for adults (0.01 mL/kg of body weight for children) plus 100 mg of intramuscular or intravenous hydrocortisone phosphate for adults[26]; antihistamines and inhaled racemic epinephrine may also be useful.[27–29] Fatal airway obstruction can occur when the larynx is affected[30]; endotracheal intubation or tracheostomy may be necessary. The acute edema usually subsides within 24 to 48 hours after initiation of treatment.

Hereditary angioedema results from an autosomal dominant deficiency of active C1 esterase inhibitor leading to recurrent attacks of localized

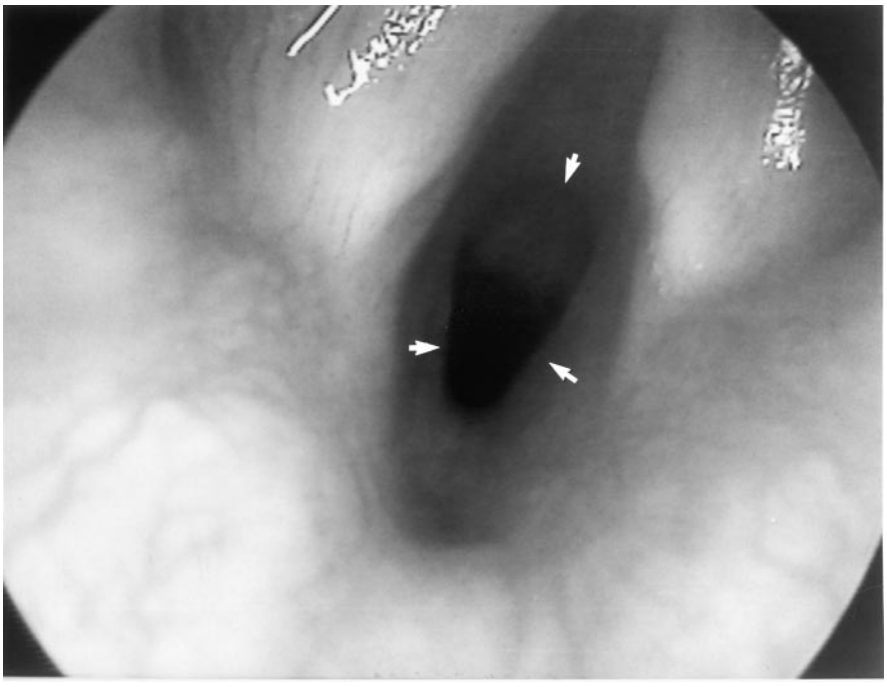


FIGURE 66–5. Endoscopic view of circumferential narrowing of the subglottis (*arrows*) in an 11-year-old boy with relapsing polychondritis.

edema or severe abdominal pain, precipitated by trauma or physical or emotional stress.[29,31] The onset is usually in childhood, and there may be a family or past medical history of self-limited attacks associated with a salmon-colored rash.[30] The deficiency of C1 esterase inhibitor protein is usually quantitative but can be qualitative.[31] Decreased levels of C4, another component of the complement activation cascade, may accompany both quantitative or qualitative C1 esterase inhibitor protein deficiencies; determination of C4 levels may be a useful screening test.[29,32] Airway occlusion is the most common cause of mortality.[31] Standard therapies for allergy-mediated angioedema, such as epinephrine, antihistamines, and corticosteroids, are less successful, and mechanical control of the airway may be needed.[31] Long-term prophylactic treatment may include attenuated androgens such as danazol or stanozolol; antifibrinolytic medications such as ε-aminocaproic acid and tranexamic acid; and C1-INH concentrate. Long-term use of attenuated androgens has been associated with hepatocellular carcinoma and should be used with caution in women, especially during pregnancy, and children. Antifibrinolytic medications are the drugs of choice for children; continuous use increases the risk of thromboembolic events.[33,34] C1-INH concentrate, effective within minutes of infusion, is used for acute management; fresh frozen plasma can also be used but is less efficacious. C1-INH concentrate or attenuated androgens are used for short-term prophylaxis before surgical and dental procedures.[29,34]

Acquired angioedema is less common and generally begins during adulthood.[31] Drug-induced angioedema most commonly occurs in patients taking angiotensin-converting enzyme (ACE) inhibitors, particularly captopril and enalapril maleate. The effects of kinase II, which stimulates the conversion of angiotensin I to the vasoconstrictive angiotensin II and inactivates bradykinin, a potent vasodepressor, are blocked by ACE inhibitors.[27]

Acute spasmodic laryngitis or nocturnal croup typically presents as a mild viral upper respiratory tract infection with nocturnal episodic wakening accompanied by inspiratory stridor and croupy cough in a toddler; it invariably subsides during the day. Examination is remarkable for the absence of severe laryngeal inflammation or fever. The etiology is unclear, but it may be related to an upper respiratory tract infection[30] or may be a manifestation of reactive airway disease.

Epiglottitis, or acute supraglottitis, is an acute bacterial infection of the epiglottis, aryepiglottic folds, and soft tissues of the arytenoid cartilages.[35] Features include the dangerously rapid onset of sore throat, high fever, muffled voice, and signs and symptoms of airway obstruction, including inspiratory stridor, inability to swallow, respiratory distress, and restlessness. The classic presentation is a pale, shocky, restless, drooling child in an upright position with head forward and tongue protruding. Examination, which is deferred until equipment and personnel are available to secure the airway, reveals fiery red edema of epiglottis and aryepiglottic folds.[26] Pediatric acute supraglottitis, now uncommon, was previously caused predominantly by *Haemophilus influenzae*. The disease is now more common in adults and may be life-threatening but is usually less severe, with slower onset and progression.

"Croup," or laryngotracheobronchitis, refers to a constellation of viral infections that may affect any or all segments of the pediatric airway. Croup is the most common illness affecting the airway in children and the most common cause of airway obstruction in children; the majority of patients are between the ages of 6 months and 4 years.[35] Patients usually have a characteristic "barking" cough, 60% have inspiratory stridor, and the voice may be hoarse but not muffled,[35] Most patients respond to medical and supportive therapy,[35] including humidification, nebulized racemic epinephrine, and systemic glucocorticoids[35]; uncommonly, intubation is necessary. Anteroposterior airway radiographs may demonstrate subglottic swelling, which changes the appearance of the square-shouldered space below the vocal folds to that of a steeple or pencil[35]; lateral neck radiographs may show subglottic haziness. Endoscopy is avoided during the acute phase of the illness but is indicated if the patient fails two extubation attempts[35] or does not improve as expected. Interval endoscopy is indicated if the patient has other risk factors for subglottic stenosis, more than three episodes of croup, onset of croup prior to 6 months of age, or an atypical course.

Recurrent respiratory papilloma is a benign but very rapidly proliferating tumor[30] that usually presents within the first few years of life with hoarseness, inspiratory stridor, and respiratory distress progressing over a period of weeks or months.[30] Human papillomavirus (HPV) subtypes 6 and 11 are most commonly found in juvenile recurrent respiratory

papilloma.[36] Treatment consists of repeated excision using carbon dioxide laser or a microdébrider.[3,36] A variety of adjunctive therapies have been tried, including interferon alpha-2a, cruciferous vegetables, indole-3-carbinol, acyclovir, ribavirin, mumps vaccine, isotretinoin, and photodynamic therapy.[36] Recent trials of cidofovir appear promising.[36] Tracheostomy may be necessary but may be associated with seeding of papillomas locally and distally.[30] Malignant degeneration of recurrent respiratory papilloma is uncommon, but the risk is greater for patients with subtypes 16/18 and 31/33[37] and those treated with radiation therapy.[38]

Bacterial or fungal laryngitis and tracheitis (Figure 66–6) can cause sudden airway obstruction.

Laryngopharyngeal reflux, or extraesophageal reflux, denotes the entry of gastric contents above the upper esophageal sphincter. More than 60% of adults with suspected laryngopharyngeal reflux do not have the classic gastroesophageal reflux symptoms of heartburn and reflux.[39,40] Otolaryngologic manifestations and sequelae of extraesophageal reflux include chronic intermittent hoarseness, vocal fatigue, sore throat, dysphagia, chronic cough, stridor, croup, subglottic stenosis, excessive mucus production, postnasal drip, globus sensation, a choking sensation resulting from laryngospasm, and frequent throat clearing.[40,41] Pulmonary conditions associated with reflux include asthma, pulmonary fibrosis, bronchiectasis, chronic bronchitis, and pneumonia.[42] Other oral cavity, nasal, and nasopharyngeal manifestations include chronic sinusitis, otalgia, and otitis media.[42] Laryngeal examination may reveal edema, which may be diffuse or limited to the vocal folds, obliteration of the ventricles, interarytenoid thickening, pseudosulcus, or vocal process granuloma. Effects within the posterior part of the larynx or trachea include erythema, edema, "cobblestoning," blunting of the carina, and friable mucosa.[39]

Two mechanisms of injury may coexist: reflex mechanisms may cause laryngospasm, bronchospasm, or apnea, and prolonged direct contact with the gastric fluid may cause mucosal damage. In infants, some gastroesophageal reflux is considered normal; however, it increases the risk of postintubation subglottic stenosis and should be controlled in candidates for laryngotracheal reconstruction and repair of cleft larynx.[35] Twenty-four-hour multichannel ambulatory esophageal pH monitoring is the most sensitive study for extraesophageal reflux; esophagoscopy with biopsy, barium esophagram, nuclear medicine scintiscan, and bronchoscopy with washings for fat-laden macrophages may be supportive.[35]

Treatment consists of a combination of antireflux behavioral and dietary modifications and acid-suppressive medications, including histamine H_2-receptor antagonists and proton pump inhibitors. Patients who are unable to tolerate the medications or those who continue to have acid reflux despite medications may benefit from surgical therapy such as Nissen fundoplication.[35]

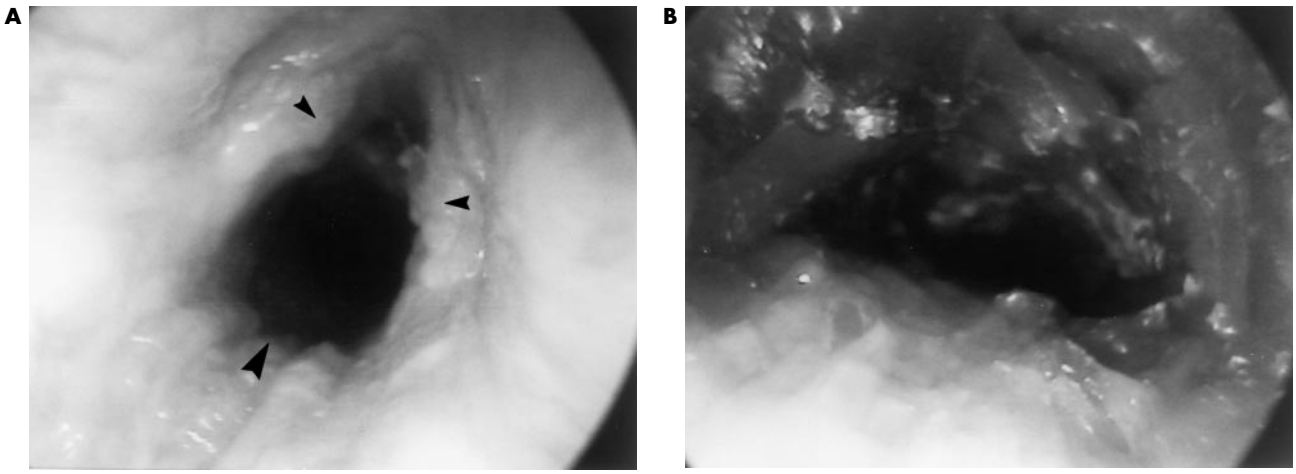


FIGURE 66–6. Endoscopic view of the larynx and trachea in a 4-year-old with diffuse *Candida* laryngitis (*A*) and tracheitis (*B*) with macrophage-activating illness. Note crusted lesions in the posterior commissure (*large arrowhead*) and obscuring the vocal folds anteriorly (*small arrowheads*) and massive crusts within the trachea.

Acute airway obstruction from laryngeal or tracheal edema is possible after ingestion of acid, alkali, or corrosive substances; following inhalation of hot air, smoke, steam, or chemical fumes[30,35]; or following direct mechanical trauma, such as a finger sweep maneuver. Traumatic epiglottitis may be treated with fluids, humidified supplemental oxygen, nebulized epinephrine, and parenteral administration of corticosteroids[30]; intubation may be necessary. Intraluminal trauma may occur from intubation or other causes and can result in acquired glottic or subglottic stenoses, subglottic cysts, synechiae, and webs. Pressure from endotracheal or tracheostomy tube cuffs may cause tracheomegaly (Figure 66–7).

External causes of laryngeal injury include blunt or penetrating trauma. The majority of blunt injuries result from motor vehicle crashes but may also occur from falls, sporting injuries, including "clothesline" injuries, or assaults. Penetrating injuries are most commonly caused by knives or bullets.[43] Minor injuries may consist of edema, hematoma, contusion, abrasion, and small lacerations; major injuries may include loss of soft tissue, large lacerations without approximation, fracture of cricoid or thyroid cartilages, displacement of the arytenoid cartilages or epiglottis, and impaired vocal fold mobility.[43] Cricoarytenoid joint dislocation, hematoma or soft tissue swelling, intralaryngeal edema, or fracture of the cartilaginous laryngeal framework[30] may cause acute respiratory obstruction. Airway obstruction may occur instantly, gradually over hours, or with great rapidity hours after the injury has occurred.[1] Computed tomography (CT) is the radiologic modality of choice; axial sections may demonstrate the extent of soft tissue edema and hematoma formation, the location of fractures, deformity of cartilaginous structures secondary to fractures, dislocation of the arytenoid cartilages, and disjunction of the cricothyroid joint.[44] Laryngoscopy is indicated when evidence of airway injury or obstruction occurs following trauma to the neck, as well as in patients sustaining high-risk injuries, even with minimal symptoms. Tracheostomy or intubation may be needed to maintain the airway.[30] The best airway and voice outcomes are achieved after penetrating injury and when treatment is instituted within 24 hours of injury.[43]

Congenital laryngeal abnormalities, including teratomas or hamartomas, laryngomalacia, subglottic stenosis, complete or partial laryngeal atresia including congenital laryngeal webs, laryngoesophagotracheal clefts, hemangiomas, and lymphangiomas, are discussed in Chapter 45.

Vocal fold paralysis may be congenital or acquired; the neural lesion may be central or peripheral. Central bilateral vocal fold paralysis may result from iatrogenic increased intracranial pressure[30] or congenital nerve compression, such as occurs with a Chiari brainstem malformation. Peripheral causes of paralysis are usually associated with injury to the recurrent laryngeal nerves, particularly during cardiovascular or thoracic surgery or operations on the

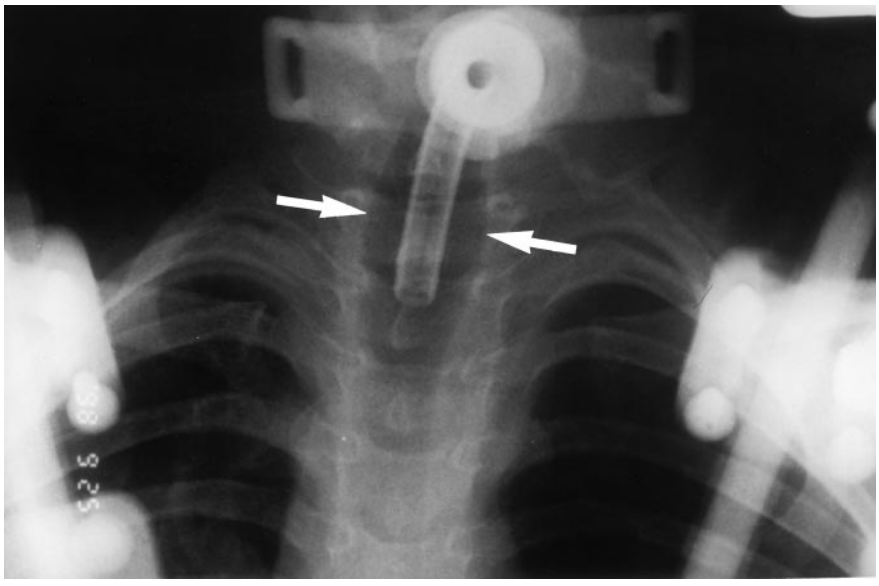


FIGURE 66–7. Chest radiograph demonstrating tracheomegaly (*arrows*) from a cuffed tracheostomy tube in a ventilator-dependent child.

neck such as thyroid surgery[30]; the left recurrent laryngeal nerve is more frequently injured. Other causes include involvement in benign or malignant tumors or inflammation in the neck or mediastinum; cardiovascular abnormalities, such as aneurysm; central nervous system disease; or neuromotor disorders such as myasthenia gravis and postpolio syndrome.[1]

Laryngeal nerve paralysis may result in a flaccid, shortened vocal fold that can be in a paramedian, intermediate, or lateral position relative to the midline and may result in loss of voice and aspiration from lack of adduction.[1,45,46] Medialization procedures may be performed for unilateral vocal fold paralysis to improve the vocal quality and decrease aspiration. Conversely, endoscopic laser arytenoidectomy or cordotomy, performed to improve the airway in bilateral vocal fold paralysis may result in a weaker voice.[3]

Speech and swallowing disorders, including dysarthria, dysphagia, or aspiration, may result from a combination of neurogenic, structural, cardiorespiratory, metabolic, and/or behavioral abnormalities.[47]

MALIGNANCIES

The larynx is the second most common site of cancer in the upper aerodigestive tract. A total of 12,500 new cases are diagnosed in the United States each year, with an estimated 5-year survival rate of 68% overall. Squamous cell carcinoma accounts for over 95% of all laryngeal malignancies, with a peak incidence in the sixth and seventh decades. Tobacco is the single most significant risk factor; ethanol is believed to have a synergistic effect with tobacco. Controversy exists as to whether ethanol alone is an independent risk factor.[48] Palliation and cure of laryngeal squamous cell carcinomas may be accomplished using radiation therapy, chemotherapy, laser,[3] or open surgical resection singly or in various combinations. Treatment decisions are based on the extent of local disease, results of previous treatment, and presence of metastases. Verrucous carcinoma, an exophytic, highly differentiated variant of squamous cell carcinoma, constitutes 1 to 3% of all laryngeal carcinomas. It rarely metastasizes, and surgical resection is the treatment of choice.[48] Laryngeal squamous cell carcinomas are rare in children.[38]

Pleomorphic carcinoma may histologically resemble other laryngeal supporting tissue neoplasms. Although treatment is predicated on the fact that the lesion is a squamous cell carcinoma, polypoid or pedunculated lesions portend a favorable outcome.[8]

Rhabdomyosarcomas are the most common type of malignant laryngeal tumor in children and adolescents.[38] Histologically, they are usually of the embryonal type and may be confused with benign lesions such as granular cell tumor and rhabdomyoma.[8] Treatment may include chemotherapy alone or with radiation therapy. Surgery is usually reserved for diagnostic biopsy or debulking lesions; extensive debilitating surgery is usually not indicated.[38]

Well-differentiated laryngeal fibrosarcomas usually do not metastasize, manifest a low recurrence rate, and have a low mortality rate. In contrast, poorly differentiated laryngeal fibrosarcomas are less common but more likely to metastasize or recur and are fatal in more than half of the reported instances.[8]

Primary tumors including adenocarcinoma, malignant fibrous histiocytoma, non-Hodgkin's lymphoma, malignant schwannoma, mucoepidermoid carcinoma, and primitive neuroectodermal tumor, as well as other sarcomas, such as synovial sarcoma, Ewing's sarcoma, chondrosarcoma, and "mixed" sarcoma, are rare.[38]

Metastases to the larynx are also rare and include metastases from testicular choriocarcinoma in adolescents;[38] additional evaluation may include a skeletal survey, bone scan, bone marrow aspirate or biopsy, CT of the chest,[38] and liver function tests.

RADIOLOGIC EVALUATION OF THE LARYNX

Routine radiologic evaluation comprises anteroposterior and lateral neck images. The lateral view of the neck provides useful information about the base of the tongue, vallecula, thyroid and cricoid cartilages, posterior pharyngeal wall, precervical soft tissues, and intralaryngeal structures including the epiglottis, aryepiglottic folds, arytenoids, false and true vocal folds, laryngeal ventricles, and subglottic space (Figure 66–8).

Fluoroscopy may be useful in assessing dynamic lesions. A cooperative patient can phonate /ee/ during inspiration to accentuate vocal cord motion and open the ventricles or breath-hold dur-

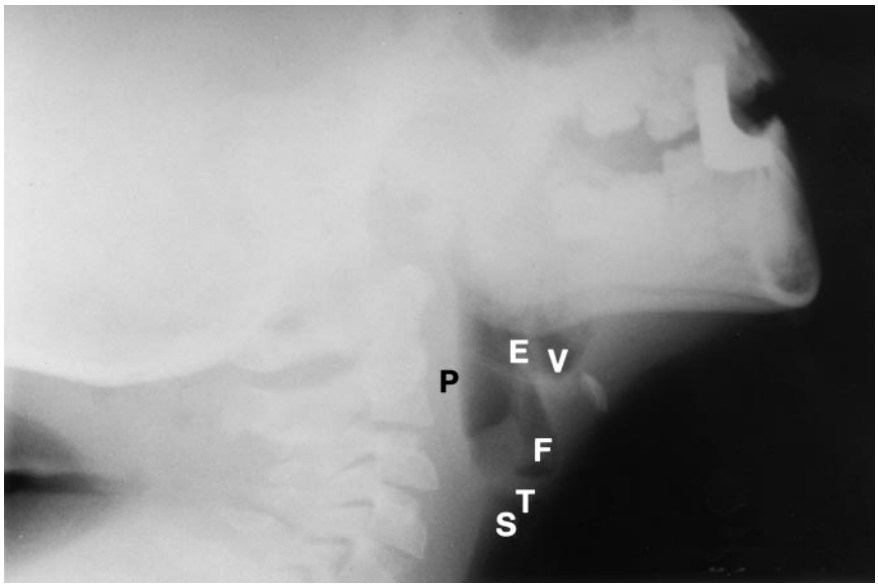


FIGURE 66–8. Lateral neck radiograph of a 6-year-old child with a large adenoid, demonstrating the valleculae (V), the epiglottis (E) crossed by the hyoid bone, the prevertebral soft tissue (P), the laryngeal ventricle between the false (F) and true (T) vocal folds, and the subglottis (S).

ing a modified Valsalva's maneuver to distend the piriform sinuses. Adding radiologic contrast agents such as barium or iohexol for tracheobronchography may enhance visualization and further define airflow dynamics and distal anatomy.[44,49]

Technological advances allow radiologic evaluation of deep tissue involvement by laryngeal disease. Spiral CT and fast magnetic resonance imaging techniques allow rapid acquisition of enhanced images of laryngeal structures, minimizing motion artifact. Direct spiral CT of the neck is performed in the axial plane; coronal, sagittal, and three-dimensional reconstruction can be computer generated. Magnetic resonance imaging has the advantage of multiplanar high-resolution imaging with an increased ability to separate various soft tissues and may be more sensitive for identifying early cartilage invasion[38,50] Both modalities are invaluable in preoperative malignancy staging, facilitating surgical planning. When possible, images should be obtained prior to biopsy to avoid artifact from postsurgical inflammation.[51]

REFERENCES

1. Proctor DF. The upper airways II. The larynx and trachea. Am Rev Respir Dis 1977;115:315–42.
2. Sataloff RT. Structural abnormalities of the larynx. In: Sataloff RT, editor. Professional voices: the science and art of clinical care. 2nd ed. New York: Singular Publishing Group; 1997. p. 509–40.
3. Keane WM, Atkins JP. CO_2 laser surgery of the upper airway. Surg Clin North Am 1984;64:955–71.
4. Lai S, Kelleher K, Sataloff RT. Vocal fold granuloma: the "ball-valve" phenomenon. Ear Nose Throat J 2000;79:836.
5. Havas TE, Priestley J, Lowinger DS. A management strategy for vocal process granulomas. Laryngoscope 1999;109:301–6.
6. Orloff LA, Goldman SN. Vocal fold granuloma: successful treatment with botulinum toxin. Otolaryngol Head Neck Surg 1999;121:410–3.
7. Ylitalo R, Lindestad PA. A retrospective study of contact granuloma. Laryngoscope 1999;109:433–6.
8. Batsakis JG, Fox JE. Supporting tissue neoplasms of the larynx. Surg Gynecol Obstet 1970;131:989–97.
9. Ferlito A, Milroy CM, Wenig BM, et al. Laryngeal paraganglioma versus atypical carcinoid tumor. Ann Otol Rhinol Laryngol 1995;104:78–83.
10. Ferlito A, Barnes L, Rinaldo A, et al. A review of neuroendocrine neoplasms of the larynx: update on diagnosis and treatment. J Laryngol Otol 1998;112: 827–34.
11. Benjamin B, Dalton C, Croxson G. Laryngoscopic diagnosis of laryngeal sarcoid. Ann Otol Rhinol Laryngol 1995;104:529–31.
12. Wenig BM. Non-neoplastic lesions of the nasal cavity and paranasal sinuses. In: Atlas of head and neck pathology. Philadelphia: WB Saunders; 1993.
13. Cordier JF, Valeyre D, Guillevin L, et al. Pulmonary Wegener's granulomatosis. A clinical and imaging study of 77 cases. Chest 1990;97:906–12.

14. Daum TE, Specks U, Colby TV, et al. Tracheobronchial involvement in Wegener's granulomatosis. Am J Respir Crit Care Med 1995;151:522–6.
15. O'Halloran LR, Lusk RP. Amyloidosis of the larynx in a child. Ann Otol Rhinol Laryngol 1994;103: 590–4.
16. Lewis JE, Olsen KD, Kurtin PJ, Kyle RA. Laryngeal amyloidosis: a clinicopathologic and immunohistochemical review. Otolaryngol Head Neck Surg 1992;106:372–7.
17. Zundel RS, Pyle GM, Voytovich M. Head and neck manifestations of amyloidosis. Otolaryngol Head Neck Surg 1999;120:553–7.
18. Woo P, Mendelsohn J, Humphrey D. Rheumatoid nodules of the larynx. Otolaryngol Head Neck Surg 1995;113:147–50.
19. Lofgren RH, Montgomery WW. Incidence of laryngeal involvement in rheumatoid arthritis. N Engl J Med 1962;267:193–5.
20. Vrabec JT, Driscoll BP, Chaljub G. Cricoarytenoid joint effusion secondary to rheumatoid arthritis. Ann Otol Rhinol Laryngol 1997;106:976–8.
21. Sorensen WT, Moller-Andersen K, Behrendt N. Rheumatoid nodules of the larynx. J Laryngol Otol 1998;112:573–4.
22. Lee-Chiong TL Jr. Pulmonary manifestations of ankylosing spondylitis and relapsing polychondritis. Clin Chest Med 1998;19:747–57.
23. Spraggs PD, Tostevin PM, Howard DJ. Management of laryngotracheobronchial sequelae and complications of relapsing polychondritis. Laryngoscope 1997;107:936–41.
24. McAdam LP, O'Hanlan MA, Bluestone R, Pearson CM. Relapsing polychondritis: prospective study of 23 patients and a review of the literature. Medicine (Baltimore) 1976;55:193–215.
25. Prasad S, Grundfast KW, Lipnick R. Airway obstruction in an adolescent with relapsing polychondritis. Otolaryngol Head Neck Surg 1990; 103:113–6.
26. Davison FW. Acute laryngeal obstruction in children. A fifty-year review. Ann Otol Rhinol Laryngol 1978;87:606–13.
27. Thompson AB, Rennard SI. Diagnostic procedures not involving the pleura. In: Baum GL, Crapo JD, Celli BR, Karlinsky JB, editors. Textbook of pulmonary diseases. 6th ed. Philadelphia: Lippincott-Raven Publishers; 1998. p. 239–45.
28. Gannon TH, Eby TL. Angioedema from angiotensin converting enzyme inhibitors: a cause of upper airway obstruction. Laryngoscope 1990;100: 1156–60.
29. Seidman MD, Lewandowski CA, Sarpa JR, et al. Angioedema related to angiotensin-converting enzyme inhibitors. Otolaryngol Head Neck Surg 1990;102:727–31.
30. Maze A, Bloch E. Stridor in pediatric patients. Anesthesiology 1979;50:132–45.
31. Moore GP, Hurley WT, Pace SA. Hereditary angioedema. Ann Emerg Med 1988;17:1082–6.
32. Deutsch ES, Zalzal GH. Quincke's edema, revisited. Arch Otolaryngol Head Neck Surg 1991;117:100–2.
33. Borum ML, Howard DE. Hereditary angioedema. Complex symptoms can make diagnosis difficult. Postgrad Med 1998;103:251–5–6.
34. Bork K, Barnstedt SE. Treatment of 193 episodes of laryngeal edema with C1 inhibitor concentrate in patients with hereditary angioedema. Arch Intern Med 2001;161:714–8.
35. Darrow DH, Holinger LD. Inflammatory illness of the pediatric airway. In: Fried MP, editor. The larynx: a multidisciplinary approach. 2nd ed. St. Louis: CV Mosby; 1996. p. 143–53.
36. Pransky SM, Magit AE, Kearns DB, et al. Intralesional cidofovir for recurrent respiratory papillomatosis in children. Arch Otolaryngol Head Neck Surg 1999;125:1143–8.
37. Rimell FL, Shoemaker DL, Pou AM, et al. Pediatric respiratory papillomatosis: prognostic role of viral typing and cofactors. Laryngoscope 1997;107:915–8.
38. Ferlito A, Rinaldo A, Marioni G. Laryngeal malignant neoplasms in children and adolescents. Int J Pediatr Otorhinolaryngol 1999;49:1–14.
39. Koufman JA. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. Laryngoscope 1991;101 Suppl 3:1–78.
40. Ulualp SO, Toohill RJ. Laryngopharyngeal reflux: state of the art diagnosis and treatment. Otolaryngol Clin North Am 2000;33:785–802.
41. Bouchard S, Lallier M, Yazbeck S, Bensoussan A. The otolaryngologic manifestations of gastroesophageal reflux: when is a pH study indicated? J Pediatr Surg 1999;34:1053–6.
42. Hogan WJ, Shaker R. Supraesophageal complications of gastroesophageal reflux. Dis Mon 2000; 46:193-232.

43. Leopold DA. Laryngeal trauma. A historical comparison of treatment methods. Arch Otolaryngol 1983;109:106–12.
44. Weber AL. Radiology of the larynx. In: Fried MP, editor. The larynx: a multidisciplinary approach. 2nd ed. St. Louis: CV Mosby; 1996. p. 101–14.
45. Woodson GE. Configuration of the glottis in laryngeal paralysis. I: Clinical study. Laryngoscope 1993; 103:1227–34.
46. Woodson GE. Configuration of the glottis in laryngeal paralysis. II: animal experiments. Laryngoscope 1993;103:1235–41.
47. Hartnick CJ, Hartley BE, Miller C, Willging JP. Pediatric fiberoptic endoscopic evaluation of swallowing. Ann Otol Rhinol Laryngol 2000;109:996–9.
48. Sinard RJ, Netterville JL, Garrett CG, Ossoff RH. Cancer of the larynx. In: Myers EN, Suen JY, editors. Cancer of the head and neck. Philadelphia: WB Saunders; 1996. p. 381–417.
49. Deutsch ES, Smergel E, Crisci K, Panitch H. Tracheobronchography in children. Laryngoscope 1996;106:1248–54.
50. Vogl TJ. Hypopharynx, larynx, thyroid, and parathyroid. In: Stark DD, Bradley WG, editors. Magnetic resonance imaging. 2nd ed. St. Louis: CV Mosby; 1991. p. 1184–210.
51. Curtin HD. Larynx. In: Som PM, Curtin HD, editors. Head and neck imaging. 3rd ed. St. Louis: CV Mosby; 1996. p. 612–707.

CHAPTER 67

Bronchoesophagology

Jane Y. Yang, MD, Ellen S. Deutsch, MD, James S. Reilly, MD

BRONCHOLOGY

The trachea and bronchi usually function so well that their contribution to the respiratory process is taken for granted, but their critical importance in the function of the airway should not be underestimated. Heffner suggested that the compliant and noncylindrical structure of the trachea, comprising incomplete cartilage rings, allows for more than simple conduit function.[1] Enlargement of the lumen during inspiration may allow sufficiently decreased airflow velocity at the periphery of the air column so as not to impede the normal egress of the mucociliary blanket. In contrast, compression of the lumen at the beginning of a cough contributes to the increased driving pressure and increased airflow velocity that promote mucus expulsion; the abrupt noise of the cough results from suddenly increased airflow turbulence. Normal airflow in the trachea is laminar; masses or stenoses obstructing the lumen cause turbulent airflow evidenced as stridor. Adding low-density helium to oxygen can decrease turbulence and increase airflow in stenotic areas.[1]

Bronchoscopy

Bronchoscopy can be a diagnostic or therapeutic procedure, providing a direct view, tactile information, and the capability to manipulate the tracheal and bronchial lumina that cannot be obtained by any other method of investigation. Indications include obtaining tissue samples, placing treatment devices, removing foreign bodies, biopsy or removal of masses, and investigation of suspected pulmonary disease that does not respond to treatment as expected. Chest radiographs are usually obtained prior to endoscopy; they may narrow the differential diagnosis and direct the endoscopist's attention to a particular area.

Indications for Bronchoscopy

Hemoptysis The first step in hemoptysis management is control and protection of the airway. Although mild hemoptysis can be evaluated with fiber-optic telescopes, massive hemoptysis should be managed by rigid, open-tube endoscopy.[2] Options include using selective bronchial intubation or double-lumen endotracheal tube, allowing preferential ventilation of the uninvolved lung. Balloon tamponade can be accomplished using endoscopic guidance. Adjunctive methods include instillation of vasopressin, sodium bicarbonate, 1:20,000 concentrations of epinephrine, or cold saline lavages. Neodymium:yttrium-aluminum-garnet (Nd:YAG) lasers may be used to control bleeding from airway neoplasms.[3] Thorough evaluation is required for patients with hemoptysis as the source may be located anywhere from the nares to the alveolar space, or even within the esophagus. Tracheobronchial causes of hemoptysis include tracheobronchitis, tracheobronchial tear, tracheal erosion, foreign body, endobronchial tumors, bronchiectasis, tuberculosis, lung abscess, and fungoma.[1] In patients with tracheostomies or endotracheal tubes, bleeding may result from suction trauma or granulation tissue.

Chronic Cough Bronchoscopy is indicated for patients with chronic cough; yield is increased if there is a history of tobacco use, chest radiograph abnormality, hemoptysis, or localized wheezing.[4] Other causes of chronic cough include asthma, evaluated with pulmonary function testing including methacholine challenge, and laryngopharyngeal (extraesophageal) reflux, evaluated with 24-hour multichannel pH monitoring and laryngeal examination. Medications, such as angiotensin-convert-

ing enzyme inhibitors, can also cause chronic cough.

Stridor Stridor is a symptom, rather than a diagnosis, caused by turbulent airflow resulting from an abnormality of the larynx, trachea, or bronchi. Stridor can be high or low pitched; inspiratory, expiratory, or biphasic; wet or dry. Lesions of the extrathoracic portion of the airway tend to cause inspiratory noises, and lesions of the intrathoracic airways tend to cause expiratory sounds; severe lesions in any location can cause biphasic sounds. Some lesions involving the supraglottis or the vocal folds, such as vocal fold paralysis or laryngomalacia in infants, tend to cause high-pitched noises, whereas laryngomalacia in children after infancy and other glottic lesions can cause low-pitched noises. Lesions in the trachea tend to cause low-pitched noises, and abnormalities of the distal, small airways cause high-pitched sounds. Wheezing is a classic example of a high-pitched, expiratory noise coming from the distal airways. Wet sounds suggest anatomic or neuromuscular difficulties in clearing secretions.

Lesions affecting the trachea and bronchi may be within the lumen, intrinsic to the walls, or extrinsic. Intraluminal lesions include strictures, neoplasms, infectious masses, granulation tissue, or foreign bodies. Intrinsic lesions include stenotic complete tracheal rings. Extrinsic compression may result from lesions within the mediastinum or esophagus, such as substernal goiter, aortic aneurysm, benign or malignant mediastinal tumors,[2] vascular rings and slings or an anomalous innominate artery, and esophageal foreign bodies[5] (Figure 67–1). Tracheal obstruction resulting from aerophagia as a manifestation of tardive dyskinesia from neuroleptic medications has been reported.[6] In children, the anterior tracheal wall may be depressed from prolonged tracheostomy.[7] Dynamic causes of stridor, such as some manifestations of tracheomalacia in children, are generally examined with fiber-optic bronchoscopy; overdiagnosis of normal respiratory dynamics should be avoided.[8]

It may be difficult to determine the origin of the stridor prior to endoscopy, so evaluation of the larynx, discussed in other chapters, is often indicated prior to tracheobronchoscopy. The suspected lesion, patient characteristics, and the circumstances of the endoscopy are considered when determining whether to proceed with fiber-optic and/or rigid tracheobronchoscopy. Airway control may necessitate rigid endoscopy, intubation, or tracheostomy; reasonable preparations should be made prior to instrumenting the airway.

Chest Trauma Bronchoscopy is indicated when a tracheal tear or bronchopleural fistula is suspected following blunt or penetrating chest trauma. Symptoms of tracheal or bronchial disruption include dyspnea, hemoptysis, subcutaneous and mediastinal emphysema, inspiratory stridor, hoarseness, cough-

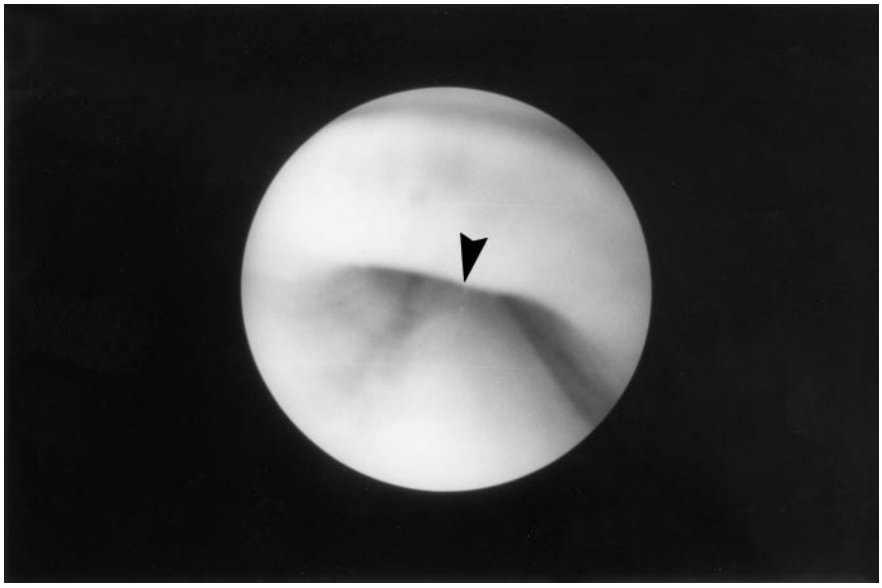


FIGURE 67–1. Endoscopic view: *arrow* shows anterior tracheal compression from an anomalous innominate artery.

ing, localized pain or tenderness, and, in severe cases, cyanosis; the most uniform early sign of tracheal disruption is subcutaneous emphysema.[9] Significant injury can occur with minimal evidence of trauma, such as from impact against a padded dashboard. Radiographic signs include pneumomediastinum, deep cervical emphysema, subcutaneous emphysema, pneumothorax, fractures limited to the upper rib cage, air surrounding a bronchus, and obstruction in the course of an air-filled bronchus.

Tracheobronchial trauma can occur at every level of the trachea and almost all of the major bronchi, but more than 80% of the injuries are within 2.5 cm of the carina.[9] The cervical trachea is protected by the mandible and sternum anteriorly and by the vertebrae posteriorly.[9] Tracheobronchial disruption has an estimated overall mortality of 30%[9]; 15% of patients sustaining cervical tracheal disruption die from hypoxia within 1 hour of injury.[9]

Bronchoscopy is the most reliable means of establishing the site, nature, and extent of bronchial disruption.[9] Thoracotomy and repair are indicated as soon as the patient's condition permits unless the bronchial tear involves less than one-third of the bronchial circumference and intercostal underwater drainage results in complete and persistent re-expansion of the lung with early cessation of air leakage.[9]

Inhalation Injury Thermal or chemical inhalation injury can produce significant edema and mucosal necrosis. Clinical findings may include facial, oropharyngeal, or nasal burns; carbonaceous sputum; wheezing; rales; rhonchi; and hoarseness. Arterial blood gas measurements may indicate hypoxemia, hypercapnia, and the presence of carboxyhemoglobin. Chest radiography is not usually helpful in the early stages of interstitial lung injury. Endoscopy is a valuable diagnostic tool and may reveal erythema, edema, ulceration, necrosis, or soot deposits; findings may be subtle in early airway injury.[10,11] If the patient's history suggests the possibility of thermal injury to the tracheobronchial tree, close observation is indicated and bronchoscopy should be considered, even in the absence of clinical findings.[10]

Respiratory Infections Bronchoscopy should be considered in adult patients who do not respond adequately to empiric antibiotics covering the most common respiratory pathogens, including *Streptococcus pneumoniae, Staphylococcus aureus, Haemophilus influenzae, Mycoplasma pneumoniae*, and *Legionella pneumophila*, as well as in critically ill or immunocompromised patients in whom atypical pathogens can cause fatal disease. Standard suction catheter aspiration may inappropriately sample microorganisms colonizing the upper airway, whereas bronchoscopy facilitates bronchial washings and broncheoalveolar lavage, allowing sterile, directed collection of specimens from distal bronchial segments. Lung abscesses can be drained endoscopically.[12] Transbronchial needle aspiration of interstitial infiltrates often provides the causative microorganism.

Abnormal Chest Radiograph Bronchoscopy performed for evaluation of patients with abnormal chest radiographs is more likely to be diagnostic when the radiograph demonstrates lobar collapse and hilar abnormalities and is less likely to be diagnostic in evaluation of nodular lesions and infiltrates.[13] Su et al found that persistent postobstructive consolidation was caused by an endobronchial mass in 65% of cases.[13]

Evaluation and Biopsy of Endobronchial Lesions Biopsy of endobronchial lesions can be accomplished using rigid or fiber-optic bronchoscopy; rigid bronchoscopy or transbronchial fine-needle aspiration may be more effective for tumors with necrotic centers. Several biopsy specimens should be obtained to improve the yield. Brushings and biopsies may be obtained for both visible and bronchoscopically invisible parenchymal lung lesions[14]; fluoroscopic guidance may be useful.[12] Endoscopy is also useful in evaluating potential complications of intubation, such as granulomas, webs, stenoses, or necrosis.[5]

Malignancy Staging and Follow-up In the United States, lung cancer is the most common malignancy causing death. Central lesions can be biopsied using a bronchoscope and forceps; bronchoalveolar lavage and cytologic brushings improve cytologic yield when performed after biopsy using forceps. The therapeutic response of endobronchial tumors can be monitored using bronchoscopic visualization and bronchoalveolar lavage in conjunction with high-resolution computed tomographic (CT) scans.

Transbronchial needle aspiration through a fiber-optic bronchoscope can be used to biopsy submucosal endobronchial lesions and peripheral nodules and masses,[14] as well as paratracheal, subcarinal, and perihilar areas for diagnosis and staging of thoracic malignancies, obviating the need for mediastinoscopy under general anesthesia. However, this technique is not widely used because of its low yield and technical difficulty.[15] Reported complications include pneumothorax and pneumomediastinum.[15]

Bronchoalveolar Lavage Bronchoalveolar lavage is used to evaluate pulmonary inflammation in diseases such as asthma, sarcoidosis, extraesophageal reflux disease, idiopathic pulmonary fibrosis, hypersensitivity pneumonitis, pulmonary alveolar proteinosis, hemosiderosis, histiocytosis X, and other interstitial lung diseases. Bronchoalveolar lavage is also used for identification of infectious lung diseases and confirmation of aspiration based on quantitative evaluation of lipid-laden macrophages.[14,16,17] Microbiologic, cytologic, or chemical studies may be performed on lavage fluids.[17]

The tip of the bronchoscope is wedged into a subsegment of the lung and multiple aliquots of 20 to 60 mL (in adults) of sterile saline solution are injected, aspirated, and collected separately.[18] In adults, a total volume of at least 100 mL is generally recommended to obtain adequate alveolar sampling.[19]

Bronchopleural Fistulae Bronchopleural fistulae are usually iatrogenic but may also result from tuberculosis, pneumonia, empyema, lung abscess, or trauma. If the site of the fistula is not readily visible, the specific location may be identified by serially inflating a balloon to occlude each bronchial segment until the source is identified or by injecting a small bolus of xenon 133 into the bronchial segments. Tracheobronchography is unique in its potential to localize fistulae.[20] Bronchoscopic application of sealants may be effective in closing small fistulae.[21]

Tracheoesophageal Fistulae Most tracheoesophageal fistulae are congenital, but acquired lesions, such as posterior erosion of the trachea, may result from intubation or tracheostomy, particularly if a foreign body, such as a nasogastric tube, resides simultaneously in the esophagus.[7]

Therapeutic Procedures

A variety of procedures can be performed endoscopically, avoiding the need for open surgical procedures. Endoscopically guided intubation is discussed in Chapter 65. Percutaneous tracheostomy is controversial because of concern over complications;[22] endoscopic guidance may reduce this risk.[23,24]

Stent Placement Endobronchial stents can be used to bypass areas of obstruction caused by lesions that may be malignant or benign and may be intrinsic or extrinsic to the lumen. In adults, stents are used as a temporizing measure to maintain airway patency while the patient is being treated for malignancy or to provide palliation for terminal disease. Expandable wire stents can be inserted with a fiber-optic bronchoscope and/or fluoroscopic guidance and are most commonly used for malignant lesions; extraction may be hazardous.[14,25] Rigid bronchoscopes are essential for the insertion, manipulation, and removal of silicone stents,[14] which can be used for benign conditions such as subglottic or tracheal stenosis, tracheobronchomalacia, and postanastomotic stenosis, as well as malignant lesions.[14] Complications of stents include migration of the stent, inspissation of mucus, ingrowth of granulation tissue into metallic stents,[14] erosion of the wall of the airway, and aortobronchial fistula, especially when inserted for treatment of airway obstruction secondary to compression by vascular structures.[26]

Laser Procedures Laser therapy can be used to resect or debulk both benign and malignant tumors that obstruct the airway or to prepare the airway for insertion of airway stents.[14] Carbon dioxide lasers allow precise destruction or excision of tissue, hemostasis of microcirculation, preservation of adjoining tissue, and minimal postoperative edema[27] but can only be applied using rigid instruments. Neodymium:yttrium-aluminum-garnet lasers[14] can be passed through fiber-optic bronchoscopes and applied using noncontact tips[28] but can cause submucosal damage, which progresses and does not become fully evident until after application is completed. Argon lasers are available with contact tips.[28]

Risks are related both to the anesthesia and the laser; anesthesia risks include hypoxemia, hypercarbia, and inadequate airway control. Risks of lasering include laser damage to unprotected areas of the operative field, patient, or operating room person-

nel; airway burns; severe hemorrhage; and perforation causing pneumothorax or pneumomediastinum.[14] Both oxygen and nitrous oxide support combustion, so the primary gas for anesthetic maintenance should consist of blended oxygen and air. The inspired oxygen content should be reduced to 30% or as close to this level as possible. Matte-finished instruments are used to minimize reflection.[27] The patient's eyes and face are protected with moistened towels, and operating room personnel must wear protective goggles appropriate for the type of laser being used.[27]

Airway burns can range from minor reflected injury of adjacent mucosa to transluminal destruction of the trachea. Ignition of the endotracheal tube can produce local thermal burns of the trachea and diffuse injuries to the lower respiratory tract from the inhalation of toxic materials, especially if polyvinyl chloride tubes are used.[27] Treatment includes immediate extubation, paralysis of the patient, and reintubation with a small endotracheal tube. Laryngoscopy and bronchoscopy are performed to evaluate the extent of damage and to remove charred debris and foreign material. Saline lavage, systemic corticosteroids and antibiotics, and careful monitoring of pulmonary function are indicated.[27] Perforation into the esophagus, mediastinum, pleural space, or major pulmonary vessels can cause severe hemorrhage, pneumothorax, or pneumomediastinum.[14,29]

Foreign-Body Removal Foreign-body aspiration is more common in young children than adults since they explore the world with their hands and mouths and have incomplete dentition, limited oromotor control, and immature judgment. The most common type of foreign body aspirated by children is food, particularly nuts.[5] Adults may aspirate when alcohol, sedatives, or head trauma alter their judgment or mental status, with cervicofacial trauma, or when neurogenic disease or physical conditions, such as dentures, impair sensation or control of the food bolus.

Patients who aspirate foreign bodies may present with gagging, coughing, and choking; however, over time, the symptoms may become quiescent. Children or impaired adults may not be able to recall or report the event, and the event may not have been witnessed. Endoscopy is indicated if there is a suggestive history despite a lack of radiographic findings, a suggestive radiograph despite a lack of supportive history, or pulmonary disease that follows an atypical course. Occasionally, a foreign body may be visible in radiographic images (Figure 67–2); sometimes only the sequelae are demonstrated. Jackson described a check-valve mechanism of airflow past a foreign body, resulting in obstructive emphysema, and a stop-valve mechanism, resulting in atelectasis (Figures 67–3 and 67–4).[30] Lack of deflation of a lung or bronchus may be demonstrated by comparing inspiratory and expiratory chest radiographs (Figure 67–5). For patients who are unable to cooperate or in whom the findings are subtle, lateral decubitus radiographs or airway fluoroscopy may be useful (Figure 67–6). In both children and adults, a high index of suspicion must be entertained to avoid missing an aspirated foreign body.[5] Delay in diagnosis can lead to hypoxemia, pneumothorax, asphyxiation, obstructive pneumonitis, lung abscess,

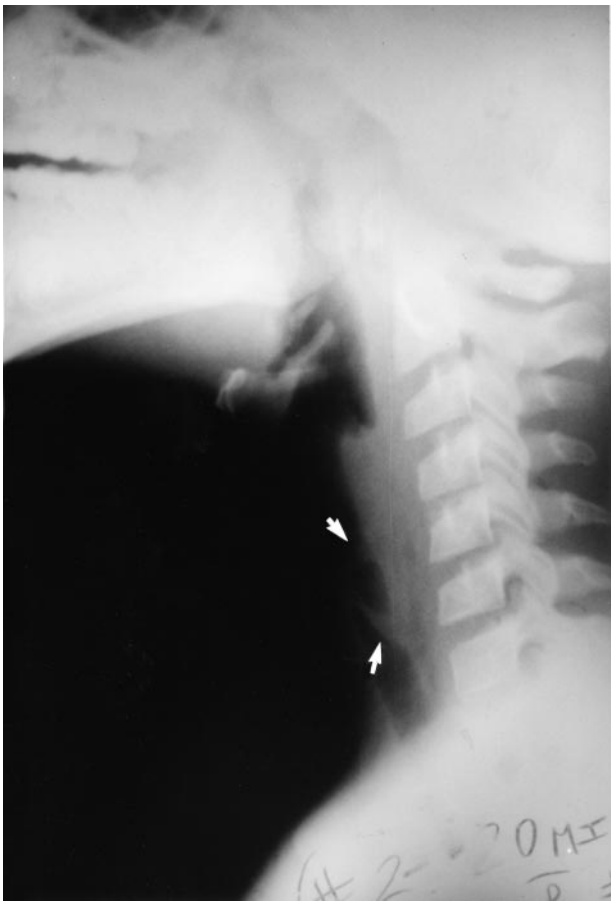


FIGURE 67–2. Radiograph of an adolescent with a chewed pen cap in his mid-trachea (*arrows*).

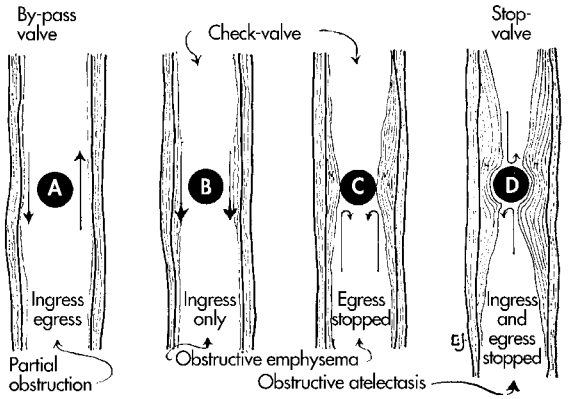


FIGURE 67–3. Three types of bronchial obstruction encountered clinically. Reproduced with permission from Jackson C. Bronchoscopy and esophagoscopy. A manual of peroral endoscopy and laryngeal surgery. 2nd ed. Philadelphia: WB Saunders; 1927.

bronchiectasis, pulmonary torsion, granuloma, and polyp formation.[31]

Although fiber-optic bronchoscopic retrieval of foreign bodies has been performed successfully,[31] rigid bronchoscopy is preferred in most cases, especially in the pediatric population. Rigid bronchoscopes with Hopkins rod telescopes offer a larger instrument internal diameter to allow easier retrieval of larger foreign bodies, superior optics, a wider variety of removal instruments, and the capability to control the airway.

To remove airway foreign bodies, optical or other forceps can be advanced just proximal to the foreign body, then the jaws are opened to distract adjacent mucosa, and the forceps are advanced beyond the widest portion of the foreign body or any portion that protrudes sufficiently to provide adequate purchase. Airway foreign bodies must be grasped in a secure manner and controlled during their removal. "Stripping off" a foreign body during its passage through the trachea or in the larynx may convert partial airway obstruction to total obstruction with an inability to ventilate the patient.[2] In children, this risk is increased in the subglottis because of its intrinsic relative narrowness. If control over a foreign body is lost within the trachea or subglottis, the foreign body can be pushed into a main bronchus, allowing ventilation of at least one lung. Visualization of a distal foreign body may be enhanced by having the anesthesiologist apply positive-pressure ventilation, similar to a Valsalva's maneuver. Attempts to remove a foreign body with a sharp end may cause additional trauma: "Advancing points perforate, trailing points do not" (Figure 67–7). Some foreign bodies can be rotated to decrease this risk; others can be sheathed by advancing the bronchoscope over them or be withdrawn into the lumen of the bronchoscope. For selected foreign bodies with a lumen, such as beads, a Fogarty catheter can be passed through the lumen of the object, the balloon inflated, and the catheter withdrawn so that the foreign body is controlled between

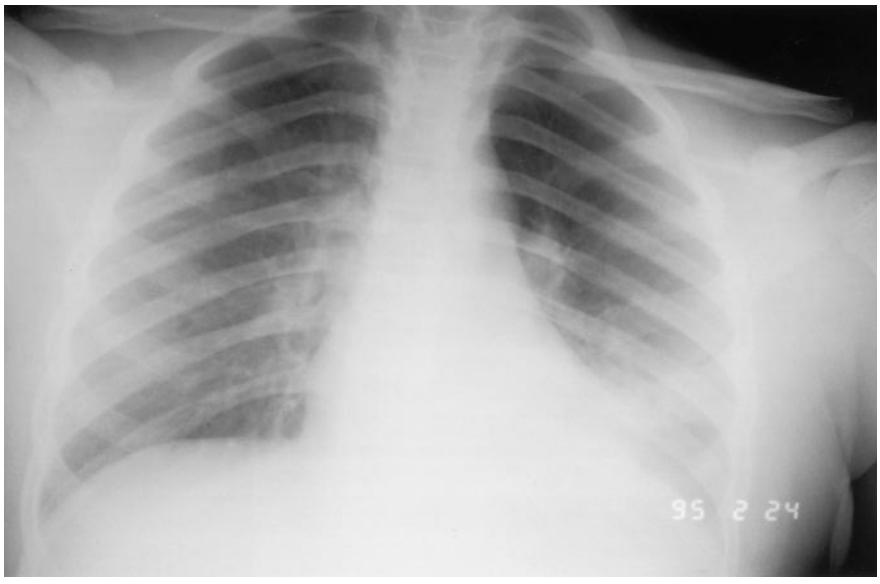


FIGURE 67–4. Chest radiograph demonstrating persistent left lower lobe atelectasis from an impacted sunflower seed.

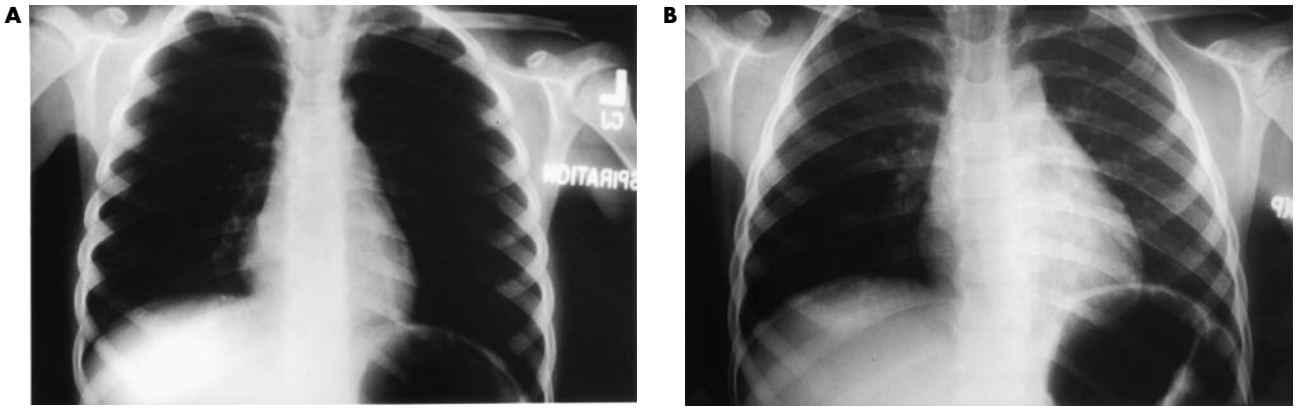


FIGURE 67–5. *A*, Inspiratory and *B*, expiratory radiographs demonstrating lack of deflation of the right lung in a child with a foreign body in the bronchus intermedius.

the balloon and the bronchoscope. Using a Fogarty catheter to dislodge a foreign body without controlling it is not recommended. Balloon breakage has been reported.[32] After removal of a foreign body, examination should be repeated in case there are additional foreign bodies; pulmonary toilet can be accomplished while the bronchoscope is still in situ.[5]

In unusual circumstances, tracheostomy may be necessary to control the airway,[33] and selected foreign bodies may be removed through a tracheostomy incision[33] or thoracotomy.

Removal of Mucus Plug and Pulmonary Toilet

Critically ill patients may not have the respiratory

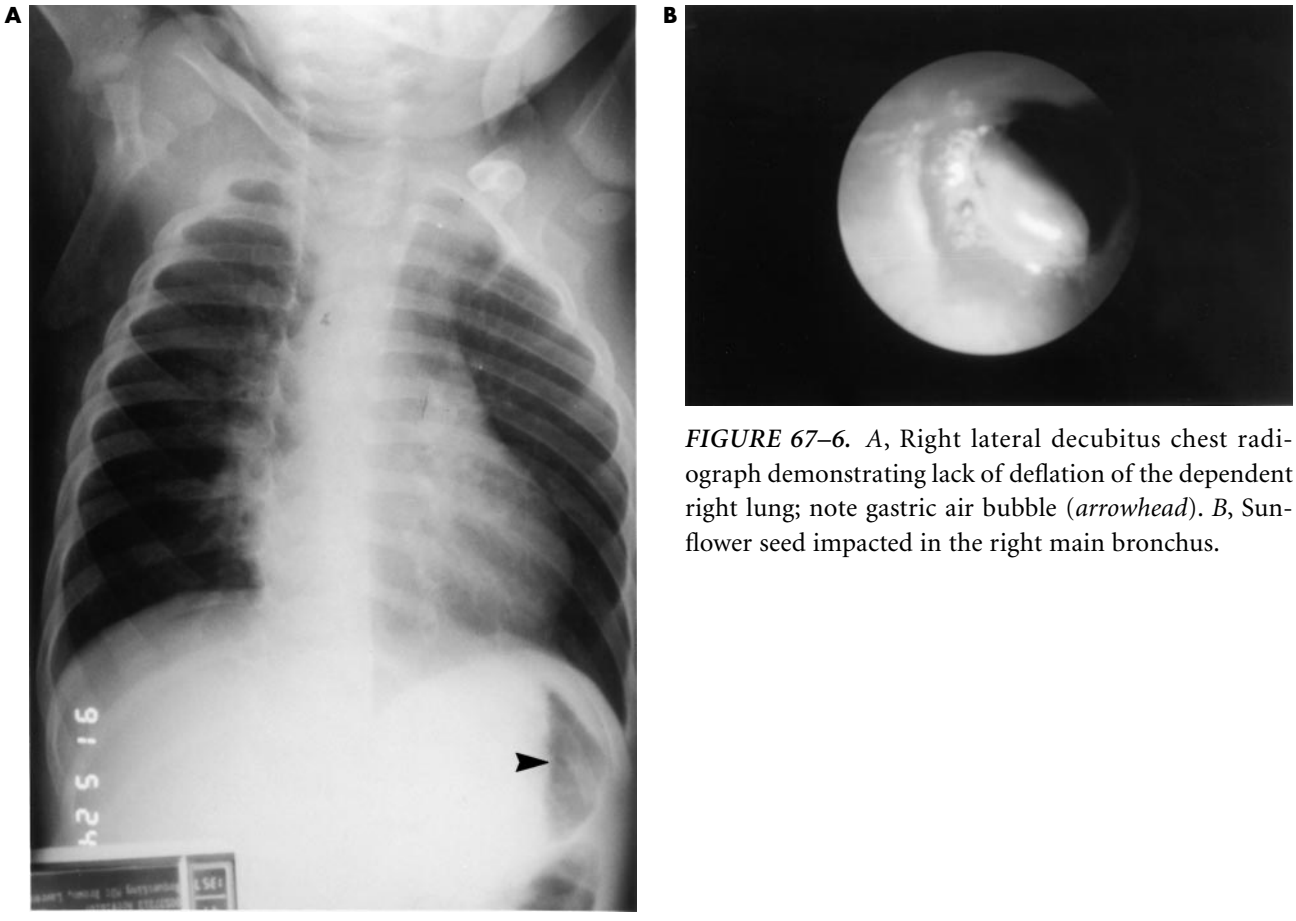


FIGURE 67–6. *A*, Right lateral decubitus chest radiograph demonstrating lack of deflation of the dependent right lung; note gastric air bubble (*arrowhead*). *B*, Sunflower seed impacted in the right main bronchus.

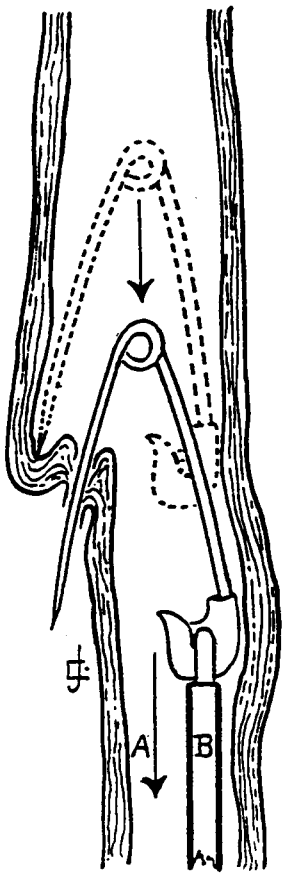


FIGURE 67–7. Schematic drawing of what will happen from ignoring the dictum "Advancing points perforate, trailing points do not." Reproduced with permission from Jackson C. Bronchoscopy and esophagoscopy. A manual of peroral endoscopy and laryngeal surgery. 2nd ed. Philadelphia: WB Saunders; 1927.

support and muscular strength to mobilize mucus plugs or other tenacious secretions interfering with ventilation. If aggressive respiratory therapy is inadequate or is contraindicated in patients with neuromuscular disease or head or spinal trauma,[31] bronchoscopy can be used to visualize, irrigate, and suction problematic secretions, improving pulmonary toilet. Furthermore, fiber-optic bronchoscopes can be used with fluoroscopic guidance to drain lung abscesses.[12]

Delivery of Medication, Radioactive Agents, or Other Therapies Surfactant and chemotherapeutic, antifungal, and antituberculous drugs have been administered directly into the lung on an investigational basis using bronchoscopic visualization.[31,34] Small bronchopleural fistulae can be closed by application of sealants using bronchoscopic guidance.[21]

Bronchoscopes can be used to guide the placement of brachytherapy catheters bearing radioactive source material into an obstructed airway.[14,35] Potential complications include hemorrhage and the formation of fistulae between the airway and other thoracic structures.[14]

Uncommon indications for bronchoscopy include application of hematoporphyrin derivatives or cryotherapy in the management of airway malignancies and bronchoscopic electrocautery to coagulate and vaporize endobronchial lesions.[14]

Tracheobronchial Conditions

Acute bronchiolitis is a viral disease of the lower respiratory tract that generally manifests in children under 2 years of age since inflammatory airway obstruction has a greater impact on young children with narrower airways. Respiratory syncytial virus (RSV) is responsible for more than 50% of cases. Following a viral prodrome, patients develop respiratory distress characterized by paroxysmal wheezy cough, dyspnea, and irritability with tachypnea, air hunger, and cyanosis. Treatment is generally supportive, including oxygen supplementation.[36]

Bacterial tracheitis is an acute infection of the upper airway that does not involve the epiglottis but can cause life-threatening sudden airway obstruction, particularly in children. Patients, generally less than 3 years of age, develop a brassy cough, high fever, and "toxicity" with gradually worsening inspiratory stridor. Thick, copious secretions complicate mucosal swelling at the level of the cricoid cartilage. Bacterial tracheitis may occur as a complication following a viral respiratory tract infection. The usual treatments for croup, a viral laryngotracheobronchitis, are ineffective, and intubation is often necessary; appropriate antimicrobial therapy is also indicated. *Staphylococcus aureus* is the most commonly isolated pathogen; *Moraxella catarrhalis* and *H. influenzae* can also cause bacterial tracheitis.[36]

Pneumonias can be divided into three groups: community acquired, nosocomial, and ventilator associated. Initial management includes appropriate empiric antibiotic therapy directed toward common pathogens based on the type of pneumonia and the patient's immune status. Cultures, obtained when patients have inadequate response despite standard

therapy, are often obtained from expectorated material, which may be contaminated by oropharyngeal flora. Bronchoscopy allows directed aspiration of secretions directly from the tracheobronchial tree and has proven to be of benefit in immunocompromised patients.[37] Proper disassembly and cleaning of equipment between patients and proper specimen collection techniques must be used to prevent contamination. Bronchoalveolar lavage and bronchoscopic protected specimen brushing can retrieve specimens adequate for quantitative analysis. Bronchoscopy is indicated for lung abscesses unresponsive to postural drainage and chest physiotherapy to rule out an underlying carcinoma or foreign body and to obtain secretions for culture.

Bronchiectasis, irreversible dilatation of the bronchial tree, most commonly presents with chronic purulent sputum production and hemoptysis. Bronchiectasis may be postinflammatory, obstructive, or congenital. Stasis of secretions leads to infections that damage the bronchial walls, resulting in further dilatation and distortion. Infectious causes include RSV, pertussis, rubella, and tuberculosis. Obstructive lesions include tumors, foreign bodies, extrinsic compression, and impacted mucus. Congenital causes include bronchial webs and atresia, immotile cilia syndrome, cystic fibrosis, and syndromes associated with abnormal cartilage formation such as Williams-Campbell syndrome (absence of annular bronchial cartilage distal to the first division of the bronchi) and Mounier-Kuhn's syndrome (congenital tracheobronchomegaly).[38] High-resolution CT or tracheobronchography are useful imaging modalities in suspected bronchiectasis.[20,38]

Mycetomas are formed by the conglomeration of fungal elements, most commonly *Aspergillus* species, within a pulmonary cavity. The diagnosis is made clinically, based primarily on chest radiographs demonstrating a solid, round, sometimes mobile, mass within a spherical or ovoid cavity; when the mass is peripheral, pleural thickening is often present. Serum precipitin to *Aspergillus* species has > 95% sensitivity for aspergilloma. Bronchoscopy is indicated for directed culture sampling if serum precipitin is negative or for evaluation of hemoptysis or progressive disease. Hemoptysis is the most common and potentially catastophic complication of mycetomas, causing 26% mortality in patients with aspergilloma. Bronchial artery embolization is only temporarily effective in managing acute hemoptysis. Surgical resection provides definitive treatment but is associated with high mortality and morbidity. Other treatment modalities, such as endobronchial instillation of antifungal agents, are under investigation.[39,40]

The incidence of *Mycobacterium tuberculosis*, a bacillus transmitted by airborne inhalation of infected droplets, has declined in recent years; an increasing proportion of new cases occurs among immigrants. Presenting pulmonary symptoms include productive or nonproductive cough, chest pain, and hemoptysis. Constitutional symptoms include chills, fever, night sweats, and loss of appetite and weight.[41] Diagnosis can be difficult but is supported by a reactive purified protein derivative skin test, chest radiograph with various abnormal findings, and identification of acid-fast bacillus by culture or histology. The sensitivity for three consecutive early morning sputums is disappointing at 50 to 80%. Bronchoscopy, with broncheoalveolar lavage and/or biopsy, has a high yield and should be considered in high-risk patients who have negative sputum cultures or patients who are unable to expectorate adequately.[41] Newer, more sensitive methods, such as polymerase chain reaction, are being investigated to detect mycobacterium in lavage fluid and sputum.[41,42]

Multidrug-resistant tuberculosis was a serious problem from 1993 to 1996, with the highest rates of resistance occurring in foreign-born, human immunodeficiency virus–positive patients and previously treated patients.[43] By instituting initial antituberculosis therapy with four first-line antibiotics, such as isoniazid (INH), rifampin (RMP), pyrazinamide (PZA), and ethambutol (EMB) or streptomycin, the incidence of new cases of multidrug resistance has declined significantly.[41,42]

Tuberculosis in children is usually contracted from adults and adolescents in the household rather than from other children in day care or school; congenital infection is rare. The presentation of primary pediatric tuberculosis may be subtle, including erythema nodosum and nonspecific constitutional symptoms. As tuberculin skin testing may be negative in 40% of these children, chest radiography demonstrating hilar and paratracheal lymphadenopathy is important in the diagnostic process. If young children cannot provide sputum specimens, three morning gastric aspirate specimens may be obtained. Children, especially those younger

than 6 years of age, have a more rapid progression of disease from inoculation to dissemination; they are treated immediately from the time of exposure until the incubation period has passed and repeat diagnostic tests are negative. The principal drugs used to treat pediatric tuberculosis include EMB, INH, PZA, and RMP; pyridoxine is given to infants to prevent neurologic complications of INH therapy. *Mycobacterium tuberculosis* in children is rarely contagious.[42]

Nontuberculous mycobacterial diseases encompass all *Mycobacterium* species other than *M. tuberculosis. Mycobacterium avium-intracellulare* complex is the most common nontuberculous mycobacterial disease and is treated principally with clarithromycin, azithromycin, RMP, and EMP.[43] *Mycobacterium kansasii* is the most virulent of the nontuberculous mycobacterial pathogens[43] and is treated with INH, RMP, and EMB. Rarely, atypical mycobacteria may present as an endobronchial mass (Figure 67–8).

Cystic fibrosis, a multisystem disorder of children and adults characterized by maldigestion and chronic obstruction and infection of the airways, is the most common life-threatening genetic trait in Caucasians. Otolaryngologists are most often involved for management of nasal polyposis resulting from chronic sinusitis. Exocrine gland dysfunction manifests as a paucity of water in mucous secretions and difficulty clearing mucous secretions, contributing to respiratory tract infection, particularly with *S. aureus* and *Pseudomonas aeruginosa*. Diagnosis is confirmed with high levels of chloride in sweat. Tracheobronchial suctioning or lavage may be temporarily helpful in treating atelectasis or mucoid impaction.[44] Biopsy samples for evaluation of cystic fibrosis or other ciliary dysmotility disorders can be obtained from the carina with a cup forceps or by brushing; specimens are placed in glutaraldehyde for electron microscopy.

Interstitial lung disease encompasses a wide variety of pulmonary diseases characterized by diffuse parenchymal opacities. Although more than 160 causes have been reported, pneumoconiosis, drug-induced disease, and hypersensitivity pneumonitis account for over 80% of cases (Table 67–1). A thorough history can elucidate patient exposure to a large variety of injurious inorganic dusts such as coal, carbon black, asbestos, or talc; chemicals such as polyvinyl chloride, sulfur dioxide, or ammonium; pharmacologic agents such as cyclophosphamide, methotrexate, certain anticonvulsants, and beta-blocking agents, etc; and radiation therapy. Open or thoracoscopic lung biopsy, considered the gold standard for diagnosis, is not always a viable option in the elderly or patients whose respiratory status is significantly compromised; less invasive bronchoscopy with bronchoalveolar lavage may provide useful information.[34,45,46]

Sarcoidosis is a non-necrotizing granulomatous disease of unknown etiology, more common in African Americans. Patients most commonly present with nonproductive cough and dyspnea; head

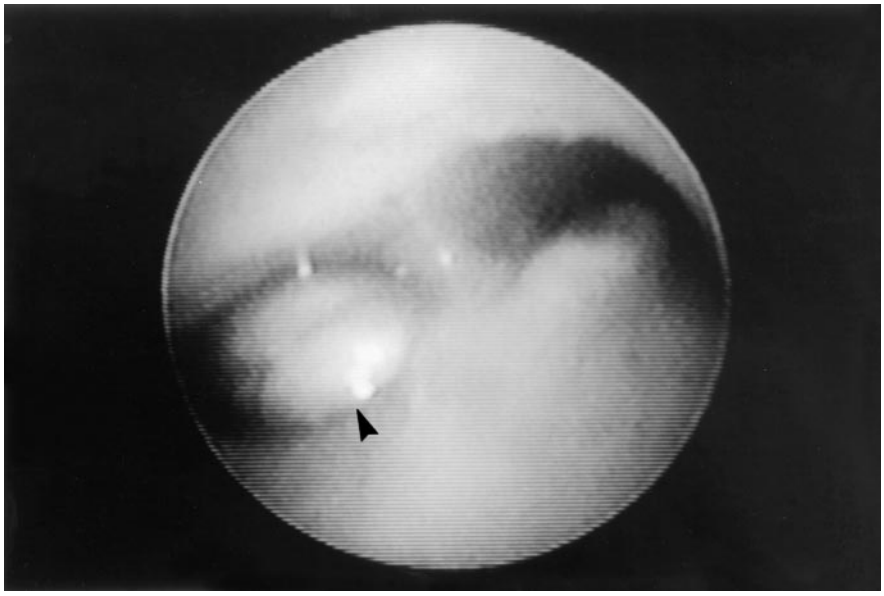


FIGURE 67–8. Endoscopic view of atypical mycobacteria presenting in a toddler as an obstructing mass in the left main bronchus (*arrowhead*); a smaller lesion is seen in the right main bronchus.

TABLE 67–1. Interstitial Lung Disease

Pneumoconiosis
Drug-induced pulmonary injury
Hypersensitivity pneumonitis
Sarcoidosis
Idiopathic pulmonary fibrosis
Bronchiolitis obliterans
Histiocytosis X
Collagen vascular disease
Rheumatoid arthritis
Systemic lupus erythematosus
Polymyositis/dermatomyositis
Mixed connective tissue disorder
Progressive systemic sclerosis
Sjögren's syndrome
Ankylosing spondylitis
Granulomatous vasculitis
Wegener's granulomatosis
Churg-Strauss syndrome
Lymphomatoid granulomatosis
Chronic eosinophilic pneumonia
Goodpasture's syndrome
Pulmonary alveolar proteinosis
Oxygen toxicity
Radiation toxicity
Berylliosis
Asbestosis
Silicosis
Coal workers' pneumoconiosis

and neck manifestations include cervical adenopathy, uveitis or episcleritis, parotid swelling, nasal obstruction, neuropathies such as facial palsy and sudden sensorineural hearing loss, epiglottic swelling, and nasal mucosal edema. Ninety to 95% of patients with sarcoidosis have an abnormal finding on chest radiography, most commonly hilar adenopathy. Occasionally, endobronchial granulomas or stenoses occur.[47] Elevated serum angiotensin-converting enzyme levels appear to correlate with the activity of the disease but should not be used in isolation. Laboratory studies may also reveal elevated liver enzymes, particularly aspartate aminotransferase and alkaline phosphatase, elevated erythrocyte sedimentation rate, eosinophilia, hypercalcemia, and hypergammaglobulinemia. Fiberoptic bronchoscopy with transbronchial biopsy is the invasive procedure of choice for diagnosis; bronchoalveolar lavage is investigational. Corticosteroids are the mainstay of therapy.[48] Pediatric sarcoidosis is rare and presents with fatigue and lethargy.[49]

Idiopathic pulmonary fibrosis is a chronic fibrosing interstitial pneumonia of unknown etiology associated with the histologic appearance of "usual" interstitial pneumonia. This diffuse parenchymal disease occurs almost exclusively in adults, usually over 50 years of age, who present with slowly progressive dyspnea and nonproductive cough. Rales, particularly at the lung bases, are noted on auscultation in 80% of patients; fever is rare, and the disease is limited to the lungs. Characteristic abnormal findings on chest radiograph include asymmetric, bilateral, peripheral areas of reticular opacification. High-resolution CT is not diagnostic but can help determine prognosis; a fibrotic appearance portends a worse prognosis than a predominantly ground-glass appearance. Diagnosis is usually presumptive, based on clinical criteria; bronchoscopy and laboratory evaluation may be indicated to exclude other pulmonary diseases. Open or thoracoscopic lung biopsy is generally obtained to establish a histologic diagnosis; bronchoalveolar lavage is investigational. Although corticosteroids are standard therapy, no clear evidence exists proving that corticosteroids or any other available treatment is efficacious.[50,51]

Relapsing polychondritis (discussed in Chapter 66) manifests with acute, recurrent, progressive inflammation and degeneration of cartilage and connective tissue, including that within the tracheobronchial tree, affecting men and women in equal numbers. Serious airway manifestations occur in about half of patients with relapsing polychondritis; bronchoscopy is useful to identify and quantify inflammation, stenosis, or dynamic collapse of the tracheobronchial tree. Bronchoscopic stent placement may be required to maintain airway patency.

Wegener's granulomatosis (discussed in Chapter 66) is a necrotizing granulomatous vasculitis,

affecting both the upper and lower respiratory tracts and the kidneys. The introduction of fiber-optic bronchoscopes has revealed tracheobronchial involvement in more than half of patients with Wegener's granulomatosis.[52,53]

Benign Neoplasms and Tumorlike Masses Although some neoplasms occurring within the trachea and bronchi are histologically benign, they may still cause airway obstruction. Recurrent respiratory papilloma has a predilection for the larynx, but the trachea and bronchi may be involved by disseminated disease (Figure 67–9). Traumatic granulomas may occur at sites of repeated mucosal trauma, such as the carina or main bronchi in patients with endotracheal or tracheostomy tubes undergoing repeated mechanical suctioning. Granulation tissue can also develop within the tracheal lumen at the superior margin of a tracheostoma; initially, the tissue is soft and friable; over time, it may become fibrotic.

In patients with *tracheopathia osteochondroplastica*, multiple submucosal nodules, consisting of cartilage and lamellar bone, can be seen projecting into the lumen of the tracheobronchial tree. Right middle lobe collapse is a common finding. The differential diagnosis of multiple nodular lesions of the tracheobronchial tree include papillomatosis, amyloidosis, and sarcoidosis.[54] If the lesions cause airway obstruction, they can be excised bronchoscopically.

Other reported benign lesions of the trachea or bronchi include inflammatory pseudotumors, plasma cell granulomas, fibrous histiocytomas, fibrolipomas, histiocytosis X, hamartomas, intratracheal ectopic thyroid tissue, pleomorphic adenomas, fibromas, fibrous histiocytomas, hemangiomas, hemangiopericytomas, paragangliomas, peripheral nerve sheath tumors, granular cell tumors, and leiomyomas.[1,14]

Malignant Neoplasms *Bronchogenic carcinoma*, often referred to as "lung cancer," is the most common malignancy in the United States. Long-term tobacco use is the single greatest risk factor for developing lung cancer; approximately 87% of all cases of lung cancer are attributable to tobacco use. Additional environmental factors, particularly exposure to asbestos and radon, increase the risk of lung cancers in smokers. The risk of a nonsmoker developing lung cancer is 1% or less of that of a smoker.[55] The symptoms of lung cancer are nonspecific; the most common symptom is cough; other pulmonary symptoms include dyspnea, hemoptysis, chest pain, and manifestations of paraneoplastic syndromes. Chest radiograph may demonstrate lung consolidation, atelectasis, infiltrate, or a solid nodule; CT of the chest with contrast often provides better definition of the lesion and is useful for assessing disease recurrence, persistence, or response to treatment. Bronchoscopy has emerged as an integral tool for the diagnosis and staging of lung cancer and may obviate the need for open biopsy. Bronchoscopy provides direct visualization of central lesions and can be combined with bronchoalveolar lavage, brushings, or biopsy to increase the diagnostic yield. Transbronchial needle aspiration of mediastinal lymph nodes can be performed to stage disease.

Bronchogenic carcinomas are divided histologically into non–small cell cancers, including squa-

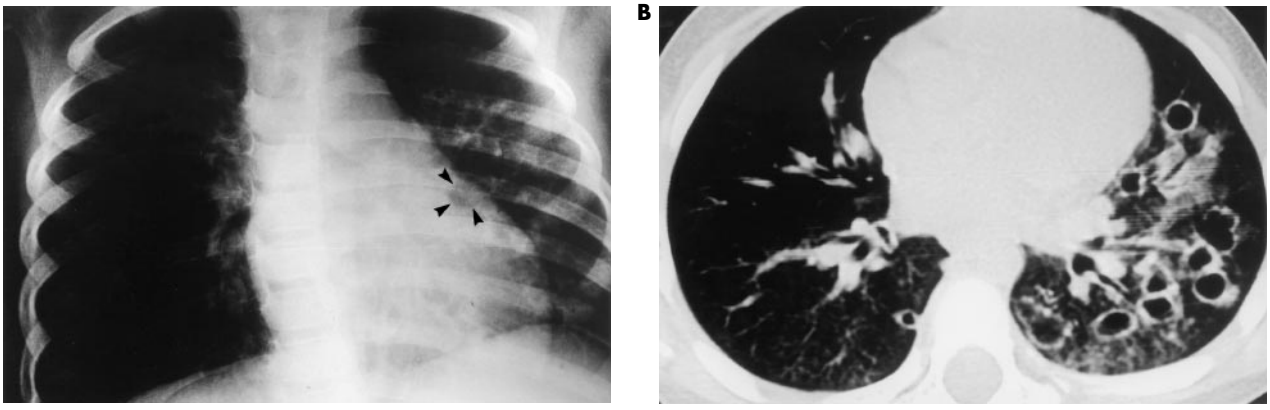


FIGURE 67–9. A, Chest radiograph and *B*, computed tomogram showing multiple cavitary pulmonary papillomas in a child with recurrent respiratory papillomatosis.

mous cell carcinoma, adenocarcinoma, and large cell carcinoma, and small cell cancers. Surgery is the primary treatment modality for non–small cell cancer; radiation therapy and chemotherapy are reserved for advanced cases and cancers not amenable to surgical resection. Small cell cancer is noted for rapid growth and early development of widespread metastases; although it is extremely sensitive to radiation and chemotherapy, 5-year survival is only 3 to 8%, and recurrence is common.[55]

Only approximately 4% of primary lung tumors are not bronchogenic carcinomas. Bronchial *carcinoid* is a neuroendocrine neoplasm comprising approximately 2% of primary lung tumors. This reddish, polypoid, endobronchial mass often presents with obstructive symptoms. As the tumors arise from Kulchitsky's cells, part of the amine precursor uptake, and decarboxylation (APUD) system, they may secrete hormones such as adrenocorticotrophic hormone, antidiuretic hormone, gastrin, somatostatin, calcitonin, and growth hormone. Carcinoid tumors are categorized as typical, which is relatively benign and is treated with conservative resection, or atypical, also referred to as neuroendocrine carcinoma, which is more aggressive and often has metastasized widely by the time of diagnosis. Aggressive local resection with lymph node dissection is recommended for locoregional disease; chemotherapy is indicated when distant metastases are present. Five-year survival is about 60% and is dependent on the histologic subtype.[56]

Other malignant neoplasms of the trachea or bronchi include spindle cell, oat cell, and adenoid cystic carcinomas; adenocarcinomas; mucoepidermoid carcinomas; malignant melanomas; sarcomas; lymphoreticular neoplasms; and malignancies invading the trachea or bronchi from adjacent structures (Figure 67–10).[1]

Other Tracheobronchial Disorders Congenital *tracheal stenoses* may result from complete tracheal rings or other cartilage deformities (Figure 67–11). Acquired stenoses may result from inhalation thermal or chemical burns or from intubation.[57] Pressure necrosis from an endotracheal or tracheostomy tube, or their attached cuffs, may result in healing by cicatrization, resulting in a spectrum of lesions.[7] Management of acute stenosis is controversial; some authorities advocate cautious dilatation,[7] and the application of mitomycin-C is under investigation.[58] The use of a laser to enlarge the airway lumen may ultimately cause more scarring and stricture formation,[7] although it may be indicated for highly selected lesions. A tracheostomy or T tube may be used for temporary or long-term management; definitive treatment involves surgical resection, expansion, and/or reconstruction.[7] Flow volume curves are of little practical use in the management of severe tracheal stenoses.[7]

Radiologic Evaluation

Specialized radiographic examinations, such as barium swallow, CT, and magnetic resonance imaging or magnetic resonance angiography, may be useful

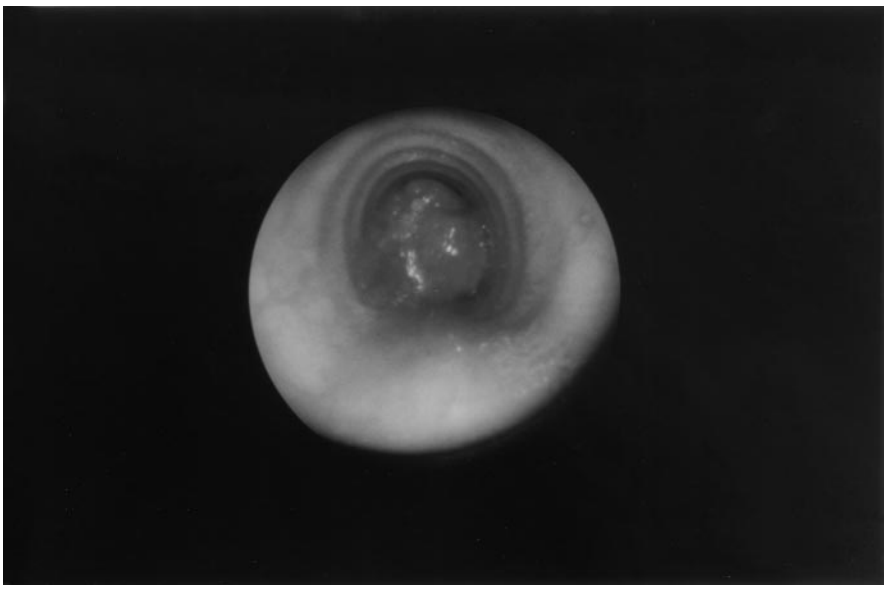


FIGURE 67–10. Endoscopic view of tracheal mucoepidermoid carcinoma in a 14-year-old child.

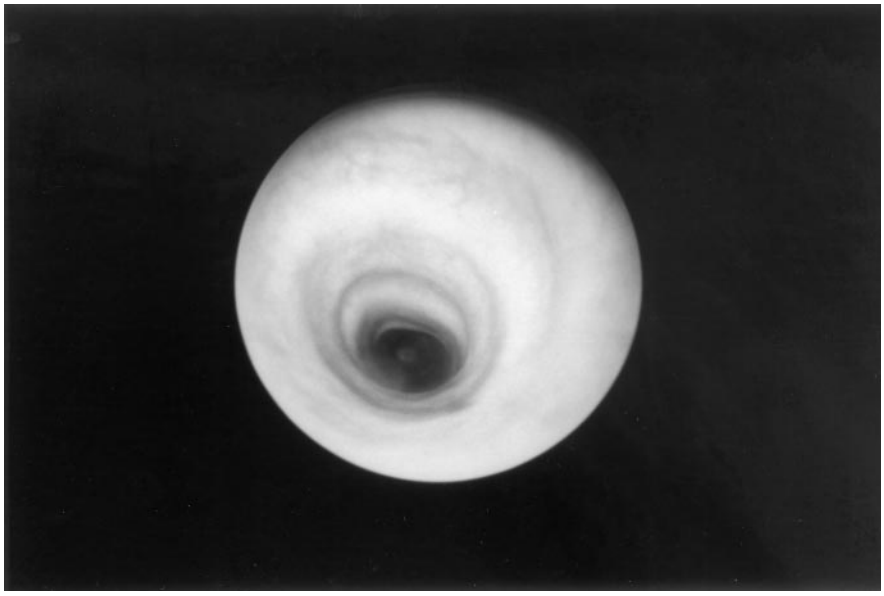


FIGURE 67–11. Complete tracheal rings in a child. Note the lack of a membranous component of the posterior tracheal wall. Photograph used with permission from Dr. Glenn Isaacson.

to define airflow dynamics, distal anatomy, or the relationship between the airway and adjacent structures or masses. Tracheobronchography is an elegant and well-tolerated procedure allowing superimposition of three-dimensional anatomy, providing a unified view of the airway rather than a planar slice image (Figure 67–12).[20] Radiographic tracheobronchial three-dimensional reconstruction and "virtual" endoscopy are investigational. Radiographic procedures can provide information about dynamic processes that is difficult to obtain by other methods but cannot provide tactile information or tissue samples.[14]

ESOPHAGOLOGY

Within the last few decades, improvements in lenses and illumination, as well as increased availability, have allowed esophagoscopes to evolve into primary diagnostic and therapeutic tools for managing many esophageal disorders. Open-tube esophagoscopes are used to examine and treat esophageal disease; flexible fiber-optic upper gastrointestinal endoscopes also allow examination of the stomach and duodenum.

Anatomy

The esophagus is a vertical muscular tube that extends from the hypopharynx to the stomach, measuring 23 to 25 cm in length in the adult. Beginning at the lower border of the cricoid cartilage, it passes through the neck, superior mediastinum, and posterior mediastinum anterior to the cervical and thoracic vertebrae, terminating at the cardiac orifice of the stomach.

The esophagus starts in the midline but then curves slightly to the left. It returns to the midline in the mediastinum and then deviates to the left as it passes anteriorly into the esophageal hiatus of the diaphragm. Importantly, the curvatures in the anteroposterior plane follow the convexity of the bodies of the thoracic vertebrae. The abdominal portion of the esophagus, which is only 1.25 cm long, is in the esophageal groove on the posterior surface of the left lobe of the liver.

In the superior mediastinum, the esophagus is posterior to the trachea and in contact with the common carotid arteries. The recurrent laryngeal nerves lie in the angle between the esophagus and the trachea, and the thoracic duct lies on its left side. Both lobes of the thyroid gland come in contact with the esophagus. The esophagus passes posterior and to the right of the aortic arch and then proceeds along the right side of the descending aorta until the inferior portion of the mediastinum is reached, where it passes anterior and slightly to the left of the aorta. Inferior to the arch of the aorta, the left bronchus crosses and indents the esophagus anteriorly. The thoracic duct is posterior to the lower portion of the esophagus, and the azygos vein is to the right side of the esophagus in the thorax. The right vagus nerve descends posterior to

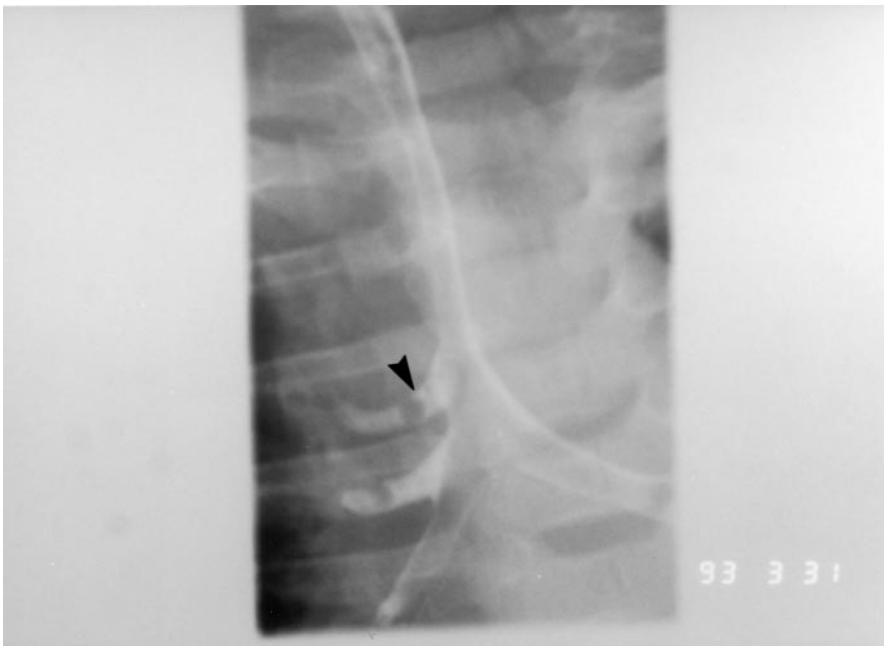


FIGURE 67–12. Tracheobronchogram demonstrating narrow tracheal caliber and duplicate right upper lobe bronchus (*arrowhead*) in a child with complete tracheal rings.

the esophagus, and the left vagus descends anterior to the esophagus.

The esophagus is lined by nonkeratinizing stratified squamous epithelium covering a thin lamina propria. The muscularis mucosae are composed of smooth muscle fibers arranged longitudinally, becoming thicker in the lower third of the esophagus. The submucosa consists primarily of thick collagenous and coarse elastic fibers and contains mucous glands and Meissner's plexus.

The muscle of the esophagus is composed of an inner circular layer, which is continuous with the inferior constrictor of the pharynx, and an outer longitudinal layer. The longitudinal fibers are arranged proximally into fascicles, which attach to the cricoid cartilage. Distally, the fascicles blend to form a uniform layer surrounding the esophagus. Auerbach's myenteric plexus lies between the two muscle layers. The muscle is striated in the upper third, mixed in the middle third, and nearly all smooth in the lower third. The outer coat, or fibrosa, consists of loose fibroelastic tissue rather than a strong serosa like that found in the gastrointestinal tract distal to the esophagus. The circular layer of muscle is slightly thicker at the level of the cricoid cartilage and blends with the cricopharyngeal muscle, which is attached to the cricoid cartilage. This upper esophageal sphincter is a high-pressure zone in esophageal manometry and corresponds with the cricopharyngeal muscle.[59] The lower esophageal sphincter is approximately 3 cm long and may descend 1 to 3 cm with normal respiration. A single distinct muscle responsible for the lower sphincter action has not been identified.

Radiologic Evaluation

Radiologic studies are generally obtained prior to endoscopy to evaluate esophageal structure and function, as well as diseases of adjacent organ systems that may affect the esophagus. Radiopaque esophageal foreign bodies can be demonstrated on posteroanterior or lateral chest and/or neck radiographs; nonradiopaque objects may be deduced from intraluminal air in the cervical esophagus. Lateral neck radiographs are taken with the patient's neck slightly flexed on the thorax, the head extended on the neck, and the shoulders positioned caudally and dorsally. Noncontrast radiographs may demonstrate swelling, air from a perforation, displacement of air-filled structures by neoplasm, cervical spine osteophytes that may cause dysphagia, or other abnormalities affecting the tissues between the cervical vertebrae and the larynx and trachea.

Radiographic contrast materials, such as barium sulfate suspensions of various consistencies, are used to outline the contours of the esophagus. Water-miscible solutions may be indicated if a fistula, per-

foration, or aspiration is suspected. Contrast studies may demonstrate congenital or acquired atresia, stenosis, and aerodigestive fistulae; extrinsic compression by aberrant vessels or cervical or mediastinal masses or foreign bodies[5]; and intraluminal abnormalities such as foreign bodies, varices, esophagitis, tumors, or ulceration. Three typical patterns are demonstrated in patients with congenital vascular anomalies. A large posterior esophageal indentation in the presence of anterior tracheal compression suggests a vascular ring consisting of a double aortic arch or the complex of a right aortic arch, left ductus arteriosus, and aberrant left subclavian artery (Figure 67–13). An oblique small retroesophageal indentation with a normal trachea is usually the result of an aberrant right subclavian artery with a left aortic arch, which is generally asymptomatic or, rarely, an aberrant left subclavian artery with a right aortic arch. The only vascular anomaly that passes between the trachea and the esophagus, causing an anterior esophageal indentation with a posterior tracheal indentation, is a "pulmonary sling," comprising an anomalous left pulmonary artery arising from the right pulmonary artery.[60]

Cinefluoroscopy is particularly valuable in demonstrating deglutition problems in the pharynx and motility and mucosal disturbances in the esophagus. A permanent record of the details, progress, and treatment response can be obtained in patients with achalasia, gastroesophageal reflux, esophageal dysmotility, esophageal spasm, Schatzki's rings, hernias, diverticula, and scleroderma.

Computed tomography and *magnetic resonance imaging* of the neck and chest have become a vital part of esophageal cancer staging, allowing distinction between intramural processes and extrinsic compression and demonstrating tumor extent and regional lymph adenopathy.

Endoscopic ultrasonography, a new, minimally invasive, innovative tool, using fiber-optic endoscopic visualization for cancer staging, provides information about intrinsic lesions, including mucosal and intramural involvement, and extrinsic masses. Ultrasonography can delineate the depth of invasion, identify possible metastatic lymph nodes, and guide fine-needle aspiration biopsies.[61,62]

Supplemental Evaluation

Twenty-four-hour multichannel pH probe studies are performed to evaluate gastric reflux. The distal pH sensor is placed 5 cm above the lower esophageal sphincter. The location of the proximal sensor has not been standardized, but some authorities advo-

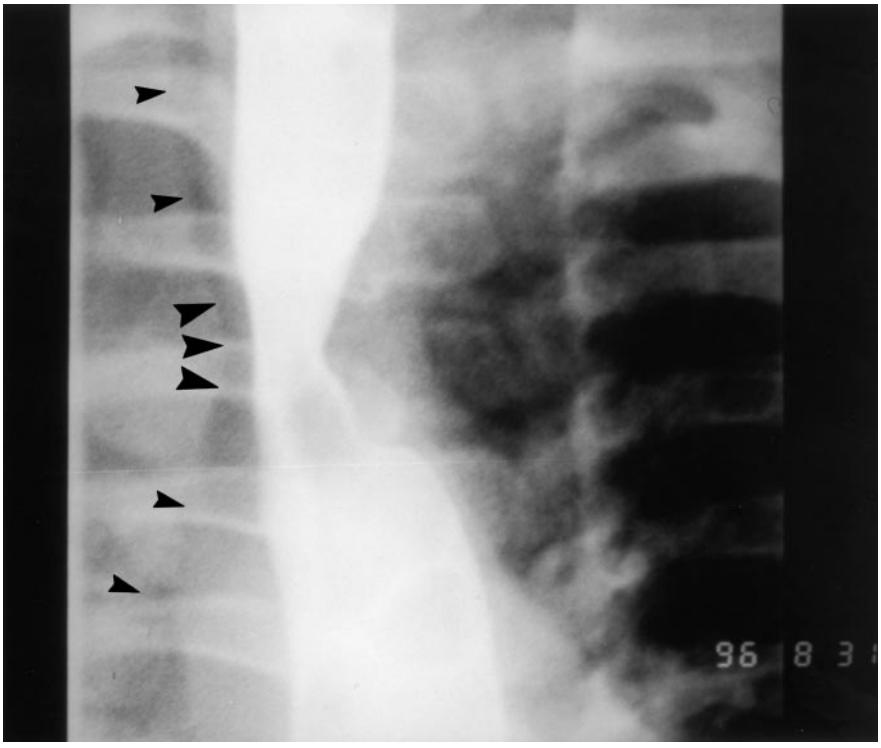


FIGURE 67–13. Barium esophagram demonstrating a double aortic arch, evidenced by a posterior filling defect in the esophagus, adjacent to a portion of the trachea that seems "tethered" to the esophagus (*larger arrowheads*). Note that the trachea widens and diverges from the esophagus both proximal and distal to the area constricted by the vascular ring.

cate placement 1 to 2 cm above the upper esophageal sphincter. Esophageal manometry of upper and lower esophageal sphincters assesses competence, spasm, or incomplete relaxation.

Indications for Esophagoscopy

Indications for esophagoscopy include complaints of dysphagia, odynophagia, regurgitation, and pyrosis; blunt trauma to the neck with associated subcutaneous air; penetrating neck trauma; radiographic evidence of esophageal masses or strictures; caustic or foreign-body ingestion; congenital anomalies; vocal fold palsies; suspicion of aerodigestive malignancies (primary, synchronous, or secondary); secondary esophageal involvement by autoimmune and idiopathic diseases; mechanical obstruction; and neuromuscular disorders. Additional interval evaluations may be indicated to monitor disease progression and treatment effects.

Esophageal *strictures* may result from caustic ingestion and as a manifestation of diseases such as Plummer-Vinson syndrome, Behçet's syndrome, gastroesophageal reflux, and Crohn's disease. Strictures may occur as a result of tracheoesophageal fistulae or malignancies or their management (Figure 67–14). Rigid or fiber-optic esophagoscopes can be used to visualize the esophageal lumen for direct dilatation or to pass a guidewire. If the esophagoscope cannot be passed through the stricture, the guidewire tip can be cautiously advanced and the stricture can be serially dilated over the guidewire; fluoroscopic guidance may be helpful. Jackson illustrated the potential for inadvertently perforating an acquired diverticulum proximal to a stricture (Figure 67–15).[30]

Stent placement using a guidewire and fiber-optic visualization may provide palliation for patients with nonoperable, obstructing esophageal cancer. Self-expanding metal stents are passed through the narrowed area of the tumor over a guidewire; once expanded, they cannot be removed.[63] To position a nonexpandable stent, the stenosis is first dilated with bougies or balloon dilators.

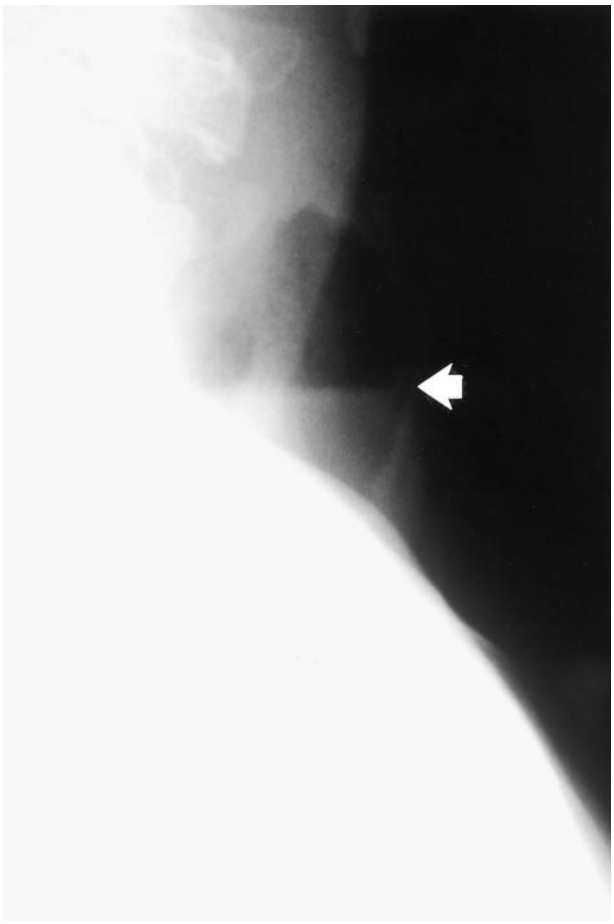


FIGURE 67–14. Esophageal stricture with esophageal air-fluid level (*arrow*). Photograph used with permission from Dr. Glenn Isaacson.

Therapeutic Esophagoscopy

Foreign Bodies Many foreign bodies pass uneventfully through the digestive system, but some become lodged in the esophagus and require surgical removal. In Western literature, meat is the most common esophageal foreign body found in adults and coins are the most common in children[64]; the ingestion is often witnessed.[65] Patients may be asymptomatic, or they may have dysphagia or emesis or develop stridor, fever, or a cough aggravated by eating.[5,65] Most children with esophageal foreign bodies are under 3 years of age, and most of the foreign bodies are coins[65] (Figures 67–16 and 67–17). Hawkins reported that 50% of children with ingested coins had them lodged for longer than 3 days, and every patient with coins lodged for longer than 7 days had respiratory symptoms, and some were febrile or had respiratory infections.[65] Lodgment occurred most commonly just below the cricopharyngeal muscle, as well as in the thoracic esophagus at the level of compression by the aortic arch or left main

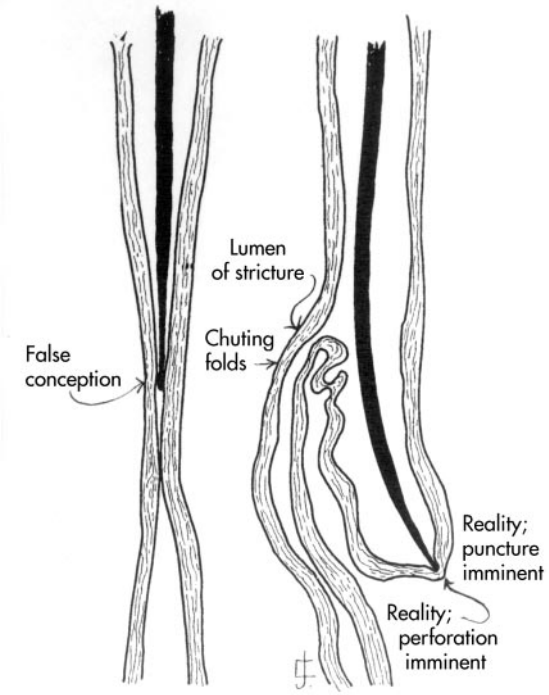


FIGURE 67–15. Mechanism of perforation by blind bougienage. On encountering resilient resistance, the operator, having a false conception, pushes on the bougie. Perforation results because, in reality, the bougie is in a pocket of a suprastrictural eccentric dilatation. Reproduced with permission from Jackson C. Bronchoscopy and esophagoscopy. A manual of peroral endoscopy and laryngeal surgery. 2nd ed. Philadelphia: WB Saunders; 1927.

bronchus or at a stricture.[65] Children and adults with atypical impactions may have predisposing conditions such as esophageal strictures or neuromuscular disturbances.[66]

Two-view neck and chest radiographs are often useful in identifying both the location and shape of a foreign body or its sequela, such as an esophageal air-fluid level. Barium is sometimes indicated to define a foreign body or to elucidate an underlying anatomic condition predisposing to foreign-body lodgment. Radiologic signs suggestive of perforation include retropharyngeal air, widening of the retropharyngeal soft tissue, leakage of contrast, or an extraluminal foreign body.[64]

Pharmacologic agents such as nifedipine have been used with varying degrees of success to facilitate the passage of an impacted foreign body by manipulating esophageal muscular tone;[67] glucagon does not appear to be effective in the dislodgment of esophageal coins in children.[68] Enzymatic digestion of a meat bolus with papain (Adolph's Meat Tenderizer) has largely been abandoned because of an unacceptable complication rate.[67] Emetic agents are ineffective and unsafe.[69]

In asymptomatic patients, a 24-hour observation period is safe; coins that reach the stomach are likely to pass through without complication and may be followed radiographically.[70] Button batteries, sharp objects, and objects causing bleeding, acute or severe airway compromise, pain, or dysphagia should be removed promptly.

Sharp objects have a perforation rate of 15 to 35%.[71] Even coins can cause stridor, esophageal erosion, aortoesophageal or tracheoesophageal fistula, mediastinitis, or paraesophageal abscess.[65,72] Other complications of esophageal foreign bodies include esophageal edema, laceration or erosion, hematoma,

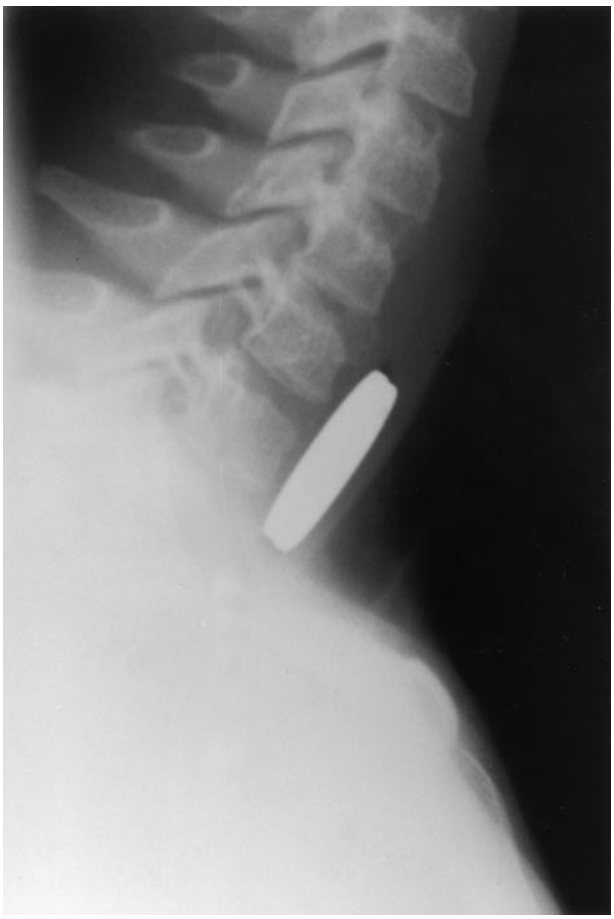


FIGURE 67–16. Lateral neck radiograph demonstrating two coins, back to back, in the esophagus.

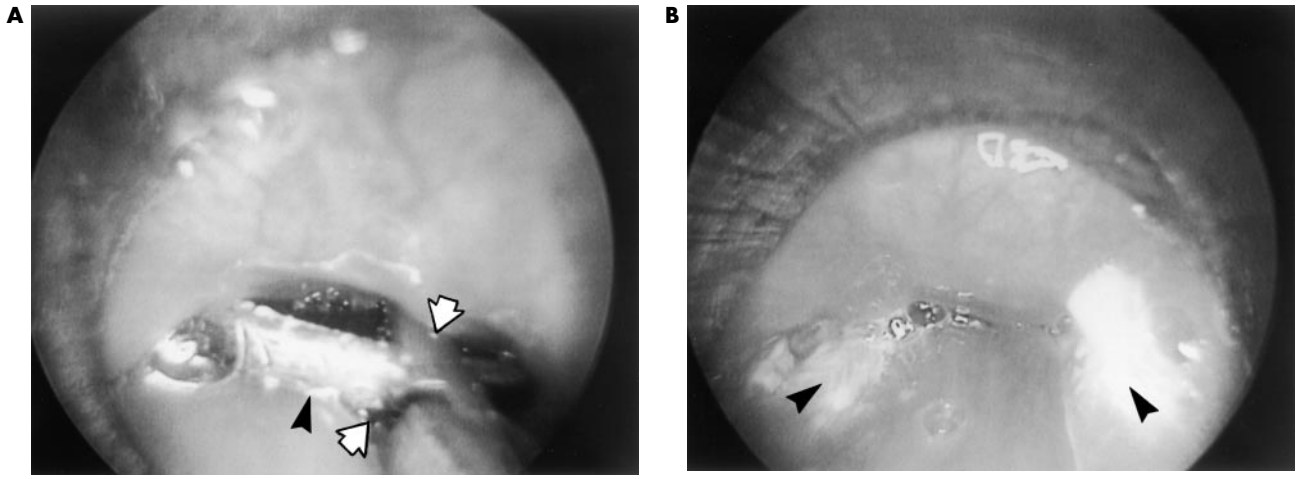


FIGURE 67–17. *A*, Coin (*black arrowhead*) grasped by side-channel alligator forceps (*white arrows*) for extraction. *B*, Minor eschar visible bilaterally after coin is removed. Note normal mucosa anteriorly, with fine vessels visible.

granulation tissue, retropharyngeal abscess, mediastinitis, migration of the foreign body into the fascial spaces of the neck, arterial-esophageal fistula with massive hemorrhage, respiratory problems, strictures, and proximal esophageal dilation; fatalities have been reported.[64,65,72]

Button batteries, commonly used in hearing aids, watches, electronic devices, toys, and even novelty clothing items, can cause significant injury to the esophagus as a result of direct current flow through tissue, local hydrolysis, and local hydroxide accumulation. The discharge state of the battery does not relate to patient outcome.[69] Patients who have ingested button cell batteries should undergo radiography to determine the location of the battery. Batteries lodged in the esophagus should be removed emergently as burns may occur as early as 4 hours after ingestion and perforation as soon as 6 hours after ingestion; fatalities have been reported.[69,73] Removal should be done endoscopically to allow direct visualization. Batteries that have passed beyond the esophagus need not be retrieved unless they remain lodged in the stomach for more than 48 hours or the patient manifests signs or symptoms of injury to the gastrointestinal tract,[69,74] The battery should be identified by size and imprint code or by evaluation of a duplicate, measurement of the battery compartment, or examination of the product, packaging, or instructions. If the chemical system cannot be identified from packaging or product instructions, it can be determined from the imprint code by calling the National Button Battery Ingestion Hotline at 202-625-3333. When mercuric oxide cells are ingested, blood and urine mercury levels are necessary only if the cell is observed to split in the gastrointestinal tract or radiopaque droplets are evident in the gut.[69]

Endoscopic removal of foreign bodies allows evaluation of any esophageal injury and visualization of multiple or radiolucent foreign bodies.[65,67] Although some foreign bodies can be safely removed using a fiber-optic esophagoscope, rigid esophagoscopes with Hopkins rod telescopes remain the gold standard for evaluation and removal of esophageal foreign bodies.

To grasp a foreign body during rigid esophagoscopy, the jaws of the forceps are opened after they have been inserted beyond the tip of the esophagoscope and may be used to push surrounding mucosa away from the foreign body. If the foreign body is too large to be withdrawn through the lumen of the esophagoscope, the esophagoscope is advanced to shield the foreign body. Sharp objects should be sheathed or rotated so that the point trails (see Figure 67–7). A blunt foreign body may be held tightly against the opening of the esophagoscope and the esophagoscope and foreign body withdrawn simultaneously in the plane of least resistance. Once proximal to the cricopharyngeus, the esophagoscope is tilted approximately 90 degrees to continue in an anterior direction through the pharynx and oral cavity. During the last phase of the maneuver, it may be more comfortable for the endoscopist to resume standing. Pediatric esophagoscopes offer the advan-

tageous capability of using side-channel instruments within the esophagoscope with the telescope simultaneously providing illumination and magnification.

Foreign bodies may escape endoscopic detection when the esophagoscope is advanced over a foreign body that is surrounded and obscured by granulation tissue; the foreign body may be seen while slowly withdrawing the esophagoscope. After foreign-body removal, it is reasonable to repeat the endoscopy to evaluate the possibility of mucosal erosions, additional foreign bodies, and underlying stricture or other lesion, especially if the foreign body was atypical.[65]

Alternative methods of removal, such as dislodging with a Foley catheter, are contraindicated if the foreign body has an unfavorable shape or if the patient has symptoms of airway involvement or any other complication.[65,75] Reported complications of Foley catheter removal include fatal airway obstruction, transient apnea, coin displacement to a main bronchus, esophageal perforation, esophageal tear, pneumomediastinum, bleeding, missed second coin, foreign body lost in the nasopharynx, aspiration pneumonia, and an inability to remove the coin.[65,75] Blind bougienage with Maloney dilators and insertion of nasogastric tubes to push the object into the stomach are occasionally used, but current opinion favors abandoning these methods as lacking safety and efficacy.[67] Other techniques include using suction or a protector hood secured to the tip of a fiber-optic esophagoscope or retrograde esophagoscopy from a gastrostomy.[71,72,76] Rarely, foreign bodies are so large or sharp that lateral pharyngotomy is necessary[77]; foreign bodies that have migrated into the soft tissues of the neck require open neck exploration.[64]

In comparing results, one must consider comparable foreign bodies, as well as whether complications result from the foreign body itself, from anesthesia or sedation, or from instrumentation. Berggreen et al compared rigid esophagoscopy to fiber-optic esophagoscopy with conscious sedation, Foley catheter technique under fluoroscopy (usually without sedation), blind bougienage with Maloney dilators, passage of a nasogastric tube to force the object into the stomach, and administration of intravenous glucagon or nifedipine. They reported that rigid esophagoscopy with the use of general anesthesia was more likely to be successful than all other techniques combined in both adults and children and that differences in complication rates were not statistically significant.[67]

Caustic Ingestion Patients who ingest caustic substances are at risk of serious esophageal injury, even in the absence of oropharyngeal burns.[74,78,79] Most ingestions occur in young children and are accidental; ingestions in adolescents and adults are often suicide gestures or attempts.[78,80,81] Initial treatment involves the immediate administration of water or milk to wash the corrosive material off the esophagus.[79] Attempts to remove the ingested material by emesis or lavage, adsorb it with activated charcoal, or neutralize it are contraindicated.[79]

Alkaline products account for the majority of ingestions and for the greatest number of serious injuries.[82] Liquid or solid drain cleaners and lye (sodium or potassium hydroxide); wall, oven, and toilet bowl cleaners (sodium hydroxide); laundry and automatic dishwasher detergents (sodium tripolyphosphate, silicates, and carbonates); Clinitest tablets (sodium hydroxide); and button batteries (potassium hydroxide) are common sources of corrosive injury. Acid ingestion is the second most common cause of severe burns. Hydrochloric acid, sulfuric acid, and swimming pool chemical ingestions can be dangerous; ammonia, chlorine bleach, and detergent ingestions, although common, are rarely responsible for significant esophageal morbidity.[82,83] Alkalis produce injury through liquefaction necrosis with penetration, whereas acids produce coagulation necrosis with a lesser degree of penetration. Liquids tend to cause wider and more severe injuries than tablets.[79]

Most authors advocate esophagoscopy to evaluate the extent and severity of injury for virtually all patients with caustic ingestions, finding that the presence or absence of symptoms is an unreliable predictor of esophageal injury.[79,80,84] Esophagoscopy is more accurate than any other means of evaluating esophageal involvement and is carried out under general endotracheal anesthesia within 24 to 48 hours of ingestion, to the upper limit of any full-thickness burn encountered.[74] First-degree (superficial) burns have mucosal hyperemia and edema or superficial mucosal desquamation; second-degree (transmucosal) burn injuries have mucosal sloughing with hemorrhages, exudates, and ulceration to and through the submucosa; and third-degree (transmural)

injuries have tissue sloughing with deep ulcerations or necrosis, including extension into the periesophageal tissues.[84] Esophagoscopy is contraindicated in the presence of a severe burn with evidence of laryngeal edema and in patients who have been on high doses of corticosteroids.[74]

In contrast, Crain et al found emesis, drooling, and stridor predictive of esophageal injury.[78] In 79 consecutive patients younger than 20 years of age, 50% (7/14) of patients with two or more of these three specific symptoms and signs had serious esophageal injury compared to normal endoscopic examination in all patients with fewer than two of these findings.[78]

Management remains controversial and may include hospitalization, antibiotics, corticosteroids, analgesics, sedation, and antireflux therapy. Corticosteroids are thought to be beneficial for moderately severe lye burns but are not indicated for severe lye and acid injury. The use of corticosteroids for third-degree burns is controversial as it may increase the likelihood of perforation or mask a perforation should one occur.[74] A prospective, randomized study of 60 children treated with and without corticosteroids following caustic ingestion with serious esophageal injury demonstrated that stricture formation was related only to the severity of the corrosive injury.[80] However, the finding of no benefit from corticosteroids may be related to the relatively low dosage of corticosteroids chosen.[74] Additional investigation and management options include interval barium esophagograms, the use of esophageal stents or the prolonged placement of a nasogastric tube, gastrostomy, anterograde or retrograde dilation of strictures, and esophageal replacement.[74,80] The long-term risk of cancer is increased in patients who require years of dilations.[74]

Selective esophagoscopy is reasonable for patients who ingest bleach or hair relaxers. Bleach ingestion does not usually result in stricture formation.[74] Hair relaxers, including "no-lye" preparations, have an alkaline pH; however, serious esophageal burns are uncommon, perhaps because of the solid nature of these products.[84] Emergency department or inpatient observation may be selectively used, with endoscopy reserved for those who do not tolerate oral intake, those with symptoms of complications, and those whose ingestion was intentional.[82,84,85]

Esophageal Diseases

Achalasia is a motor disorder of insidious onset and unknown etiology characterized by loss of peristalsis of esophageal smooth muscle and dysfunction of the lower esophageal sphincter, causing incomplete relaxation. In the early stages of achalasia, barium swallow may demonstrate spasm without dilation. The esophagus eventually becomes dilated and retains food; at this point, barium contrast may demonstrate esophageal dilatation, with a smooth tapered bird-beak appearance of the lower esophageal sphincter because of incomplete relaxation. Inhalation of amyl nitrite, a smooth muscle relaxant, results in sudden relaxation of the lower esophageal sphincter leading to rapid passage of the contrast through the sphincter.[86,87] Dysphagia is the most common symptom followed by regurgitation of food; other symptoms include aspiration, foul eructation, pyrosis, substernal pain, weight loss, and hemorrhage.[86,87] Histologically, the ganglion cells in Auerbach's plexus are significantly reduced in quantity or absent.

Endoscopy should be performed to assess the severity of retention esophagitis and to exclude malignancy. Achalasia is considered a premalignant condition; esophageal cancer eventually develops in 2 to 8% of patients.[63] Management options include dilating the lower esophageal sphincter or surgical myotomy; complications for both procedures include perforation and gastroesophageal reflux. Injection of botulinum toxin into the lower esophageal sphincter has shown promising results.[88]

Patients with *amyloidosis* often have neural or muscular involvement, resulting in abnormal or absent peristalsis. Lower esophageal sphincter pressures may be high, low, or normal, but relaxation is impaired. Symptoms can mimic achalasia.[89]

Pharyngoesophageal dysphagia, caused by cricopharyngeal spasm, a cricopharyngeal bar, cervical osteophytes, or cervical diverticula, may result in discomfort, weight loss, or aspiration. Diagnosis can be made with videoradiography and intraluminal manometry. Treatment options include dilatation, cricopharyngeal myotomy, and botulinum toxin injections.

Esophageal *diverticula* are categorized by their location and mode of development. Pharyngoesophageal ("Zenker's") and lower esophageal diverticula are thought to be related to disordered

motility of the cricopharyngeus and esophagus and to congenital or acquired weakness of the muscular walls; they do not contain a muscular layer and are classified as pulsion diverticula.[90] Traction diverticula, usually in the middle third of the esophagus, contain muscle and are thought to result from inflammatory processes adjacent to the esophagus, causing contracture and deformity of the entire wall. Traction diverticula are frequently asymptomatic incidental radiographic findings.

The cause of Zenker's diverticula remains controversial; however, cricopharyngeal muscle spasm is thought to contribute. Esophageal mucosa and minimal submucosal tissue herniate posteriorly through a weak area known as Killian's dehiscence or triangle, bound by the oblique layer of the inferior pharyngeal constrictor muscle, the transverse layer of the cricopharyngeus muscle, and the midline raphe. Patients report dysphagia, regurgitation of undigested food, aspiration pneumonia, and weight loss. Diagnosis is made radiographically by coating the diverticular sac with barium; treatment is surgical. Open cervical approaches consist of diverticulectomy or diverticulopexy with cricopharyngeal myotomy. Myotomy alone may be performed for small diverticula; diverticulopexy may be advantageous for debilitated patients because the sac is never violated so the patient may resume feeding quickly.[91] Feeley et al reported a 38% complication rate for diverticulectomy and myotomy, including wound infection, salivary fistula, and mediastinitis.[92] The "Dohlman procedure," endoscopic division of the esophagodiverticular wall incorporating division of the cricopharyngeus, was popularized by Dohlman and Mattsson in 1960 and was initially performed with electrocautery and then with a carbon dioxide laser.[90,93,94] Recently, endoscopic stapling has been advocated as a safer method, allowing division of the common septum while simultaneously sealing the transected edges. This minimally invasive, quick procedure facilitates a short hospital stay and early alimentation and is advantageous for elderly or debilitated patients.[95]

Inflammation, Esophagitis, and Gastroesophageal Reflux Disease An incompetent lower esophageal sphincter, increased intra-abdominal pressure, gastric outlet obstruction, dietary habits, or tobacco use may contribute to *gastroesophageal reflux*; the resultant contact between gastric acid and esophageal mucosa can cause chronic inflammation, strictures, webs, and dysmotility. The most common symptom is "heartburn"; other symptoms include chest pain, "backwash," or foul taste in the hypopharynx. Management may include dietary changes, weight loss, medications such as proton pump inhibitors and histamine H_2-receptor antagonists, or surgical procedures such as fundoplication.

Barrett's metaplasia comprises replacement of the normal esophageal stratified squamous epithelium with gastric-type columnar mucosa. This premalignant condition develops as a consequence of prolonged chronic gastroesophageal reflux. The diagnosis can be difficult to determine endoscopically because of the normal transition into gastric mucosa at the gastroesophageal junction. The presence of columnar epithelium 3 cm above the most proximal margin of the gastric folds is diagnostic and should be confirmed histologically. The incidence of adenocarcinoma in American patients with Barrett's esophagus is approximately 800 cases per 100,000 patients per year. Regular endoscopic surveillance for dysplasia and early carcinoma is indicated because of this 40-fold increase in risk.[96]

Infectious esophagitis, most commonly caused by herpes simplex, *Candida*, or cytomegalovirus, usually occurs in immunocompromised patients and diabetics. The primary symptoms are dysphagia and odynophagia. Endoscopy with biopsies, brushings, and culture may be necessary to establish the diagnosis. In bone marrow transplant patients with chronic *graft-versus-host disease*, an epithelial reaction of the esophageal mucosa can cause dysphagia and odynophagia.[97]

Connective Tissue Disorders The primary esophageal manifestations of connective tissue disorders are motility dysfunctions.

Scleroderma. Progressive systemic sclerosis, or scleroderma, is a progressive vasculitis and fibrosis involving multiple organs. Esophageal involvement, present in at least 70 to 80% of patients,[98] manifests with pyrosis, regurgitation, dysphagia, and odynophagia. The motor disorder manifestations include low-amplitude esophageal contractions that initially involve the smooth muscles of the lower two-thirds of the esophagus, with eventual progression to the striated muscle of the upper one-third of the esophagus. The weakened lower esophageal

sphincter, increased acid secretion, and delayed gastric emptying predispose these patients to esophageal reflux.[99]

Dermatomyositis. Polymyositis and *dermatomyositis* are inflammatory processes primarily affecting striated muscles. The proximal third of the esophagus, containing striated muscle, is the most commonly involved region of the gastrointestinal tract. Smooth muscle can also be affected, resulting in decreased pharyngeal and upper esophageal sphincter contraction pressures. Patients often present with nasal regurgitation and aspiration[98]; other common symptoms include dysphagia and esophageal regurgitation. Some authors believe that patients with these diseases are at increased risk of gastrointestinal malignancies.[98,99]

Other Connective Tissue Disorders. Sjögren's syndrome is characterized by the presence of two of the three following criteria: keratoconjunctivitis sicca, xerostomia, and chronic inflammatory connective tissue disease. Histologic evaluation demonstrates atrophy of the acinar structures, extensive lymphoplasmocytic infiltration, and sclerosis. Dysphagia is most commonly caused by xerostomia; however, rarely, esophageal dysmotility may occur.

Both *systemic lupus erythematosus* and *rheumatoid arthritis* may include esophageal dysmotility. *Mixed connective tissue disease* is a distinct clinical entity with features of systemic lupus erythematosus, progressive systemic sclerosis, myositis, and rheumatoid arthritis. Serologic studies may demonstrate an increase of antiribonucleoprotein antibodies. Seventy-five percent of patients have motor dysfunction of the esophagus; some also have pharyngeal involvement.[98]

Dermatologic Vesicobullous Diseases *Epidermolysis bullosa* comprises a group of hereditary disorders characterized by blister formation at sites of minor trauma. Bullae form on the skin or mucous membranes and result in ulcers that are slow to heal and ultimately form scars. Autosomal recessive dystrophic epidermolysis bullosa is most frequently associated with oral and esophageal involvement.[100] Oral blisters and dysphagia are the most common extracutaneous complaints.[101] Esophageal involvement often occurs in the cervical esophagus, probably from mechanical trauma, with dysphagia developing insidiously over the first decade of life.[101] Subsequent scarring and stricture formation are difficult to manage as conservative attempts to dilate the stricture may lead to additional trauma and bullae formation. Endoscopy is generally not recommended because of the possibility of causing mechanical trauma; evaluation by barium esophagram is preferred. A soft diet helps to decrease the incidence of mucosal trauma.[102]

Pemphigus vulgaris is a rare intraepidermal bullous disease marked histologically by suprabasal acantholysis. Oral lesions occur in 70% of patients; the oral cavity is the only site affected in more than 50%.[103] Esophageal involvement may cause dysphagia. Nikolsky's sign, separation of the epidermal layer from underlying tissue after applying lateral pressure to the skin, may be observed. Esophagoscopy is the diagnostic method of choice, and treatment is primarily corticosteroids.[103]

Benign mucous membrane (cicatricial) pemphigoid is a chronic, bullous disease of the elderly, causing webs and strictures most commonly involving the oral cavity and conjunctiva. Esophageal involvement is less common and may be evaluated with a barium esophagram and careful endoscopy; induction of bullous lesions owing to endoscopy has been reported.[104] Corticosteroids are the primary mode of therapy. Strictures may be treated with careful dilatation; however, the risk of perforation exists. Colon interposition has been successfully performed for severe strictures.[105]

Stevens-Johnson syndrome is a severe systemic hypersensitivity erythema multiforme reaction that presents with iris-like, target lesions and vesicobullous lesions of skin and mucous membranes. Although over 50% of cases of erythema multiforme are idiopathic, the most common causative agents in adults are medications such as sulfonamides, phenytoin, barbiturates, phenylbutazone, digitalis, and penicillin and, in children, infections with microorganisms such as *Streptococcus*, herpes simplex, and *Mycoplasma*. Extensive cutaneous and mucosal lesions are associated with conjunctivitis, fever, leukopenia, and bullous stomatitis; mucosal involvement can be painful enough to cause poor oral intake, and dehydration is a serious concern[106,107] (Figure 67–18). Treatment consists of removal of the underlying cause and supportive care. Corticosteroid therapy is often used in severe disease, but its efficacy is unproven.[106,107]

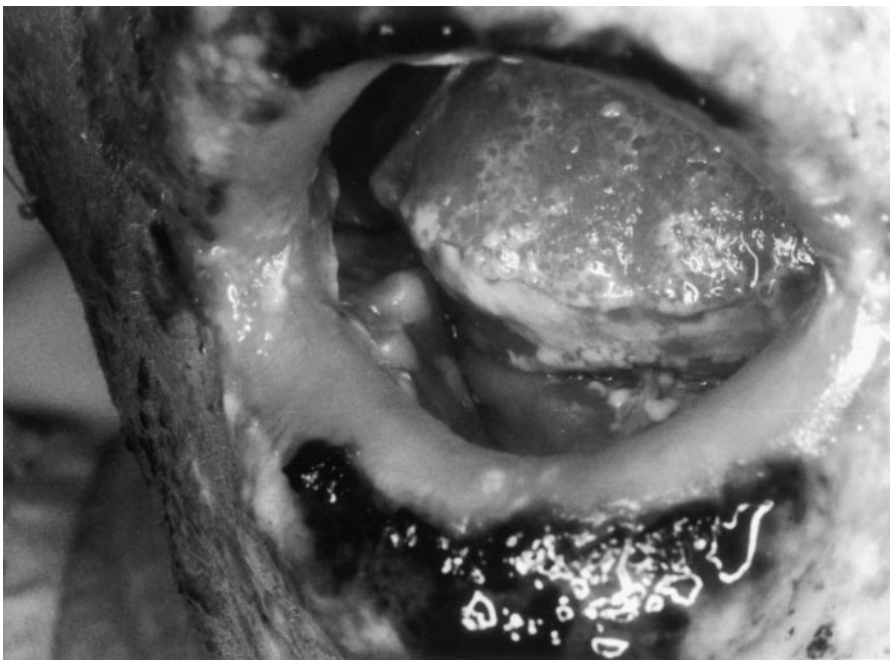


FIGURE 67–18. Oral cavity of a child with Stevens-Johnson syndrome, demonstrating skin, lip, tongue, and mucosal involvement.

Benign Tumors Benign esophageal tumors, less common than malignant esophageal tumors, are divided into epithelial and nonepithelial tumors. *Squamous papillomas* are the only benign epithelial tumor. Endoscopically, they appear as a warty tumor that usually remains small and asymptomatic. Benign nonepithelial tumors include leiomyomas, lipomas, fibromas, neurofibromas, gliomas, granular cell tumors, and osteochondromas.[97,108,109]

More than half of benign esophageal tumors are *leiomyomas*, presenting as intramural, encapsulated tumors arising from the muscularis mucosa. As they are usually of smooth muscle origin, they are most often found in the middle or lower third of the esophagus. No clinically characteristic features of leiomyoma exist; over half of patients are asymptomatic.[110] When symptoms are present, the most common complaints are dysphagia, pain, heartburn, and weight loss.[111] A chest radiograph may show a posterior mediastinal tumor or diffuse widening of the lower mediastinum. A barium esophagram will usually show a smooth filling defect; large tumors may cause irregular, lobulated filling defects. Computed tomography with contrast will show an intramural, eccentric, mediastinal mass with homogeneous enhancement. Endoscopic examination will reveal a submucosal mass. Histologically, a leiomyoma may resemble a malignant leiomyosarcoma.

Hemangiomas account for 2 to 3% of benign esophageal neoplasms. They appear pale blue and empty when compressed.[108] *Adenomatous polyps* are rare and are associated with Barrett's metaplasia.[97]

Esophageal cysts are also uncommon and may present with respiratory symptoms in infants and dysphagia and pain in adults. Cysts in infants are often lined with gastric mucosa; ulceration into the airway can cause hemoptysis. Appearance on a barium swallow study is similar to that of a leiomyoma except that the edges of an esophageal cyst have a more tapered appearance.[110]

Malignant Tumors *Squamous cell carcinoma* is the most common malignant esophageal tumor. Screening endoscopy should be performed in patients with known risk factors such as alcohol and tobacco abuse, achalasia, Plummer-Vinson syndrome, tylosis (palmar and plantar keratoderma), chronic stricture from ingestion of lye or corrosive substances, and celiac disease.[108,112] Often advanced disease presents at the time of diagnosis. The most common symptom is dysphagia; other symptoms include

aspiration pneumonia, hoarseness, unilateral neck mass, cough, fever, or a choking sensation. Endoscopically, lesions are usually found in the middle and lower third of the esophagus; they may be polypoid, ulcerative, or infiltrating. The most common complications of the carcinoma and its treatment are strictures, fistulae, and hemorrhage.[108]

Adenocarcinoma is the second most common esophageal epithelial malignancy, usually occurring in the distal esophagus in association with Barrett's metaplasia. Symptoms are similar to those of squamous cell carcinoma.

Other malignant epithelial tumors are rare and include variants of squamous cell carcinoma, such as spindle cell carcinoma, verrucous carcinoma or pseudosarcoma, adenoid cystic carcinoma, mucoepidermoid carcinoma, argyrophil cell carcinoma, and melanoma.[108]

Nonepithelial malignant esophageal tumors include leiomyosarcoma, rhabdomyosarcoma, and fibrosarcoma.[97,108] Direct esophageal infiltration or compression may be caused by thyroid, lung, or lymph node malignancies.[97,108]

Hematemesis

Esophageal varices, usually limited to the distal half of the esophagus, develop as a result of portal hypertension. Approximately 50% of cirrhotics with esophageal varices will bleed in their lifetime.[113] Mortality is highest in the first few days to weeks following the hemorrhage; early intervention is crucial for survival.[113–116] Endoscopic treatment options include sclerotherapy, ligation, or electro- or laser coagulation of bleeding points; ligation is considered the superior method.[77,114–116]

Mallory-Weiss tears develop in the gastric cardia and gastroesophageal junction after an episode of retching, vomiting, and coughing and are usually limited to the mucosa. Spontaneous cessation of bleeding occurs in 90% of patients.[117] This previously fatal condition now has a good prognosis with treatment; options include endoscopic bipolar cautery, sclerotherapy, and laser photocoagulation.

Boerhaave's syndrome describes the spontaneous, full-thickness esophageal tear that results most commonly from a prolonged bout of violent emesis. The patient may present with excruciating lower chest pain, shock, dyspnea, and extreme thirst. Survival is dependent on rapid diagnosis and treatment. Chest radiography may reveal pneumomediastinum or mediastinal widening; a swallowing study with radiopaque contrast will demonstrate the site of rupture. Esophagoscopy should not be performed as it will not add any additional information and may cause further damage.[118]

Other rare diseases that can present with esophagitis or stenosis include Wegener's granulomatosis, Crohn's disease, tuberculosis, syphilis, and Behçet's syndrome.[119–123]

REFERENCES

1. Heffner DK. Diseases of the trachea. In: Barnes L, editor. Surgical pathology of the head and neck. Vol 1. New York: Dekker; 1985. p. 437–531.
2. Proctor DF. The upper airways II. The larynx and trachea. Am Rev Respir Dis 1977;115:315–42.
3. Patel SR, Stoller JK. The role of bronchoscopy in hemoptysis. In: Wang KP, Mehta AC, editors. Flexible bronchoscopy. Cambridge (MA): Blackwell Science; 1995. p. 298–321.
4. Poe RH, Israel RH, Utell MJ, Hall WJ. Chronic cough: bronchoscopy or pulmonary function testing? Am Rev Respir Dis 1982;126:160–2.
5. Maze A, Bloch E. Stridor in pediatric patients. Anesthesiology 1979;50:132–45.
6. Haben CM, Abdulhadi K, Sadeghi N. Aerophagia resulting in acute airway obstruction. Otolaryngol Head Neck Surg 2000;123:650–1.
7. Grillo HC, Donahue DM. Postintubation tracheal stenosis. Chest Surg Clin N Am 1996;6:725–31.
8. Wood RE. Pitfalls in the use of the flexible bronchoscope in pediatric patients. Chest 1990;97:199–203.
9. Kirsh MM, Orringer MB, Behrendt DM, Sloan H. Management of tracheobronchial disruption secondary to nonpenetrating trauma. Ann Thorac Surg 1976;22:93–101.
10. Masanes MJ, Legendre C, Lioret N, et al. Using bronchoscopy and biopsy to diagnose early inhalation injury. Macroscopic and histologic findings. Chest 1995;107:1365–9.
11. Ramzy PI, Barret JP, Herndon DN. Thermal injury. Crit Care Clin 1999;15:333–52.
12. Atkins JP, Rowe LD, Jafek BW. Newer techniques in fiberoptic bronchoscopy. Ann Otol Rhinol Laryngol 1976;85:646–51.
13. Su WJ, Lee PY, Perng RP. Chest roentgenographic guidelines in the selection of patients for fiberoptic bronchoscopy. Chest 1993;103:1198–201.

14. Prakash UB. Advances in bronchoscopic procedures. Chest 1999;116:1403–8.
15. Dasgupta A, Mehta AC. Transbronchial needle aspiration: an underused diagnostic technique. Clin Chest Med 1999;20:39–52.
16. Ahrens P, Noll C, Kitz R, et al. Lipid-laden alveolar macrophages (LLAM): a useful marker of silent aspiration in children. Pediatr Pulmonol 1999;28: 83–8.
17. de Blic J, Scheinmann P. Fibreoptic bronchoscopy in infants. Arch Dis Child 1992;67:159–61.
18. Meduri GU, Chastre J. The standardization of bronchoscopic techniques for ventilator-associated pneumonia. Chest 1992;102:557S–64S.
19. Thompson AB, Rennard SI. Diagnostic procedures not involving the pleura. In: Baum GL, Crapo JD, Celli BR, Karlinsky JB, editors. Textbook of pulmonary diseases. 6th ed. Philadelphia: Lippincott-Raven; 1998. p. 239–45.
20. Deutsch ES, Smergel E, Crisci K, Panitch H. Tracheobronchography in children. Laryngoscope 1996;106:1248–54.
21. Baumann MH, Sahn SA. Medical management and therapy of bronchopleural fistulas in the mechanically ventilated patient. Chest 1990;97:721–8.
22. Wang MB, Berke GS, Ward PH, et al. Early experience with percutaneous tracheotomy. Laryngoscope 1992;102:157–62.
23. Hinerman R, Alvarez F, Keller CA. Outcome of bedside percutaneous tracheostomy with bronchoscopic guidance. Intensive Care Med 2000;26:1850–6.
24. Portex. Per-fit™ percutaneous dilational tracheostomy kit: a physician's clinical guide. Keen (NH): SIMS Portex; 1999.
25. Mehta AC, Dasgupta A. Airway stents. Clin Chest Med 1999;20:139–52.
26. Cook CH, Bhattacharyya N, King DR. Aortobronchial fistula after expandable metal stent insertion for pediatric bronchomalacia. J Pediatr Surg 1998;33:1306–8.
27. Keane WM, Atkins JP. CO_2 laser surgery of the upper airway. Surg Clin North Am 1984;64:955–71.
28. Oho K, Amemiya R. Applications of fiberoptic bronchoscopy. In: Practical fiberoptic bronchoscopy. 2nd ed. Tokyo: Igaku-Shoin; 1984. p. 151–206.
29. Turner JF, Wang KP. Endobronchial laser therapy. Clin Chest Med 1999;20:107–22.
30. Jackson C. Bronchoscopy and esophagoscopy. A manual of peroral endoscopy and laryngeal surgery. Philadelphia: WB Saunders; 1922.
31. Liebler JM, Markin CJ. Fiberoptic bronchoscopy for diagnosis and treatment. Crit Care Clin 2000;16: 83–100.
32. Good GM, Deutsch ES. Method for removing endobronchial beads. Ann Otol Rhinol Laryngol 1998; 107:291–2.
33. Zeitlin J, Myer CM. Foreign body in the trachea. Ann Otol Rhinol Laryngol 2000;109:1007–8.
34. Ahmad M, Dweik RA. Future of flexible bronchoscopy. Clin Chest Med 1999;20:1–17.
35. Sheski FD, Mathur PN. Cryotherapy, electrocautery, and brachytherapy. Clin Chest Med 1999;20:123–38.
36. Stern RC. Lower respiratory tract: acquired diseases. In: Behrman RE, Kleigman RM, Nelson WE, Vaughn VC, editors. Nelson's textbook of pediatrics. 14th ed. Philadelphia: WB Saunders; 1992. p. 1064–101.
37. Broughton WA, Middleton RM, Kirkpatrick MB, Bass JB. Bronchoscopic protected specimen brush and bronchoalveolar lavage in the diagnosis of bacterial pneumonia. Infect Dis Clin North Am 1991;5:437–52.
38. Stanford W, Galvin JR. The diagnosis of bronchiectasis. Clin Chest Med 1988;9:691–9.
39. Stevens DA, Kan, VL, Judson MA, et al. Practice guidelines for diseases caused by *Aspergillus*. Clin Infect Dis 2000;30:696–709.
40. Guleria R, Gupta D, Jindal S. Treatment of pulmonary aspergilloma by endoscopic intracavitary instillation of ketoconazole. Chest 1993;103:1301–2.
41. McDonald RJ, Reichman LB. Tuberculosis. In: Baum GL, Crapo JD, Celli BR, Karlinski JB, editors. Textbook of pulmonary diseases. 6th ed. Philadelphia: Lippincott-Raven; 1998. p. 603–30.
42. Correa AG. Unique aspects of tuberculosis in the pediatric population. Clin Chest Med 1997;18:89–98.
43. Griffith DE. Mycobacteria as pathogens of respiratory infections. Infect Dis Clin North Am 1998; 12:593–611.
44. Boat TF. The respiratory system, the lower respiratory tract. In: Behrman RE, Kleigman RM, Nelson WE, Vaughn VC, editors. Nelson's textbook of pediatrics. 14th ed. Philadelphia: WB Saunders; 1992. p. 1102–16.
45. Tasker AD, Flower CD. Imaging the airways: hemoptysis, bronchiectasis, and small airway diseases. Clin Chest Med 1999;20:761–73.
46. DePaso WJ, Winterbauer RH. Interstitial lung disease. Dis Mon 1991;37:63–133.
47. Lavergne F, Clerici C, Sadoun D, et al. Airway obstruction in bronchial sarcoidosis: outcome with treatment. Chest 1999;116:1194–9.

48. Lynch JP III, Kazerooni EA, Gay SE. Pulmonary sarcoidosis. Clin Chest Med 1997;18:755–85.
49. Kenny TJ, Werkhaven J, Netterville JL. Sarcoidosis of the pediatric larynx. Arch Otolaryngol Head Neck Surg 2000;126:536–9.
50. Michaelson JE, Aguayo SM, Roman J. Idiopathic pulmonary fibrosis: a practical approach for diagnosis and management. Chest 2000;118:788–94.
51. King TE, Costabel U, Cordier J, et al. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. Am J Respir Crit Care Med 2000;161:646–64.
52. Cordier J, Valeyre D, Guillevin L, et al. Pulmonary Wegener's granulomatosis: a clinical and imaging study of 77 cases. Chest 1990;97:906–12.
53. Daum TE, Specks U, Colby TV, et al. Tracheobronchial involvement in Wegener's granulomatosis. Am J Respir Crit Care Med 1995;151:522–6.
54. Hodges MK, Israel E. Tracheobronchopathia osteochondroplastica presenting as right middle lobe collapse. Diagnosis by brochoscopy and computerized tomography. Chest 1988;94:842–4.
55. Strauss GM. Bronchogenic carcinoma. In: Baum GL, Crapo JD, Celli BR, Karlinski JB, editors. Textbook of pulmonary diseases. 6th ed. Philadelphia: Lippincott-Raven; 1998. p. 1329–81.
56. Merrick SH. Tumors of the lung other than bronchogenic carcinoma. In: Baum GL, Crapo JD, Celli BR, Karlinski JB, editors. Textbook of pulmonary diseases. 6th ed. Philadelphia: Lippincott-Raven; 1998. p. 1383–98.
57. Yang JY, Yang WG, Chang LY, Chuang SS. Symptomatic tracheal stenosis in burns. Burns 1999;25: 72–80.
58. Rahbar R, Shapshay SM, Healy GB. Mitomycin: effects on laryngeal and tracheal stenosis, benefits, and complications. Ann Otol Rhinol Laryngol 2001;110:1–6.
59. Meyer GW, Castell DO. Current concepts of esophageal function. Am J Otolaryngol 1980;1:440–6.
60. Berdon WE, Baker DH. Vascular abnormalities and the infant lung: rings, slings, and other things. Semin Roetgenol 1972;7:39–64.
61. Menzel J, Hoepffner N, Nottberg H, et al. Preoperative staging of esophageal carcinoma: miniprobe sonography versus conventional endoscopic ultrasound in a prospective histopathologically verified study. Endoscopy 1999;31:291–7.
62. Reed CE, Mishra G, Sahai AV, et al. Esophageal cancer staging: improved accuracy by endoscopic ultrasound of celiac lymph nodes. Ann Thorac Surg 1999;67:319–22.
63. Tytgat GN. Upper gastrointestinal endoscopy. In: Yamada T, Alpers DH, Owyang C, et al, editors. The textbook of gastroenterology. 2nd ed. Philadelphia: JB Lippincott; 1995. p. 2545–70.
64. Loh KS, Tan LK, Smith JD, et al. Complications of foreign bodies in the esophagus. Otolaryngol Head Neck Surg 2000;123:613–6.
65. Hawkins DB. Removal of blunt foreign bodies from the esophagus. Ann Otol Rhinol Laryngol 1990;99: 935–40.
66. Christie DL, Ament ME. Removal of foreign bodies from esophagus and stomach with flexible fiberoptic panendoscopes. Pediatrics 1976;57:931–4.
67. Berggreen PJ, Harrison E, Sanowski RA, et al. Techniques and complications of esophageal foreign body extraction in children and adults. Gastrointest Endosc 1993;39:626–30.
68. Mehta DI, Attia MW, Quintana EC, Cronan KM. Glucagon use for esophageal coin dislodgment in children: a prospective, double-blind, placebo-controlled trial. Acad Emerg Med 2001;8:200–3.
69. Litovitz T, Schmitz BF. Ingestion of cylindrical and button batteries: an analysis of 2382 cases. Pediatrics 1992;89(4 Pt 2):747–57.
70. Lehmann CU , Elitsur Y. Keep the change, doc!: coins in the upper GI tract of children. W V Med J 1995;91:13–5.
71. Bertoni G, Sassatelli R, Conigliaro R, Bedogni G. A simple latex protector hood for safe endoscopic removal of sharp-pointed gastroesophageal foreign bodies. Gastrointest Endosc 1996;44:458–61.
72. Winkler AR, McClenathan DT, Borger JA, Ahmed N. Retrograde esophagoscopy for foreign body removal. J Pediatr Gastroenterol Nutr 1989;8:536–40.
73. Sheikh A. Button battery ingestion in children. Pediatr Emerg Care 1993;9:224–9.
74. Holinger LD. Caustic ingestion, esophageal injury and stricture. In: Holinger LD, Lusk RP, Green CG, editors. Pediatric laryngology and bronchoesophagology. Philadelphia: Lippincott-Raven; 1997. p. 295–303.
75. Myer CM III. Potential hazards of esophageal foreign body extraction. Pediatr Radiol 1991;21:97–8.
76. Nijhawan S, Rai RR, Napalia S, Pokharna R. Suction retrieval of esophageal foreign bodies. Endoscopy 1998;30 Suppl:59.
77. Steiner W. Techniques of diagnostic and operative endoscopy of the head and neck. Tracheoscopy,

bronchoscopy, esophagoscopy, mediastinoscopy, interdisciplinary panendoscopy. Endoscopy 1979;11: 151–7.
78. Crain EF, Gershel JC, Mezey AP. Caustic ingestion. Symptoms as predictors of esophageal injury. Am J Dis Child 1984;138:863–9.
79. Lovejoy FH. Corrosive injury of the esophagus in children: failure of corticosteroid treatment reemphasizes prevention. N Engl J Med 1990;323:668–70.
80. Anderson KD, Rouse TM, Randolph JG. A controlled trial of corticosteroids in children with corrosive injury of the esophagus. N Engl J Med 1990;323:637–40.
81. Scott JC, Jones B, Eisele DW, Ravich WJ. Caustic ingestion injuries of the upper aerodigestive tract. Laryngoscope 1992;102:1–8.
82. Forsen JW, Muntz HR. Hair relaxer ingestion: a new trend. Ann Otol Rhinol Laryngol 1993;102:781–4.
83. Einhorn A, Horton L, Altieri M, et al. Serious respiratory consequences of detergent ingestion in children. Pediatrics 1989;84:472–4.
84. Cox AJ III, Eisenbeis JF. Ingestion of caustic hair relaxer: is endoscopy necessary? Laryngoscope 1997;107:897–902.
85. Ahsan S, Haupert M. Absence of esophageal injury in pediatric patients after hair relaxer ingestion. Arch Otolaryngol Head Neck Surg 1999;125:953–5.
86. Skinner DB, Belsey RH. Achalasia of the cardia. In: Management of esophageal disease. Philadelphia: WB Saunders; 1988. p. 453–84.
87. Wong RK, Maydonovitch CL. Achalasia. In: Castell DO, editor. The esophagus. 2nd ed. New York: Little, Brown, and Company; 1995. p. 219–45.
88. Pasricha PJ, Ravich WJ, Hendrix TR, et al. Intrasphincteric botulinum toxin for the treatment of achalasia. N Engl J Med 1995;322:774–8.
89. Lefkowitz JR, Brand DL, Schuffler MD, Brugge WR. Amyloidosis mimics achalasia's effect on lower esophageal sphincter. Dig Dis Sci 1989;34:630–5.
90. Van Overbeek JJ. Meditations on the pathogenesis of hypopharyngeal (Zenker's) diverticulum and report of endoscopic treatment in 545 patients. Ann Otol Rhinol Laryngol 1994;103:178–85.
91. Konowitz PM, Biller HF. Diverticulopexy and cricopharyngeal myotomy: treatment for the high-risk patient with a pharyngoesophageal (Zenker's) diverticulum. Otolaryngol Head Neck Surg 1989;100:146–53.
92. Feeley MA, Righi PD, Weisberger EC, et al. Zenker's diverticulum: analysis of surgical complications from diverticulectomy and cricopharyngeal myotomy. Laryngoscope 1999;109:858–61.
93. Knegt PP, de Jong PC, van der Schans EJ. Endoscopic treatment of the hypopharyngeal diverticulum with the CO_2 laser. Endoscopy 1985;17:205–6.
94. Dohlman G, Mattsson O. The endoscopic operation for hypopharyngeal diverticula. Arch Otolaryngol 1960;71:744–52.
95. Philippsen LP, Weisberger EC, Whiteman TS, Schmidt JL. Endoscopic stapled diverticulotomy: treatment of choice for Zenker's diverticulum. Laryngoscope 2000;110:1283–6.
96. Spechler SJ. Complications of gastroesophageal reflux disease. In: Castell DO, editor. The esophagus. 2nd ed. New York: Little, Brown, and Company; 1995. p. 533–45.
97. Silverstein FE, Tytgat GN. Esophagus II: neoplasms, infections, and involvement in graft-versus-host disease. In: Atlas of gastrointestinal endoscopy. 2nd ed. Philadelphia: JB Lippincott; 1991. p. 59–91.
98. Hostein J, Fournet J. Gastrointestinal manifestations of collagen diseases. Dig Dis 1986;4:240–52.
99. Neil GA, Weinstock JV. Gastrointestinal manifestations of systemic diseases. In: Yamada T, Alpers DH, Owyang C, et al, editors. Textbook of gastroenterology. 2nd ed. Philadelphia: JB Lippincott; 1995. p. 2419–55.
100. Gryboski JD, Touloukian R, Campanella RA. Gastrointestinal manifestations of epidermolysis bullosa in children. Arch Dermatol 1988;124:746–52.
101. Ergun GA, Lin AN, Dannenberg AJ, Carter DM. Gastrointestinal manifestations of epidermolysis bullosa. A study of 101 patients. Medicine 1992;71:121–7.
102. Orlando RC, Bozymski EM, Briggaman RA, Bream CA. Epidermolysis bullosa: gastrointestinal manifestations. Ann Intern Med 1974;81:203–6.
103. Mignogna MD, Lo Muzio LL, Galloro G, et al. Oral pemphigus: clinical significance of esophageal involvement: report of eight cases. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1997;84:179–84.
104. Witte JT, Icken JN, Lloyd ML. Induction of esophageal bullae in benign mucous membrane pemphygoid. Gastrointest Endosc 1989;35:566–8.
105. Isolauri J, Airo I. Benign mucous membrane pemphigoid involving the esophagus: a report of two cases treated with dilation. Gastrointest Endosc 1989;35:569–71.
106. Fitzpatrick TB, Johnson RA, Polano MK, et al, editors. Color atlas and synopsis of clinical dermatology. 2nd ed. New York: McGraw-Hill; 1992.
107. Gluckman JL, Righi PD. Inflammatory disease of the mouth and pharynx. In: Bluestone CD, Stool SE,

Kenna MA, editors. Pediatric otolaryngology. 3rd ed. Philadelphia: WB Saunders; 1996. p. 1043–53.
108. Miller LS. Endoscopy of the esophagus. In: Castell DO, editor. The esophagus. 2nd ed. New York: Little, Brown, and Company; 1995. p. 93–132.
109. Cox MA, Cooper BT, Sagar G. Endoscopy and computed tomography in the diagnosis and follow up of oesophageal leiomyoma. Gut 1995;37: 288–91.
110. Postlethwait RW. Benign tumors and cysts of the esophagus. Surg Clin North Am 1983;63:925–31.
111. Skinner DB, Belsey RH. Benign tumors of the esophagus. In: Management of esophageal disease. Philadelphia: WB Saunders; 1988. p. 717–27.
112. Allen JW, Richardson JC, Edwards MJ. Squamous cell carcinoma of the esophagus: a review and update. Surg Oncol 1997;6:193–200.
113. Young HS, Matsui SM, Gregory PB. Endoscopic control of upper gastrointestinal variceal bleeding. In: Yamada T, Alpers DH, Owyang C, et al, editors. Textbook of gastroenterology. 2nd ed. Philadelphia: JB Lippincott; 1995. p. 2969–91.
114. Djurdjevic D, Janosevic S, Dapcevic B, et al. Combined ligation and sclerotherapy versus ligation alone for eradication of bleeding esophageal varices: a randomized and prospective trial. Endoscopy 1999;31:286–90.
115. Umehara M, Onda M, Tajiri T, et al. Sclerotherapy plus ligation versus ligation for the treatment of esophageal varices: a prospective randomized study. Gastrointest Endosc 1999;50:7–12.
116. Al Traif IA, Fachartz FS, Al Jumah AA, et al. Randomized trial of ligation versus combined ligation and sclerotherapy for bleeding esophageal varices. Gastrointest Endosc 1999;50:1–6.
117. Gilliam JH III. Esophageal varices and Mallory-Weiss tears. In: Castell DO, editor. The esophagus. 2nd ed. New York: Little, Brown, and Company; 1995. p. 757–67.
118. Skinner DB, Belsey RH. Spontaneous rupture and Boerhaave's syndrome. In: Management of esophageal disease Philadelphia: WB Saunders; 1988. p. 802–13.
119. Bottomley WW, Dakkak M, Walton S, Bennett JR. Esophageal involvement in Behçet's disease. Is endoscopy necessary? Dig Dis Sci 1992;37:594–7.
120. Mori S, Yoshihira A, Kawamura H, et al. Esophageal involvement in Behçet's disease. Am J Gastroenterol 1983;78:548–53.
121. Gheorghe C, Aposteanu G, Popescu C, et al. Long esophageal stricture in Crohn's disease: a case report. Hepatogastroenterology 1998;45:738–41.
122. Geboes K, Janssens J, Rutgeerts P, Vantrappen G. Crohn's disease of the esophagus. J Clin Gastroenterol 1986;8:31–7.
123. Spiera RF, Filippa DA, Bains MS, Paget SA. Esophageal involvement in Wegener's granulomatosis. Arthritis Rheum 1994;37:1404–7.

Index