PART ONE -INTRODUCTION TO CLINICAL MEDICINE

1. THE PRACTICE OF MEDICINE - *The Editors*

WHAT IS EXPECTED OF THE PHYSICIAN

The practice of medicine combines both science and art. The role of *science in medicine* is clear. Science-based technology and deductive reasoning form the foundation for the solution to many clinical problems; the spectacular advances in genetics, biochemistry, and imaging techniques allow access to the innermost parts of the cell and the most remote recesses of the body. Highly advanced therapeutic maneuvers are increasingly a major part of medical practice. Yet skill in the most sophisticated application of laboratory technology and in the use of the latest therapeutic modality alone does not make a good physician. One must be able to identify the crucial elements in a complex history and physical examination and extract the key laboratory results from the crowded computer printouts of laboratory data in order to determine in a difficult case whether to "treat" or to "watch." Deciding when a clinical clue is worth pursuing, or when it should be dismissed as a "red herring," and estimating in any given patient whether a proposed treatment entails a greater risk than the disease are essential to the decision-making process that the skilled clinician must exercise many times each day. This combination of medical knowledge, intuition, and judgment defines the *art of medicine*, which is as necessary to the practice of medicine as is a sound scientific base.

The editors of the first edition of this book articulated what is expected of the physician in words that, although they reflect the gender bias of that era, still ring true as a universal principle:

No greater opportunity, responsibility, or obligation can fall to the lot of a human being than to become a physician. In the care of the suffering he needs technical skill, scientific knowledge, and human understanding. He who uses these with courage, with humility, and with wisdom will provide a unique service for his fellow man, and will build an enduring edifice of character within himself. The physician should ask of his destiny no more than this; he should be content with no less.

Tact, sympathy and understanding are expected of the physician, for the patient is no mere collection of symptoms, signs, disordered functions, damaged organs, and disturbed emotions. He is human, fearful, and hopeful, seeking relief, help and reassurance.

THE PATIENT-PHYSICIAN RELATIONSHIP

It may seem trite to emphasize that physicians need to approach patients not as "cases" or "diseases" but as individuals whose problems all too often transcend their physical complaints. Most patients are anxious and frightened. Physicians should instill confidence and reassurance, overtly and in their demeanor, but without an air of arrogance. A professional attitude, coupled with warmth and openness, can do much to alleviate the patients' anxiety and to encourage them to share parts of their history that may be embarrassing. Some patients "use" illness to gain attention or to serve as a

crutch to extricate themselves from a stressful situation; some even feign physical illness; others may be openly hostile. Whatever the patient's attitude, the physician needs to consider the setting in which an illness occurs -- in terms not only of the patients themselves but also of their families and social and cultural backgrounds. The ideal patient-physician relationship is based on thorough knowledge of the patient, on mutual trust, and on the ability to communicate with one another.

The direct, one-to-one patient-physician relationship, which has traditionally characterized the practice of medicine, is increasingly in jeopardy because of the increasing complexity of medicine and change in health care delivery systems. Often the management of the individual patient is a team effort involving a number of several different physicians and professional personnel. The patient can benefit greatly from such collaboration, but *it is the duty of the patient's principal physician to guide them through an illness*. To carry out this difficult task, this physician must be familiar with the techniques, skills, and objectives of specialist physicians and of colleagues in the fields allied to medicine. In giving the patient an opportunity to benefit from scientific advances, the primary physician must, in the last analysis, retain responsibility for the major decisions concerning diagnosis and treatment.

Patients are increasingly cared for by groups of physicians in clinics, hospitals, integrated health care delivery systems, and health maintenance organizations (HMOs). Whatever the potential advantages of such organized medical groups, there are also drawbacks, chiefly the loss of the clear identification of the physician who is primarily and continuously responsible for the patient. Even under these circumstances, it is essential for each patient to have a physician who has an overview of the problems and who is familiar with the patient's reaction to the illness, to the drugs given, and to the challenges that the patient faces.

The practice of medicine in a "managed care" setting puts additional stress on the classic paradigm of the patient-physician relationship. Many physicians must deal with a patient within a restricted time frame, with limited access to specialists, and under organizational guidelines that may compromise their ability to exercise their individual clinical judgment. As difficult as these restrictions may be, it is the ultimate responsibility of the physician to determine what is best for the patient. This responsibility cannot be relinquished in the name of compliance with organizational guidelines.

The physician must also bear in mind that the modern hospital constitutes an intimidating environment for most patients. Lying in a bed surrounded by air jets, buttons, and lights; invaded by tubes and wires; beset by the numerous members of the health care team -- nurses, nurses' aides, physicians' assistants, social workers, technologists, physical therapists, medical students, house officers, attending and consulting physicians, and many others; sharing rooms with other patients who have their own problems, visitors, and physicians; transported to special laboratories and imaging facilities replete with blinking lights, strange sounds, and unfamiliar personnel -- it is little wonder that patients may lose their sense of reality. In fact, the physician is often the only tenuous link between the patient and the real world, and a strong personal relationship with the physician helps to sustain the patient in such a stressful situation.

Many trends in contemporary society tend to make medical care impersonal. Some of these have been mentioned already and include (1) vigorous efforts to reduce the escalating costs of health care; (2) the growing number of managed care programs, which are intended to reduce costs but in which the patient may have little choice in selecting a physician; (3) increasing reliance on technologic advances and computerization for many aspects of diagnosis and treatment; (4) increased geographic mobility of both patients and physicians; (5) the need for numerous physicians to be involved in the care of most patients who are seriously ill; and (6) an increasing tendency on the part of patients to express their frustrations with the health care system by legal means (i.e., by malpractice litigation). Given these changes in the medical care system, it is a major challenge for physicians to maintain the *humane* aspects of medical care. The American Board of Internal Medicine has defined humanistic qualities as encompassing integrity, respect, and compassion. Availability, the expression of sincere concern, the willingness to take the time to explain all aspects of the illness, and a nonjudgmental attitude when dealing with patients whose cultures, lifestyles, attitudes, and values differ from those of the physician are just a few of the characteristics of the humane physician. Every physician will, at times, be challenged by patients who evoke strongly negative (or strongly positive) emotional responses. Physicians should be alert to their own reactions to such patients and situations and should consciously monitor and control their behavior so that the patients' best interests remain the principal motivation for their actions at all times.

An important aspect of patient care involves an appreciation of the "quality of life," a subjective assessment of what each patient values most. Such an assessment requires detailed, sometimes intimate knowledge of the patient, which can usually be obtained only through deliberate, unhurried, and often repeated conversations. It is in these situations that the time constraints of a managed care setting may prove problematic.

The famous statement of Dr. Francis Peabody is even more relevant today than when delivered more than three-quarters of a century ago:

The significance of the intimate personal relationship between physician and patient cannot be too strongly emphasized, for in an extraordinarily large number of cases both the diagnosis and treatment are directly dependent on it. One of the essential qualities of the clinician is interest in humanity, ***for the secret of the care of the patient is in caring for the patient.***

CLINICAL SKILLS

History Taking The written history of an illness should embody all the facts of medical significance in the life of the patient. Recent events should be given the most attention. The patient should, at some point, have the opportunity to tell his or her own story of the illness without frequent interruption and, when appropriate, receive expressions of interest, encouragement, and empathy from the physician. The physician must be alert to the possibility that any event related by the patient, however trivial or apparently remote, may be the key to the solution of the medical problem.

An informative history is more than an orderly listing of symptoms; something is always gained by listening to patients and noting the way in which they describe their

symptoms. Inflections of voice, facial expression, gestures, and attitude may reveal important clues to the meaning of the symptoms to the patient. Taking history often involves much data gathering. Patients vary in their medical sophistication and ability to recall facts. Medical history should therefore be corroborated whenever possible. The family and social history can also provide important insights into the types of diseases that should be considered. In listening to the history, the physician discovers not only something about the disease but also something about the patient. The process of history taking provides an opportunity to observe the patient's behavior and to watch for features to be pursued more thoroughly during the physical examination.

The very act of eliciting the history provides the physician with the opportunity to establish or enhance the unique bond that is the basis for the ideal patient-physician relationship. It is helpful to develop an appreciation of the patient's perception of the illness, the patient's expectations of the physician and the medical care system, and the financial and social implications of the illness to the patient. The confidentiality of the patient-physician relationship should be emphasized, and the patient should be given the opportunity to identify any aspects of the history that should not be disclosed.

Physical Examination Physical signs are objective indications of disease whose significance is enhanced when they confirm a functional or structural change already suggested by the patient's history. At times, however, the physical signs may be the only evidence of disease.

The physical examination should be performed methodically and thoroughly, with consideration for the patient's comfort and modesty. Although attention is often directed by the history to the diseased organ or part of the body, the examination of a new patient must extend from head to toe in an objective search for abnormalities. Unless the physical examination is systematic, important segments may be omitted. The results of the examination, like the details of the history, should be recorded at the time they are elicited, not hours later when they are subject to the distortions of memory. Skill in physical diagnosis is acquired with experience, but it is not merely technique that determines success in eliciting signs. The detection of a few scattered petechiae, a faint diastolic murmur, or a small mass in the abdomen is not a question of keener eyes and ears or more sensitive fingers but of a mind alert to these findings. Since physical findings are subject to changes, the physical examination should be repeated as frequently as the clinical situation warrants.

Laboratory Tests The availability of a wide array of laboratory tests has increased our reliance on these studies for the solution of clinical problems. The accumulation of laboratory data does not relieve the physician from the responsibility of careful observation, examination, and study of the patient. It is also essential to bear in mind the limitations of such tests. By virtue of their impersonal quality, complexity, and apparent precision, they often gain an aura of authority regardless of the fallibility of the tests themselves, the instruments used in the tests, and the individuals performing or interpreting them. Physicians must weigh the expense involved in the laboratory procedures they order relative to the value of the information they are likely to provide.

Single laboratory tests are rarely ordered. Rather, they are generally obtained as "batteries" of multiple tests, which are often useful. For example, abnormalities of

hepatic function may provide the clue to such nonspecific symptoms as generalized weakness and increased fatigability, suggesting the diagnosis of chronic liver disease. Sometimes a single abnormality, such as an elevated serum calcium level, points to particular diseases, such as hyperparathyroidism or underlying malignancy.

The thoughtful use of screening tests should not be confused with indiscriminate laboratory testing. The use of screening tests is based on the fact that a group of laboratory determinations can be carried out conveniently on a single specimen of blood at relatively low cost. Screening tests are most useful when they are directed towards common diseases or disorders in which the result directs other useful tests or interventions that would otherwise be costly to perform. Biochemical measurements, together with simple laboratory examinations such as blood count, urinalysis, and sedimentation rate, often provide the major clue to the presence of a pathologic process. At the same time, the physician must learn to evaluate occasional abnormalities among the screening tests that may not necessarily connote significant disease. An in-depth workup following a report of an isolated laboratory abnormality in a person who is otherwise well is almost invariably wasteful and unproductive. Among the more than 40 tests that are routinely performed on patients, one or two are often slightly abnormal. If there is no suspicion of an underlying illness, these tests are ordinarily repeated to ensure that the abnormality does not represent a laboratory error. If an abnormality is confirmed, it is important to consider its potential significance in the context of the patient's condition and other test results.

Imaging Techniques The availability of ultrasonography, a variety of scans that employ isotopes to visualize organs heretofore inaccessible, computed tomography, and magnetic resonance imaging has opened new diagnostic vistas and has benefited patients because these new techniques have largely supplanted more invasive ones. While the enthusiasm for noninvasive technology is understandable, the expense entailed in performing these tests is often substantial and should be considered when assessing the potential benefits of the information provided.

PRINCIPLES OF PATIENT CARE

Medical Decision-Making Both during and in particular after the physician has taken the history, performed the physical examination, and reviewed the laboratory and imaging data, the challenging process of the differential diagnosis and medical decision-making begins. Formulating a differential diagnosis requires not only a broad knowledge base but also the ability to assess the relative probabilities of various diseases and to understand the significance of missing diagnoses that may be less likely. Arriving at a diagnosis requires the application of the scientific method. Hypotheses are formed, data are collected, and objective conclusions are reached concerning whether to accept or reject a particular diagnosis. Analysis of the differential diagnosis is an iterative process. As new information or test results are acquired, the group of disease processes being considered can be contracted or expanded appropriately. Medical decision-making occurs throughout the diagnostic and treatment process. It involves the ordering of additional tests, requests for consults, and decisions regarding prognosis and treatment. This process requires an in-depth understanding of the natural history and pathophysiology of disease, explaining why these features are strongly emphasized in this textbook. As described below, medical decision-making

should be evidence-based, thereby ensuring that patients derive the full benefit of the scientific knowledge available to physicians.

Evidence-Based Medicine Sackett has defined evidence-based medicine as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients." Rigorously obtained evidence is contrasted with anecdotal experience, which is often biased. Even the most experienced physicians can be influenced by recent experiences with selected patients, unless they are attuned to the importance of using larger, more objective studies for making decisions. The prospectively designed, double-blind, randomized clinical trial represents the "gold standard" for providing evidence regarding therapeutic decisions, but it is not the only source. Valuable evidence about the natural history of disease and prognosis can come from prospective cohort studies and analytic surveys. Persuasive evidence on the accuracy of diagnostic tests can be derived from cross-sectional studies of patients in whom a specific disorder is suspected. Evidence is strengthened immensely when it has been confirmed by multiple investigations, which can be compared with one another and presented in a meta-analysis or systemic overview.

In failing to apply the best and most current evidence, the physician places the patient at unnecessary risk. However, a knowledge of or rapid access to the best available evidence is not sufficient for optimal care. The physician must know whether the evidence is relevant to the patient in question and, when it is, the consequences of applying it in any particular situation. The skills and judgment required to apply sound evidence represent an increasing challenge. Indeed, one might redefine a "good doctor" as one who uses the ever-growing body of rigorously obtained evidence (the science of medicine) in a sensible, compassionate manner (the art of medicine).

While an understanding of biologic and physiologic mechanisms forms the basis of contemporary medicine, when a therapeutic modality is selected, the highest priority must often be placed on improving *clinical outcome* rather than interrupting what is believed to be the underlying process. For example, for decades patients who had suffered myocardial infarction were treated intuitively with drugs that suppress frequent ventricular extrasystoles, since these were believed to be harbingers of ventricular fibrillation and sudden death. Clinical trials, however, have provided firm evidence that the antiarrhythmic agents actually increase the risk of death in such patients. This finding suggests that the extrasystoles are *markers* of high risk rather than the *cause* of fatal events.

Practice Guidelines Physicians are faced with a large, increasing, and often bewildering body of evidence pointing to potentially useful diagnostic techniques and therapeutic choices. The intelligent and cost-effective practice of medicine consists of making selections most appropriate to a particular patient and clinical situation. Professional organizations and government agencies are developing formal clinical practice guidelines in an effort to aid physicians and other caregivers in this endeavor. When guidelines are current and properly applied, they can provide a useful framework for managing patients with particular diagnoses or symptoms. They can protect patients -- particularly those with inadequate health care benefits -- from receiving substandard care. Guidelines can also protect conscientious caregivers from inappropriate charges of malpractice and society from the excessive costs associated with the overuse of

medical resources. On the other hand, clinical guidelines tend to oversimplify the complexities of medicine. Different groups with differing perspectives may develop divergent recommendations regarding issues as basic as the need for periodic sigmoidoscopy in middle-aged persons. Furthermore, guidelines do not -- and cannot be expected to -- take into account the uniqueness of each individual and of his or her illness. The challenge for the physician is to integrate into clinical practice the useful recommendations offered by the experts who prepare clinical practice guidelines without accepting them blindly or being inappropriately constrained by them.

Assessing the Outcome of Treatment Clinicians generally use *objective* and readily measurable parameters to judge the outcome of a therapeutic intervention. For example, findings on physical or laboratory examination -- such as the level of blood pressure, the patency of a coronary artery on an angiogram, or the size of a mass on a radiologic examination -- can provide information of critical importance. However, patients usually seek medical attention for *subjective* reasons; they wish to obtain relief from pain, to preserve or regain function, and to enjoy life. The components of a patient's health status or quality of life can include bodily comfort, capacity for physical activity, personal and professional function, sexual function, cognitive function, and overall perception of health. Each of these important areas can be assessed by means of structured interviews or specially designed questionnaires. Such assessments also provide useful parameters by which the physician can judge the patient's subjective view of his or her disability and the response to treatment, particularly in chronic illness. The practice of medicine requires consideration and integration of both objective and subjective outcomes.

Care of the Elderly Over the next several decades, the practice of medicine will be greatly influenced by the health care needs of the growing elderly population. In the United States the population over age 65 will almost triple over the next 30 years. It is essential that we understand and appreciate the physiologic processes associated with aging; the different responses of the elderly to common diseases; and disorders that occur commonly with aging, such as depression, dementia, frailty, urinary incontinence, and fractures. The elderly have more adverse reactions to drugs, in large part due to altered pharmacokinetics and pharmacodynamics. Commonly used medications such as digoxin and aminoglycosides have prolonged half-lives in the elderly, and tissues such as the central nervous system are more sensitive to certain drugs, such as the benzodiazepines and narcotics. The large number of drugs used by the elderly increases the risk of unwanted interactions, especially when care is provided by several physicians in an uncoordinated manner.

Diseases in Women versus Men In the past, many epidemiologic studies and clinical trials focused on men. It is now appreciated that there are significant gender differences in diseases that afflict both men and women. Mortality rates are substantially higher in women than in men under the age of 50 suffering acute myocardial infarction. Hypertension is more prevalent in African-American women than in their male counterparts (and in African-American than in white males); osteoporosis is more common in women, reflecting the menopausal loss of estrogen; diseases involving the immune system, such as lupus erythematosus, multiple sclerosis, and primary biliary cirrhosis, occur more frequently in women; and the average life expectancy of women is greater than that of men. Recently, considerable attention has been paid to women's

health issues, a subject that regrettably did not receive sufficient attention in the past. Ongoing study should enhance our understanding of the mechanisms of gender differences in the course and outcome of certain diseases.

Iatrogenic Disorders In an *iatrogenic disorder*, the deleterious effects of a therapeutic or diagnostic maneuver cause pathology independent of the condition for which the intervention was performed. Adverse drug reactions occur in at least 5% of hospitalized patients, and the incidence increases with use of a large number of drugs. No matter what the clinical situation, it is the responsibility of the physician to use powerful therapeutic measures wisely, with due regard for their beneficial action, potential dangers, and cost. Every medical procedure, whether diagnostic or therapeutic, has the potential for harm, but it would be impossible to provide the benefits of modern scientific medicine if reasonable steps in diagnosis and therapy were withheld because of possible risks. *Reasonable* implies that the physician has weighed the pros and cons of a procedure and has concluded, on the basis of objective evidence whenever possible, that it is necessary for establishing a diagnosis, for the relief of discomfort, or for the cure of disease. However, the harm that a physician can do is not limited to the imprudent use of medication or procedures. Equally important are ill-considered or unjustified remarks. Many a patient has developed a cardiac neurosis because the physician ventured a grave prognosis on the basis of a misinterpreted finding of a heart murmur. Not only the diagnostic procedure or the treatment but the physician's words and behavior are capable of causing injury.

Informed Consent Patients often require diagnostic and therapeutic procedures that are painful and that pose some risk. For many such procedures, patients are required to sign a consent form. The patient must understand clearly the risks entailed in these procedures; this is the definition of *informed consent*. It is incumbent on the physician to explain the procedures in a clear and understandable manner and to ascertain that the patient comprehends both the nature of the procedure and the attendant risks. The dread of the unknown that is inherent in hospitalization can be mitigated by such explanations.

Incurability and Death No problem is more distressing than that presented by the patient with an incurable disease, particularly when premature death is inevitable. What should the patient and family be told, what measures should be taken to maintain life, what can be done to maintain the quality of life, and how is death to be defined?

The concept of incurable illness and terminal care often evokes examples of cancer. However, patients with many other end-stage diseases including chronic obstructive pulmonary disease, congestive heart failure, renal or hepatic failure, and overwhelming infection face similar issues. The same principles of terminal care should be applied in each of these cases. Doing seemingly small things, focused on the needs of the patient, can do much to restore comfort or dignity during a person's final weeks or days. In the same way that pain should be attentively managed with analgesia, every effort should be made to alleviate shortness of breath and to provide good skin care.

Although some would argue otherwise, there is no ironclad rule that the patient must immediately be told "everything," even if the patient is an adult with substantial family responsibilities. How much is told should depend on the individual's ability to deal with

the possibility of imminent death; often this capacity grows with time, and whenever possible, gradual rather than abrupt disclosure is the best strategy. A wise and insightful physician is often guided by an understanding of what a patient wants to know and when he or she wants to know it. The patient's religious beliefs may also be taken into consideration. The patient must be given an opportunity to talk with the physician and ask questions. Patients may find it easier to share their feelings about death with their physician, who is likely to be more objective and less emotional, than with family members. As William Osler wrote:

One thing is certain; it is not for you to don the black cap and, assuming the judicial function, take hope away from any patient...hope that comes to us all.

Even when the patient directly inquires, "Am I dying?" the physician must attempt to determine whether this is a request for information or a demand for reassurance. Only open communication between the patient and the physician can resolve this question and guide the physician in what to say and how to say it.

The physician should provide or arrange for emotional, physical, and spiritual support and must be compassionate, unhurried, and open. There is much to be gained by the laying on of hands. Pain should be adequately controlled, human dignity maintained, and isolation from the family avoided. These aspects of care tend to be overlooked in hospitals, where the intrusion of life-sustaining apparatus can so easily detract from attention to the whole person and encourage concentration instead on the life-threatening disease, against which the battle will ultimately be lost in any case. In the face of terminal illness, the goal of medicine must shift from *cure* to *care*, in the broadest sense of the term. In offering care to the dying patient, the physician must be prepared to provide information to family members and to deal with their guilt and grief. It is important for the doctor to assure the family that everything possible has been done.

"Do Not Resuscitate" Orders and Cessation of Therapy When carried out in a timely and expert manner, cardiopulmonary resuscitation is often useful in the prevention of sudden, unexpected death. However, unless there are reasons to the contrary, this procedure should not be used merely to prolong the life of a patient with terminal, incurable disease. The decision whether or not to resuscitate or even to treat an incurably and terminally ill patient must be reviewed frequently and must take into consideration any unexpected changes in the patient's condition. In this context, the administration of fluids or food is considered therapy that may be withdrawn or withheld. These decisions must also take into account both the underlying medical condition, especially its reversibility, and the wishes of the patient, especially if these have been expressed in a living will or advance directive. If the patient's wishes cannot be ascertained directly, a close relative or another surrogate who can be relied on to transmit the patient's wishes and to be guided by the patient's best interests should be consulted. The patient's autonomy -- whether the choice is to continue or discontinue treatment or to be resuscitated or not in the event of a cardiopulmonary arrest -- must be paramount. The courts have ruled that competent patients may refuse therapy and that an incompetent patient's previously stated wishes regarding life support should therefore be respected. The issues involving death and dying are among the most difficult in medicine. In approaching them rationally and consistently, the physician must

combine both the science and the art of medicine.

THE EXPANDING ROLE OF THE PHYSICIAN

Genetics and Medicine The genomic era is likely to lead to a revolution in the practice of medicine. Obtaining the DNA sequence of the entire human genome may help to elucidate the genetic components of common chronic diseases -- hypertension, diabetes, atherosclerosis, many cancers, dementias, and behavioral and autoimmune disorders. This information should make it possible to determine individual susceptibility to these conditions early in life and to implement individualized prevention programs. Subclassification of many diseases on a genetic basis may allow the selection of appropriate therapy for each patient. As the response to drugs becomes more predictable, pharmacotherapy should become more rational. In short, the completion of the Human Genome Project is likely to lead to a substantial increase in physicians' ability to influence their patient's health and well-being.

Patients will be best served if physicians play an active role in applying this powerful, sensitive new information rather than being passive bystanders who are intimidated by the new technology. This is a rapidly evolving field, and physicians and other health care professionals must remain updated to apply this new knowledge. Genetic testing requires wise counsel based on an understanding of the value and limitations of the tests as well as the implications of their results for specific individuals.

Medicine on the Internet The explosion in use of the Internet through personal computers is having an important impact on many practicing physicians. The Internet makes a wide range of information available to physicians almost instantaneously at any time of the day or night and from anywhere in the world. This medium holds enormous potential for delivering up-to-date information, practice guidelines, state-of-the-art conferences, journal contents, textbooks (including this text), and direct communications with other physicians and specialists, thereby expanding the depth and breadth of information available to the physician about the diagnosis and care of patients. Most medical journals are now accessible on-line, providing rapid and comprehensive sources of information. Patients, too, are turning to the Internet in increasing numbers to derive information about their illnesses and therapies and to join Internet-based support groups. Physicians are increasingly challenged by dealing with patients who are becoming more sophisticated in their understanding of illness. At this time, there is one critically important caveat. Virtually anything can be published on the Internet, thus circumventing the peer-review process that is an essential feature of quality publications. Physicians or patients who search the Internet for medical information must be aware of this danger. Notwithstanding this limitation, appropriate use of the Internet is revolutionizing information access for physicians and is a positive force in the practice of medicine.

Delivering Cost-Effective Medical Care As the cost of medical care has risen, it has become necessary to establish priorities in the expenditure of resources. In some instances, preventive measures offer the greatest return for the expenditure; outstanding examples include vaccination, improved sanitation, reduction in accidents and occupational hazards, and biochemical- and DNA-based screening of newborns. For example, the detection of phenylketonuria by newborn screening may result in a net

saving of many thousands of dollars.

As resources become increasingly constrained, the physician must weigh the possible benefits of performing costly procedures that provide only a limited life expectancy against the pressing need for more primary care for those persons who do not have adequate access to medical services. For the individual patient, it is important to reduce costly hospital admissions as much as possible if total health care is to be provided at a cost that most can afford. This policy, of course, implies and depends on close cooperation among patients, their physicians, employers, payers, and government. It is equally important for physicians to know the cost of the diagnostic procedures they order and the drugs and other therapies they prescribe and to monitor both costs and effectiveness. The medical profession should provide leadership and guidance to the public in matters of cost control, and physicians must take this responsibility seriously without being or seeming to be self-serving. However, the economic aspects of health care delivery must not interfere with the welfare of patients. The patient must be able to rely on the individual physician as his or her principal advocate in matters of health care.

Accountability Medicine is a satisfying but demanding profession. Physicians must understand the characteristics of the populations they serve, and they must appreciate their patients' social and cultural attitudes to health, disease, and death. As the public has become more educated and more sophisticated regarding health matters, their expectations of the health system in general and of their physicians in particular have risen. Physicians are expected to maintain mastery of their rapidly advancing fields (the *science* of medicine) while considering their patient's unique needs (the *art* of medicine). Thus, physicians are held accountable not only for the technical aspects of the care that they provide but also for their patient's satisfaction with the delivery and costs of care.

In the United States, there are increasing demands for physicians to account for the way in which they practice medicine by meeting certain standards prescribed by federal and state governments. The hospitalization of patients whose health care costs are reimbursed by the government and other third parties is subjected to utilization review. Thus the physician must defend the cause for and duration of a patient's hospitalization if it falls outside certain "average" standards. Authorization for reimbursement is increasingly based on documentation of the nature and complexity of an illness, as reflected by recorded elements of the history and physical examination. The purpose of these regulations is both to improve standards of health care and to contain spiraling health care costs. This type of review is being extended to all phases of medical practice and is profoundly altering the practice of medicine. Physicians are also expected to give evidence of their continuing competence through mandatory continuing education, patient-record audits, recertification by examination, or relicensing.

Continued Learning The conscientious physician must be a perpetual student because the body of medical knowledge is constantly expanding and being refined. The profession of medicine should be inherently linked to a career-long thirst for new knowledge that can be used for the good of the patient. It is the responsibility of a physician to pursue continually the acquisition of new knowledge by reading, attending conferences and courses, and consulting colleagues and the Internet. This is often a difficult task for a busy practitioner; however, such a commitment to continued learning is an integral part of being a physician and must be given the highest priority.

Research and Teaching The title *doctor* is derived from the Latin *docere*, "to teach," and physicians should share information and medical knowledge with colleagues, with students of medicine and related professions, and with their patients. The practice of medicine is dependent on the sum total of medical knowledge, which in turn is based on an unending chain of scientific discovery, clinical observation, analysis, and interpretation. Advances in medicine depend on the acquisition of new information, i.e., on research, which often involves patients; improved medical care requires the transmission of this information. As part of broader societal responsibilities, the physician should encourage patients to participate in ethical and properly approved clinical investigations if they do not impose undue hazard, discomfort, or inconvenience. To quote Osler once more:

To wrest from nature the secrets which have perplexed philosophers in all ages, to track to their sources the causes of disease, to correlate the vast stores of knowledge, that they may be quickly available for the prevention and cure of disease -- these are our ambitions.

(Bibliography omitted in Palm version)

Back to Table of Contents

2. ETHICAL ISSUES IN CLINICAL MEDICINE - *Bernard Lo*

Physicians frequently confront ethical issues in clinical practice that are perplexing, time-consuming, and emotionally draining. Experience, common sense, and simply being a good person do not guarantee that physicians can identify or resolve ethical dilemmas. Knowledge about common ethical dilemmas is also essential.

FUNDAMENTAL ETHICAL GUIDELINES

Physicians should follow two fundamental but frequently conflicting ethical guidelines: respecting patient autonomy and acting in the patient's best interests.

RESPECTING PATIENT AUTONOMY

Competent, informed patients may refuse recommended interventions and choose among reasonable alternatives.

Informed Consent Informed consent requires physicians to discuss with patients the nature of the proposed care, the alternatives, the risks and benefits of each, the likely consequences, and to obtain the patient's agreement to care. Informed consent involves more than obtaining signatures on consent forms. Physicians need to educate patients, answer questions, make recommendations, and help them deliberate. Patients can be overwhelmed with medical jargon, needlessly complicated explanations, or too much information at once.

Nondisclosure of Information Physicians may consider withholding a serious diagnosis, misrepresenting it, or limiting discussions of prognosis or risks because they fear that a patient will develop severe anxiety or depression or refuse needed care. Patients should not be forced to receive information against their will. Most people, however, want to know their diagnosis and prognosis, even if they are terminally ill. Generally, physicians should provide relevant information, offer empathy and hope, and help patients cope with bad news.

Emergency Care Informed consent is not required when patients cannot give consent and when delay of treatment would place their life or health in peril. People are presumed to want such emergency care, unless they have previously indicated otherwise.

Futile Interventions Autonomy does not entitle patients to insist on whatever care they want. Physicians are not obligated to provide futile interventions that have no physiologic rationale or have already failed. For example, cardiopulmonary resuscitation would be futile in a patient with progressive hypotension despite maximal therapy. But physicians should be wary of using the term "futile" in looser senses to justify unilateral decisions to forego interventions when they believe that the probability of success is too low, no worthwhile goals can be achieved, the patient's quality of life is unacceptable, or the costs are too high. Such looser usages of the term are problematic because they may be inconsistent and mask value judgments.

ACTING IN THE BEST INTERESTS OF PATIENTS

The guideline of *beneficence* requires physicians to act for the patient's benefit. Laypeople do not possess medical expertise and may be vulnerable because of their illness. They justifiably rely on physicians to provide sound advice and to promote their well-being. Physicians encourage such trust. Hence, physicians have a fiduciary duty to act in the best interests of their patients. The interests of the patient should prevail over physicians' self-interest or the interests of third parties, such as hospitals or insurers. These fiduciary obligations of physicians contrast sharply with business relationships, which are characterized by "let the buyer beware," not by trust and reliance. The guideline of "*do no harm*" forbids physicians from providing ineffective interventions or acting without due care. This precept, while often cited, provides only limited guidance, because many beneficial interventions also have serious risks.

CONFLICTS BETWEEN BENEFICENCE AND AUTONOMY

Patients' refusals of care may thwart their own goals or cause them serious harm. For example, a young man with asthma may refuse mechanical ventilation for reversible respiratory failure. Simply to accept such refusals, in the name of respecting autonomy, seems morally constricted. Physicians can elicit patients' expectations and concerns, correct misunderstandings, and try to persuade them to accept beneficial therapies. If disagreements persist after discussions, the patient's informed choices and view of his or her best interests should prevail. While refusing recommended care does not render a patient incompetent, it may lead the physician to probe further to ensure that the patient is able to make informed decisions.

PATIENTS WHO LACK DECISION-MAKING CAPACITY

Patients may not be able to make informed decisions because of unconsciousness, dementia, delirium, or other conditions. Physicians should ask two questions regarding such patients: Who is the appropriate surrogate? What would the patient want done?

ASSESSING CAPACITY TO MAKE MEDICAL DECISIONS

All adults are considered legally competent unless declared incompetent by a court. In practice, physicians usually determine that patients lack the capacity to make health care decisions and arrange for surrogates to make them, without involving the courts. By definition, competent patients can express a choice and appreciate the medical situation, the nature of the proposed care, the alternatives, and the risks, benefits, and consequences of each. Their choices should be consistent with their values and should not result from delusions or hallucinations. Psychiatrists may help in difficult cases because they are skilled at interviewing mentally impaired patients and can identify treatable depression or psychosis. When impairments are fluctuating or reversible, decisions should be postponed if possible until the patient recovers decision-making capacity.

CHOICE OF SURROGATE

If a patient lacks decision-making capacity, physicians routinely ask family members to serve as surrogates. Most patients want their family members to be surrogates, and

family members generally know the patient's preferences and have the patient's best interests at heart. Patients may designate a particular individual to serve as proxy; such choices should be respected. Some states have established a prioritized list of which relative may serve as surrogate if the patient has not designated a proxy.

STANDARDS FOR SURROGATE DECISION MAKING

Advance Directives These are statements by competent patients to direct care if they lose decision-making capacity. They may indicate (1) what interventions they would refuse or accept or (2) who should serve as surrogate. Following the patient's advance directives, surrogate respects patients' autonomy.

Oral conversations are the most frequent form of advance directives. While such conversations are customarily followed in clinical practice, casual or vague comments may not be trustworthy.

Living wills direct physicians to forego or provide life-sustaining interventions if the patient develops a terminal condition or persistent vegetative state. Generally patients may refuse only interventions that "merely prolong the process of dying."

A health care proxy is someone appointed by the patient to make health care decisions if he or she loses decision-making capacity. It is more flexible and comprehensive than the living will, applying whenever the patient is unable to make decisions.

Physicians can encourage patients to provide advance directives, to indicate both what they would want and who should be surrogate, and to discuss their preferences with surrogates. In discussions with patients, physicians can ensure that advance directives are informed, up-to-date, and address likely clinical scenarios. Such discussions are best carried out in the ambulatory setting. The federal Patient Self-Determination Act requires hospitals and health maintenance organizations to inform patients of their right to make health care decisions and to provide advance directives.

Substituted Judgment In the absence of clear advance directives, surrogates and physicians should try to decide as the patient would under the circumstances, using all information that they know about the patient. While such substituted judgments try to respect the patient's values, they may be speculative or inaccurate. A surrogate may be mistaken about the patient's preferences, particularly when they have not been discussed explicitly.

Best Interests When the patient's preferences are unclear or unknown, decisions should be based on the patient's best interests. Patients generally take into account the quality of life as well as the duration of life when making decisions for themselves. It is understandable that surrogates would also consider quality of life of patients who lack decision-making capacity. Judgments about quality of life are appropriate if they reflect the patient's own values. Bias or discrimination may occur, however, if others project their values onto the patient or weigh the perceived social worth of the patient. Most patients with chronic illness rate their quality of life higher than their family members and physicians do.

Legal Issues Physicians need to know pertinent state laws regarding patients who lack decision-making capacity. A few state courts allow doctors to forego life-sustaining interventions only if patients have provided written advance directives or very specific oral ones.

Disagreements Disagreements may occur among potential surrogates or between the physician and surrogate. Physicians can remind everyone to base decisions on what the patient would want, not what they would want for themselves. Consultation with the hospital ethics committee or with another physician often helps resolve disputes. Such consultation is also helpful when patients have no surrogate and no advance directives. The courts should be used only as a last resort when disagreements cannot be resolved in the clinical setting.

DECISIONS ABOUT LIFE-SUSTAINING INTERVENTIONS

Although medical technology can save lives, it can also prolong the process of dying. Competent, informed patients may refuse life-sustaining interventions. Such interventions may also be withheld from patients who lack decision-making capacity on the basis of advance directives or decisions by appropriate surrogates. Courts have ruled that foregoing life-sustaining interventions is neither suicide nor murder.

MISLEADING DISTINCTIONS

People commonly draw distinctions that are intuitively plausible but prove untenable on closer analysis.

Extraordinary and Ordinary Care Some physicians are willing to forego "extraordinary" or "heroic" interventions, such as surgery, mechanical ventilation, or renal dialysis, but insist on providing "ordinary" ones, such as antibiotics, intravenous fluids, or feeding tubes. However, this distinction is not logical because all medical interventions have both risks and benefits. Any intervention may be withheld, if the burdens for the individual patient outweigh the benefits.

Withdrawing and Withholding Interventions Many health care providers find it more difficult to discontinue interventions than to withhold them in the first place. Although such emotions need to be acknowledged, there is no logical distinction between the two acts. Justifications for withholding interventions, such as refusal by patients or surrogates, are also justifications for withdrawing them. In addition, an intervention may prove unsuccessful or new information about the patient's preferences or condition may become available after the intervention is started. If interventions could not be discontinued, patients and surrogates might not even attempt treatments that might prove beneficial.

DO NOT RESUSCITATE (DNR) ORDERS

When a patient suffers a cardiopulmonary arrest, cardiopulmonary resuscitation (CPR) is initiated unless a DNR order has been made. Although CPR can restore people to vigorous health, it can also disrupt a peaceful death. After CPR is attempted on a general hospital service, only 14% of patients survive to discharge, and even fewer in

certain subgroups. DNR orders are appropriate if the patient or surrogate requests them or if CPR would be futile. To prevent misunderstandings, physicians should write DNR orders and the reasons for them in the medical record. "Slow" or "show" codes that merely appear to provide CPR are deceptive and therefore unacceptable. Although a DNR order signifies only that CPR will be withheld, the reasons that justify DNR orders may lead to a reconsideration of other plans for care.

ASSISTED SUICIDE AND ACTIVE EUTHANASIA

Proponents of these controversial acts believe that competent, terminally ill patients should have control over the end of life and that physicians should relieve refractory suffering. Opponents assert that such actions violate the sanctity of life, that suffering can generally be relieved, that abuses are inevitable, and that such actions are outside the physician's proper role. These actions are illegal throughout the United States, except that physician-assisted suicide is legal in Oregon under certain circumstances. Whatever their personal views, physicians should respond to patients' inquiries with compassion and concern. Physicians should elicit and address any underlying problems, such as physical symptoms, loss of control, or depression. Often, additional efforts to relieve distress are successful, and after this is done patients generally withdraw their requests for these acts.

CARE OF DYING PATIENTS

Patients often suffer unrelieved pain and other symptoms during their final days of life. Physicians may hesitate to order high doses of narcotics and sedatives, fearing they will hasten death. Relieving pain in terminal illness and alleviating dyspnea when patients forego mechanical ventilation enhances patient comfort and dignity. If lower doses of narcotics and sedatives have failed to relieve suffering, increasing the dose to levels that may suppress respiratory drive is ethically appropriate because the physician's intention is to relieve suffering, not hasten death. Physicians can also relieve suffering by spending time with dying patients, listening to them, and attending to their psychological distress.

CONFLICTS OF INTEREST

Acting in the patient's best interests may conflict with the physician's self-interest or the interests of third parties such as insurers or hospitals. The ethical ideal is to keep the patient's interests paramount. Even the appearance of a conflict of interest may undermine trust in the profession.

FINANCIAL INCENTIVES

In managed care systems, physicians may serve as gatekeepers or bear financial risk for expenditures. Although such incentives are intended to reduce inefficiency and waste, there is concern that physicians may withhold beneficial care in order to control costs. In contrast, physicians have incentives to provide more care than indicated when they receive fee-for-service reimbursement or when they refer patients to medical facilities in which they have invested. Regardless of financial incentives, physicians should recommend available care that is in the patient's best interests -- no more and

no less.

DENIALS OF COVERAGE

Utilization review programs designed to reduce unnecessary services may also deny coverage for care that the physician believes will benefit the patient. Physicians should inform patients when a plan is not covering standard care and act as patient advocates by appealing such denials of coverage. Patients may ask physicians to misrepresent their condition to help them obtain insurance coverage or disability. While physicians understandably want to help patients, such misrepresentation undermines physicians' credibility and violates their integrity.

GIFTS FROM PHARMACEUTICAL COMPANIES

Physicians may be offered gifts ranging from pens and notepads to lavish entertainment. Critics worry that any gift from drug companies may impair objectivity, increase the cost of health care, and give the appearance of conflict of interest. A helpful rule of thumb is to consider whether patients would approve if they knew physicians had accepted such gifts.

OCCUPATIONAL RISKS

Some health care workers, fearing fatal occupational infections, refuse to care for persons with HIV infection or multidrug-resistant tuberculosis. Such fears about personal safety need to be acknowledged, and institutions should reduce occupational risk by providing proper training, equipment, and supervision. Physicians should provide appropriate care within their clinical expertise, despite personal risk.

MISTAKES

Mistakes are inevitable in clinical medicine. They may cause serious harm to patients or result in substantial changes in management. Physicians and students may fear that disclosing such mistakes could damage their careers. Without disclosure, however, patients cannot understand their clinical situation or make informed choices about subsequent care. Similarly, unless attending physicians are informed of trainees' mistakes, they cannot provide optimal care and help trainees learn from mistakes.

LEARNING CLINICAL SKILLS

Learning clinical medicine, particularly learning to perform invasive procedures, may present inconvenience or risk to patients. To ensure patient cooperation, students may be introduced as physicians, or patients may not be told that trainees will be performing procedures. Such misrepresentation undermines trust, may lead to more elaborate deception, and makes it difficult for patients to make informed choices about their care. Patients should be told who is providing care, what benefits and burdens can be attributed to trainees, and how trainees are supervised. Most patients, when informed, allow trainees to play an active role in their care.

IMPAIRED PHYSICIANS

Physicians may hesitate to intervene when colleagues impaired by alcohol abuse, drug abuse, or psychiatric or medical illness place patients at risk. However, society relies on physicians to regulate themselves. If colleagues of an impaired physician do not take steps to protect patients, no one else may be in a position to do so.

CONFLICTS FOR TRAINEES

Medical students and residents may fear that they will receive poor grades or evaluations if they act on the patient's behalf by disclosing mistakes, avoiding misrepresentation of their role, and reporting impaired colleagues. Discussing such dilemmas with more senior physicians can help trainees check their interpretation of the situation and obtain advice and assistance.

ADDITIONAL ETHICAL ISSUES

MAINTAINING CONFIDENTIALITY

Maintaining the confidentiality of medical information respects patients' autonomy and privacy, encourages them to seek treatment and to discuss their problems candidly, and prevents discrimination. Physicians need to guard against inadvertent breaches of confidentiality, as when talking about patients in elevators. Maintaining confidentiality is not an absolute rule. The law may require physicians to override confidentiality in order to protect third parties, for example, reporting to government officials persons with specified infectious conditions, such as tuberculosis and syphilis; persons with gunshot wounds; and victims of elder abuse and domestic violence. Computerized medical records raise additional concerns because breaches of confidentiality may affect many patients.

ALLOCATING RESOURCES JUSTLY

Allocation of limited health care resources is problematic. Ideally, allocation decisions should be made as public policy, with physician input. At the bedside, physicians generally should act as patient advocates within constraints set by society, reasonable insurance coverage, and sound practice. *Ad hoc* rationing by the individual physician at the bedside may be inconsistent, discriminatory, and ineffective. In some cases, however, two patients may compete for the same limited resources, such as physician time or a bed in intensive care. When this occurs, physicians should ration their time and resources according to patients' medical needs and the probability of benefit.

ASSISTANCE WITH ETHICAL ISSUES

Discussing perplexing ethical issues with other members of the health care team, colleagues, or the hospital ethics committee often clarifies issues and suggests ways to improve communication and to deal with strong emotions. When struggling with difficult ethical issues, physicians may need to reevaluate their basic convictions, tolerate uncertainty, and maintain their integrity while respecting the opinions of others.

(Bibliography omitted in Palm version)

Back to Table of Contents

3. DECISION-MAKING IN CLINICAL MEDICINE - *Daniel B. Mark*

To the medical student who requires 2 h to collect a patient's history and perform a physical examination, and several additional hours to organize them into a coherent presentation, the experienced clinician's ability to reach a diagnosis and decide on a management plan in a fraction of the time seems extraordinary. While medical knowledge and experience play a significant role in the senior clinician's ability to arrive at a differential diagnosis and plan quickly, much of the process involves skill in clinical decision-making. The first goal of this chapter is to provide an introduction to the study of clinical reasoning.

Equally bewildering to the student are the proper use of diagnostic tests and the integration of the results into the clinical assessment. The novice medical practitioner typically uses a "shotgun" approach to testing, hoping to a hit a target without knowing exactly what that target is. The expert, on the other hand, usually has a specific target in mind and efficiently adjusts the testing strategy to it. The second goal of this chapter is to review briefly some of the crucial basic statistical concepts that govern the proper interpretation and use of diagnostic tests; quantitative tools available to assist in clinical decision-making will also be discussed.

CLINICAL DECISION-MAKING

CLINICAL REASONING

The most important clinical actions are not procedures or prescriptions but the judgments from which all other aspects of clinical medicine flow. In the modern era of large randomized trials, it is easy to overlook the importance of this elusive mental activity and focus instead on the algorithmic practice guidelines constructed to improve care. One reason for this apparent neglect is that much more research has been done on how doctors *should* make decisions (e.g., using a Bayesian model discussed below) than on how they actually *do*. Thus, much of what we know about clinical reasoning comes from empirical studies of nonmedical problem-solving behavior.

Despite the great technological advances of the twentieth century, uncertainty still plays a pivotal role in all aspects of medical decision-making. We may know that a patient does not have long to live, but we cannot be certain how long. We may prescribe a potent new receptor blocker to reverse the course of a patient's illness, but we cannot be certain that the therapy will do so without side effects. Uncertainty in medical outcomes creates the need for probabilities and other mathematical/statistical tools to help guide decision-making. (These tools are reviewed later in the chapter.)

Uncertainty is compounded by the information overload that characterizes modern medicine. Today's experienced clinician needs close to 2 million pieces of information to practice medicine. Doctors subscribe to an average of 7 journals, representing over 2500 new articles each year. Computers offer the obvious solution both for management of information and for better quantitation and management of the daily uncertainties of medical care. While the technology to computerize medical practice is available, many practical problems remain to be solved before patient information can be standardized and integrated with medical evidence on a single electronic platform.

The following three examples introduce the subject of clinical reasoning:

· A 46-year-old man presents to his internist with a chief complaint of hemoptysis. The physician knows that the differential diagnosis of hemoptysis includes over 100 different conditions, including cancer and tuberculosis (Chap. 33). The examination begins with some general background questions, and the patient is asked to describe his symptoms and their chronology. By the time the examination is completed, and even before any tests are run, the physician has formulated a working diagnostic hypothesis and planned a series of steps to test it. In an otherwise healthy and nonsmoking patient recovering from a viral bronchitis, the doctor's hypothesis would be that the acute bronchitis is responsible for the small amount of blood-streaked sputum the patient observed. In this case, a chest x-ray and purified protein derivative (PPD) skin test may be sufficient.

· A second 46-year-old patient with the same chief complaint who has a 100-pack-year smoking history, a productive morning cough, and episodes of blood-streaked sputum may generate the principal diagnostic hypothesis of carcinoma of the lung. Consequently, along with the chest x-ray andPPD skin test, the physician refers this patient for bronchoscopy.

· A third 46-year-old patient with hemoptysis who is from a developing country is evaluated with an echocardiogram as well, because the physician thinks she hears a soft diastolic rumble at the apex on cardiac auscultation, suggesting rheumatic mitral stenosis.

These three vignettes illustrate two aspects of expert clinical reasoning: (1) the use of cognitive shortcuts, or *heuristics*, as a way to organize the complex unstructured material that is collected in the clinical evaluation; and (2) the use of diagnostic hypotheses to consolidate the information and indicate appropriate management steps.

THE USE OF COGNITIVE SHORTCUTS

Heuristics reduce the complexity of a problem to a manageable level. Psychologists have found that people rely on three basic types of heuristics. For example, when assessing a patient, clinicians often weigh the probability that this patient's clinical features match those of the class of patients with the leading diagnostic hypotheses being considered. In other words, the clinician is searching for the diagnosis for which the patient appears to be a representative example; this cognitive shortcut is called the *representativeness heuristic*. It may take only a few characteristics from the history for an expert clinician using the representativeness heuristic to arrive at a sound diagnostic hypothesis. For example, an elderly patient with new-onset fever, cough productive of copious sputum, unilateral pleuritic chest pain, and dyspnea is readily identified as fitting the pattern for acute pneumonia, probably of bacterial origin. Evidence of focal pulmonary consolidation on the physical examination will increase the clinician's confidence in the diagnosis because it fits the expected pattern of acute bacterial pneumonia. Knowing this allows the experienced clinician to conduct an efficient, directed, and therapeutically productive patient evaluation although there may be little else in the history or physical examination of direct relevance. The inexperienced medical student or resident, who has not yet learned the patterns most prevalent in

clinical medicine, must work much harder to achieve the same result and is often at risk of missing the important clinical problem in a sea of compulsively collected but unhelpful data.

However, physicians using the representativeness heuristic can reach erroneous conclusions if they fail to consider the underlying prevalence of two competing diagnoses. Consider a patient with pleuritic chest pain, dyspnea, and a low-grade fever. A clinician might consider acute pneumonia and acute pulmonary embolism to be the two leading diagnostic alternatives. Clinicians using the representativeness heuristic might judge both diagnostic candidates to be equally likely, although to do so would be wrong if pneumonia was much more prevalent in the underlying population. Mistakes may also result from a failure to consider that a pattern based on a small number of prior observations will likely be less reliable than one based on larger samples.

A second commonly used cognitive shortcut, the *availability heuristic*, involves judgments made on the basis of how easily prior similar cases or outcomes can be brought to mind. For example, the experienced clinician may recall 20 elderly patients seen over the past few years who presented with painless dyspnea of acute onset and were found to have acute myocardial infarction. The novice clinician may spend valuable time seeking a pulmonary cause for the symptoms before considering and discovering the cardiac diagnosis. In this situation, the patient's clinical pattern does not fit the expected pattern of acute myocardial infarction, but experience with this atypical presentation, and the ability to recall it, can help direct the physician to the diagnosis.

Errors with the availability heuristic can come from several sources of recall bias. For example, rare catastrophes are likely to be remembered with a clarity and force out of proportion to their value, and recent experience is, of course, easier to recall and therefore more influential on clinical judgments.

The third commonly used cognitive shortcut, the *anchoring heuristic*, involves estimating a probability by starting from a familiar point (the anchor) and adjusting to the new case from there. For example, a clinician may judge the probability of colorectal cancer to be extremely high after an elevated screening carcinoembryonic antigen (CEA) result because the prediction of colorectal cancer is anchored to the test result. Yet, as discussed below, this prediction would be inaccurate if the clinical picture of the patient being tested indicates a low probability of disease (for example, a 30-year-old woman with no risk factors). Anchoring can be a powerful tool for diagnosis but is often used incorrectly (see "Measures of Disease Probability and Bayes' Theorem," below).

DIAGNOSTIC HYPOTHESIS GENERATION

Cognitive scientists studying the thought processes of expert clinicians have observed that clinicians group data into packets or "chunks," which are stored in their memories and manipulated to generate diagnostic hypotheses. Because short-term memory can typically hold only 7 to 10 items at a time, the number of packets that can be actively integrated into hypothesis-generating activities is similarly limited. The cognitive shortcuts discussed above play a key role in the generation of diagnostic hypotheses, many of which are discarded as rapidly as they are formed.

A diagnostic hypothesis sets a context for diagnostic steps to follow and provides testable predictions. For example, if the enlarged and quite tender liver felt on physical examination is due to acute hepatitis (the hypothesis), certain specific liver function tests should be markedly elevated (the prediction). If the tests come back normal, the hypothesis may need to be discarded or substantially modified.

One of the factors that makes teaching diagnostic reasoning so difficult is that expert clinicians do not follow a fixed pattern in patient examinations. From the outset, they are generating, refining, and discarding diagnostic hypotheses. The questions they ask in the history are driven by the hypotheses they are working with at the moment. Even the physical examination is driven by specific questions rather than a preordained checklist. While the student is palpating the abdomen of the alcoholic patient, waiting for a finding to strike him, the expert clinician is on a focused search mission. Is the spleen enlarged? How big is the liver? Is it tender? Are there any palpable masses or nodules? Each question focuses the attention of the examiner to the exclusion of all other inputs until answered, allowing the examiner to move on to the next specific question.

Negative findings are often as important as positive ones in establishing and refining diagnostic hypotheses. Chest discomfort that is not provoked or worsened by exertion in an active patient reduces the likelihood that chronic ischemic heart disease is the underlying cause. The absence of a resting tachycardia and thyroid gland enlargement reduces the likelihood of hyperthyroidism in a patient with paroxysmal atrial fibrillation.

While the representativeness and availability heuristics may play the major roles in shaping early diagnostic hypotheses, the acuity of a patient's illness can also be very influential. For example, clinicians are taught to consider aortic dissection routinely as a possible cause of acute severe chest discomfort along with myocardial infarction, even though the typical history of dissection is different from myocardial infarction and dissection is far less prevalent (Chap. 247). This recommendation is based on the recognition that a relatively rare but catastrophic diagnosis like aortic dissection is very difficult to make unless it is explicitly considered. If the clinician fails to elicit any of the characteristic features of dissection by history and finds equivalent blood pressures in both arms and no pulse deficits, he or she may feel comfortable in discarding the aortic dissection hypothesis. If, however, the chest x-ray shows a widened mediastinum, the hypothesis may be reinstated and a diagnostic test ordered [e.g., thoracic computed tomography (CT) scan, transesophageal echocardiogram] to evaluate it more fully. In noncritical situations, the prevalence of potential alternative diagnoses should play a much more prominent role in diagnostic hypothesis generation. The value of conducting a rapid systematic clinical survey of symptoms and organ systems to avoid missing important but inapparent clues cannot be overstated.

Because the generation and evaluation of appropriate diagnostic hypotheses is a skill that not all clinicians possess to an equal degree, errors in this process can occur, and in the patient with serious acute illness these may lead to tragic consequences. Consider the following hypothetical example. A 45-year-old male patient with a 3-week history of a "flulike" upper respiratory infection (URI) presented to his physician with symptoms of dyspnea and a productive cough. Based on the presenting complaint, the clinician pulled out a "URI Assessment Form" to improve quality and efficiency of care. The physician quickly completed the examination components outlined on this

structured form, noting in particular the absence of fever and a clear chest examination. He then prescribed an antibiotic for presumed bronchitis, showed the patient how to breathe into a paper bag to relieve his "hyperventilation," and sent him home with the reassurance that his illness was not serious. After a sleepless night with significant dyspnea unrelieved by rebreathing into a bag, the patient developed nausea and vomiting and collapsed. He was brought into the Emergency Department in cardiac arrest and could not be resuscitated. Autopsy showed a posterior wall myocardial infarction and a fresh thrombus in an atherosclerotic right coronary artery. What went wrong? The clinician decided, even before starting the history, that the patient's complaints were not serious. He therefore felt confident that he could perform an abbreviated and focused examination using the URI assessment protocol rather than considering the full range of possibilities and performing appropriate tests to confirm or refute his initial hypotheses. In particular, by concentrating on the "URI," the clinician failed to elicit the full dyspnea history, which would have suggested a far more serious disorder, and did not even search for other symptoms that could have directed him to the correct diagnosis.

This example illustrates how patients can diverge from textbook symptoms and the potential consequences of being unable to adapt the diagnostic process to real-world challenges. The expert, while recognizing that common things occur commonly, approaches each evaluation on high alert for clues that the initial diagnosis may be wrong. Patients often provide information that "does not fit" with any of the leading diagnostic hypotheses being considered. Distinguishing real clues from false trails can only be achieved by practice and experience. A less experienced clinician who tries to be too efficient (as in the above example) can make serious judgment errors.

MAJOR INFLUENCES ON CLINICAL DECISION-MAKING

More than a decade of research on variations in clinician practice patterns has shed much light on forces that shape clinical decisions. The use of heuristic "shortcuts," as detailed above, provides a partial explanation, but several other key factors play an important role in shaping diagnostic hypotheses and management decisions. These factors can be grouped conceptually into three overlapping categories: (1) factors related to physician personal characteristics and practice style, (2) factors related to the practice setting, and (3) economic incentive factors.

Practice Style Factors One of the key roles of the physician in medical care is to serve as the patient's agent to ensure that necessary care is provided at a high level of quality. Factors that influence this role include the physician's knowledge, training, and experience. It is obvious that physicians cannot practice evidence-based medicine if they are unfamiliar with the evidence. As would be expected, specialists generally know the evidence in their field better than do generalists. Surgeons may be more enthusiastic about recommending surgery than medical doctors because their belief in the beneficial effects of surgery is stronger. For the same reason, invasive cardiologists are much more likely to refer chest pain patients for diagnostic catheterization than are noninvasive cardiologists or generalists. The physician beliefs that drive these different practice styles are based on personal experience, recollection, and interpretation of the available medical evidence. For example, heart failure specialists are much more likely than generalists to achieve target angiotensin-converting enzyme (ACE) inhibitor

therapy in their heart failure patients because they are more familiar with what the targets are (as defined by large clinical trials), have more familiarity with the specific drugs (including dosages and side effects), and are less likely to overreact to foreseeable problems in therapy such as a rise in creatinine levels or symptomatic hypotension. Other intriguing research has shown a wide distribution of acceptance times of antibiotic therapy for peptic ulcer disease following widespread dissemination of the "evidence" in the medical literature. Some gastroenterologists accepted this new therapy before the evidence was clear (reflecting, perhaps, an aggressive practice style), and some gastroenterologists lagged behind (a conservative practice style, associated in this case with older physicians). As a group, internists lagged several years behind gastroenterologists.

The opinion of influential leaders can also have an important effect on practice patterns. Such influence can occur at both the national level (e.g., expert physicians teaching at national meetings) and the local level (e.g., local educational programs, "curbside consultants"). Opinion leaders do not have to be physicians. When conducting rounds with clinical pharmacists, physicians are less likely to make medication errors and more likely to use target levels of evidence-based therapies.

The patient's welfare is not the only concern that drives clinical decisions. The physician's perception about the risk of a malpractice suit resulting from either an erroneous decision or a bad outcome creates a style of practice referred to as *defensive medicine*. This practice involves using tests and therapies with very small marginal returns to preclude future criticism in the event of an adverse outcome. For example, a 40-year-old woman who presents with a long-standing history of intermittent headache and a new severe headache along with a normal neurologic examination has a very low likelihood of structural intracranial pathology. Performance of a headCT or magnetic resonance imaging (MRI) scan in this situation would constitute defensive medicine. On the other hand, the results of the test could provide reassurance to an anxious patient.

Practice Setting Factors Factors in this category relate to the physical resources available to the physician's practice and the practice environment. *Physician-induced demand* is a term that refers to the repeated observation that physicians have a remarkable ability to accommodate to and employ the medical facilities available to them. A classic early study in this area showed that physicians in Boston had an almost 50% higher hospital admission rate than did physicians in New Haven, despite there being no obvious differences in the health of the cities' inhabitants. The physicians in New Haven were not aware of using fewer hospital beds for their patients, nor were the Boston physicians aware of using less stringent criteria to admit patients.

Other environmental factors that can influence decision-making include the local availability of specialists for consultations and procedures, "high tech" facilities such as angiography suites, a heart surgery program, andMRImachines.

Economic Incentives Economic incentives are closely related to the other two categories of practice-modifying factors. Financial issues can exert both stimulatory and inhibitory influences on clinical practice. In general, physicians are paid on a fee-for-service, capitation, or salary basis (Chap. 4). In fee-for-service, the more the physician does, the more the physician gets paid. The incentive in this case is to do

more. When fees are reduced (discounted fee-for-service), doctors tend to increase the number of services billed for. Capitation, in contrast, provides a fixed payment per patient per year, encouraging physicians to take on more patients but to provide each patient with fewer services. Expensive services are more likely to be affected by this type of incentive than inexpensive preventive services. Salary compensation plans pay physicians the same regardless of the amount of clinical work performed. The incentive here is to see fewer patients. Recognizing these powerful shapers of physician behavior, managed care plans have begun to explore combinations of the three reimbursement types with the goal of improving individual physician productivity while restraining their use of expensive tests and therapies.

In summary, expert clinical decision-making can be appreciated as a complex interplay between cognitive devices used to simplify large amounts of complex information interacting with physician biases reflecting education, training, and experience, all of which are shaped by powerful, sometimes perverse, external forces. In the next section, we will review a set of statistical tools and concepts that can assist in making clinical decisions under uncertainty.

QUANTITATIVE METHODS TO AID CLINICAL DECISION-MAKING

The process of medical decision-making can be divided into two parts: (1) defining the available courses of action and estimating the likely outcomes with each, and (2) assessing the desirability of the outcomes. The former task involves integrating key information about the patient along with relevant evidence from the medical literature to create the structure of a decision problem. The remainder of this chapter will present some quantitative tools to assist the clinician in these activities. These tools can be divided into those that assist the clinician in making better outcome predictions, which are then used to make decisions, and those that support the decision process directly. While these tools are not yet used routinely in daily clinical practice, the computerization of medicine is creating the required substrate for their future widespread dissemination.

QUANTITATIVE MEDICAL PREDICTIONS

Diagnostic Testing The purpose of performing a test on a patient is to reduce uncertainty about the patient's diagnosis or prognosis and to aid the clinician in making management decisions. Although diagnostic tests are commonly thought of as laboratory tests (e.g., measurement of serum amylase level) or procedures (e.g., colonoscopy or bronchoscopy), any technology that changes our understanding of the patient's problem qualifies as a diagnostic test. Thus, even the history and physical examination can be considered a form of diagnostic test. In clinical medicine, it is common to reduce the results of a test to a dichotomous outcome, such as positive or negative, normal or abnormal. In many cases, this simplification results in the waste of useful information. However, such simplification makes it easier to demonstrate some of the quantitative ways in which test data can be used.

To characterize the accuracy of diagnostic tests, four terms are routinely used (Table 3-1). The *true-positive rate*, i.e., the sensitivity, provides a measure of how well the test correctly identifies patients with disease. The *false-negative rate* is calculated as (1-sensitivity). The *true-negative rate*, i.e., the specificity, reflects how well the test

correctly identifies patients without disease. The *false-positive rate* is (1- specificity). A perfect test would have a sensitivity of 100% and a specificity of 100% and would completely separate patients with disease from those without it.

Calculating sensitivity and specificity require selection of a cutpoint value for the test to separate "normal" from "diseased" subjects. As the cutpoint is moved to improve sensitivity, specificity typically falls and vice versa. This dynamic tradeoff between more accurate identification of subjects with versus those without disease is often displayed graphically as a receiver operating characteristic (ROC) curve. An ROC curve plots sensitivity (*y*-axis) versus 1 -specificity (*x*-axis). Each point on the curve represents a potential cutpoint with an associated sensitivity and specificity value. The area under the ROC curve is often used as a quantitative measure of the information content of a test. Values range from 0.5 (no diagnostic information at all, test is equivalent to flipping a coin) to 1.0 (perfect test).

In the diagnostic testing literature, ROC areas are often used to compare alternative tests. The test with the highest area (i.e., closest to 1.0) is presumed to be the most accurate. However, ROC curves are not a panacea for evaluation of diagnostic test utility. Like Bayes' theorem, they are typically focused on only one possible test parameter (e.g., ST segment response in a treadmill exercise test) to the exclusion of other potentially relevant data. In addition, ROC area comparisons do not simulate the way test information is actually used in clinical practice. Finally, biases in the underlying population used to generate the ROC curves (e.g., related to an unrepresentative test sample) can bias the ROC area and the validity of a comparison among tests.

Measures of Disease Probability and Bayes' Theorem Unfortunately, there are no perfect tests; after every test is completed the true disease state of the patient remains uncertain. Quantitating this residual uncertainty can be done with Bayes' theorem. This theorem provides a simple mathematical way to calculate the posttest probability of disease from three parameters: the pretest probability of disease, the test sensitivity, and the test specificity (Table 3-2). The pretest probability is a quantitative expression of the confidence in a diagnosis before the test is performed. In the absence of more relevant information it is usually estimated from the prevalence of the disease in the underlying population. For some common conditions, such as coronary artery disease (CAD), nomograms and statistical models have been created to generate better estimates of pretest probability from elements of the history and physical examination. The posttest probability, then, is a revised statement of the confidence in the diagnosis, taking into account both what was known before and after the test.

To understand how Bayes' theorem creates this revised confidence statement, it is useful to examine a nomogram version of Bayes' theorem that uses the same three parameters to predict the posttest probability of disease (Fig. 3-1). In this nomogram, the accuracy of the diagnostic test in question is summarized by the likelihood ratio for a positive test, which is the ratio of the true-positive rate to the false-positive rate [or sensitivity/(1 - specificity)]. For example, a test with a sensitivity of 0.90 and a specificity of 0.90 has a likelihood ratio of 0.90/(1 - 0.90), or 9. Thus, for this hypothetical test, a "positive" result is 9 times more likely in a patient with the disease than in a patient without it. The more accurate the test, the higher the likelihood ratio. However, if sensitivity is excellent but specificity is less so, the likelihood ratio will be substantially

reduced (e.g., with a 90% sensitivity but a 60% specificity, the likelihood ratio is 2.25). Most tests in medicine have likelihood ratios for a positive result between 1.5 and 20.

Consider two tests commonly used in the diagnosis ofCAD, an exercise treadmill and an exercise thallium-201 single photon emission CT (SPECT) test (Chap. 244). Meta-analysis has shown the treadmill to have an average sensitivity of 66% and an average specificity of 84%, yielding a likelihood ratio of 4.1 [0.66/(1 - 0.84)]. If we use this test on a patient with a pretest probability of CAD of 10%, the posttest probability of disease following a positive result rises only to about 30%. If a patient with a pretest probability of CAD of 80% has a positive test result, the posttest probability of disease is about 95%.

The exercise thalliumSPECTtest is a more accurate test for the diagnosis ofCAD. For our purposes, assume that it has both a sensitivity and specificity of 90%, yielding a likelihood ratio of 9.0 [0.90/(1 - 0.90)]. If we again test our low pretest probability patient and he has a positive test, using Fig. 3-1 we can demonstrate that the posttest probability of CAD rises from 10 to 50%. However, from a decision-making point of view, the more accurate test has not been able to improve diagnostic confidence enough to change management. In fact, the test has moved us from being fairly certain that the patient did not have CAD to being completely undecided (a 50:50 chance of disease). In a patient with a pretest probability of 80%, using the more accurate thallium SPECT test raises the posttest probability to 97% (compared with 95% for the exercise treadmill). Again, the more accurate test does not provide enough improvement in posttest confidence to alter management, and neither test has improved much upon what was known from clinical data alone.

If the pretest probability is low (e.g., £20%), even a positive result on a very accurate test will not move the posttest probability to a range high enough to rule in disease (e.g., ³80%). Conversely, with a high pretest probability, a negative test will not adequately rule out disease. Thus, the largest gain in diagnostic confidence from a test occurs when the clinician is most uncertain before performing it (e.g., pretest probability between 30 and 70%). For example, if a patient has a pretest probability forCAD of 50%, a positive exercise treadmill test will move the posttest probability to 80% and a positive exercise thalliumSPECTtest will move it to 90% (Fig. 3-1).

Bayes' theorem, as presented above, employs a number of important simplifications that should be considered. First, few tests have only two useful outcomes, positive or negative, and many tests provide numerous pieces of data about the patient. Even if these can be integrated into a summary result, multiple levels of useful information may be present (e.g., strongly positive, positive, indeterminate, negative, strongly negative). While Bayes' theorem can be adapted to this more detailed test result format, it is computationally complex to do so. Second, Bayes' theorem assumes that the information from the test is completely unique and nonoverlapping with information used to estimate the pretest probability. This independence assumption, however, is often wrong. In many cases, test results are correlated with patient characteristics. For example, the findings of cardiomegaly and pulmonary edema on chest x-ray are correlated with the historic features of heart failure and with the physical findings of a displaced left ventricular apical impulse, an S_3gallop, and rales. The unique predictive information contributed by the test in this case (the chest x-ray) is only a fraction of its

total information because much had already been learned about the probability of heart failure before the test was done.

Finally, it has long been thought that sensitivity and specificity are prevalence-independent parameters of test accuracy, and many texts still make this assertion. This statistically useful assumption, however, is clinically wrong. For example, a treadmill exercise test has a sensitivity in a population of patients with one-vesselCAD of around 30%, whereas the sensitivity in severe three-vessel CAD approaches 80%. Thus, the best estimate of sensitivity to use in a particular decision will often vary depending on the distribution of disease stages present in the tested population. A hospitalized population typically has a higher prevalence of disease and in particular a higher prevalence of more advanced disease stages than an outpatient population. As a consequence, test sensitivity will tend to be higher in hospitalized patients, whereas test specificity will be higher in outpatients.

Statistical Prediction Models Bayes' theorem, as presented above, deals with a clinical prediction problem that is unrealistically simple relative to most problems a clinician faces. Prediction models, based on multivariable statistical models, can handle much more complex problems and substantially enhance predictive accuracy for specific situations. Their particular advantage is the ability to take into account many overlapping pieces of information and assign a relative weight to each based on its unique contribution to the prediction in question. For example, a logistic regression model to predict the probability ofCADtakes into account all of the relevant independent factors from the clinical examination and diagnostic testing instead of the small handful of data that clinicians can manage in their heads or with Bayes' theorem. However, despite this strength, the models are too complex computationally to use without a calculator or computer (although this limit may be overcome when medicine is practiced from a fully computerized platform.) To date, only a handful of prediction models have been developed and properly validated. The importance of independent validation in a population separate from the one used to develop the model cannot be overstated. Unfortunately, most published models have not been properly validated, making their utility in clinical practice uncertain at best.

When statistical models have been compared directly with expert clinicians, they have been found to be more consistent, as would be expected, but not significantly more accurate. Their biggest promise, then, would seem to be to make less-experienced clinicians more accurate predictors of outcome.

DECISION SUPPORT TOOLS

DECISION SUPPORT SYSTEMS

Over the past 30 years, many attempts have been made to develop computer systems to help clinicians make decisions and manage patients. Conceptually, computers offer a very attractive way to handle the vast information load that today's physicians face. The computer can help by making accurate predictions of outcome, simulating the whole decision process, or providing algorithmic guidance. Computer-based predictions using Bayesian or statistical regression models inform a clinical decision but do not actually reach a "conclusion" or "recommendation." Artificial intelligence systems attempt to

simulate or replace human reasoning with a computer-based analogue. To date, such approaches have achieved only limited success. Reminder or protocol-directed systems do not make predictions but use existing algorithms, such as practice guidelines, to guide clinical practice. In general, however, decision support systems have shown little impact on practice. Reminder systems, although not yet in widespread use, have shown the most promise, particularly in correcting drug dosing and in promoting guideline adherence. The full potential of these approaches will only be achieved when computers are fully integrated into medical practice.

DECISION ANALYSIS

Compared with the methods discussed above, decision analysis represents a completely different approach to decision support. Its principal application is in decision problems that are complex and involve a substantial risk, a high degree of uncertainty in some key area, or an idiosyncratic feature that does not "fit" the available evidence. Three general steps are involved. First, the decision problem must be clearly defined. Second, the elements of the decision must be made explicit. This involves specifying the alternatives being considered, their relevant outcomes, the probabilities attached to each outcome, and the relative desirability (called "utility") of each outcome. Cost can also be assigned to each branch of the decision tree, allowing calculation of cost-effectiveness (Chap. 4).

An example of a decision tree used to evaluate strategies for management of the risk of infective endocarditis after catheter-associated *Staphylococcus aureus* bacteremia is shown in Fig. 3-2. Approximately 35,000 cases of *S. aureus* bacteremia occur each year in the United States. The development of complicating endocarditis, which occurs in about 6% of cases, is associated with high morbidity (31% mortality, 21% stroke rate) and medical costs. The three choices for management of the bacteremia are (1) transesophageal echocardiography (TEE), (2) a 4-week course of intravenous antibiotics (long-course), or (3) a 2-week course of intravenous antibiotics (short-course). In the TEE strategy, a 4-week course of antibiotics is given if endocarditis is evident and a 2-week course is given if it is not. With each strategy, there is a risk that the patient will develop endocarditis with or without major complications. In this analysis, the longest quality-adjusted survival (5.47 quality-adjusted life-years) was associated with the 4-week antibiotic course strategy, which also had the highest costs ($14,136 per patient), whereas the lowest costs ($9830 per patient) and worst outcomes (5.42 quality-adjusted life-years) were associated with the 2-week antibiotic course strategy. From a clinical point of view (ignoring costs), the 4-week antibiotic course was best. From a cost-effectiveness point of view, the TEE strategy (5.46 quality-adjusted life-years and $10,051 per patient costs) provided the best balance of added benefits and costs. Thus, decision analysis can be extremely helpful in clarifying tradeoffs in outcomes and costs in difficult management areas such as the above where it is highly unlikely that an adequate randomized trial will ever be done.

The data needed to fill in a decision tree (Fig. 3-2) are typically cobbled together from a variety of sources, including the literature (randomized trials, meta-analyses, observational studies) and expert opinion. Once the decision tree is finished, the decision is "analyzed" by calculating the average value of each limb of the tree. The decision arm with the highest net value (or expected utility) is the preferred choice. The

value of this exercise, however, is not so much in developing a prescription for action as it is in exploring the key elements and pressure points of a complex or difficult decision. The process of building the decision tree forces the analyst to be explicit about the choices being considered and all their relevant outcomes. Areas of high uncertainty are readily identified. Sensitivity analyses are an integral part of decision analysis and involve systematically varying the value of each key parameter in the model alone (one-way sensitivity analysis,) in pairs (two-way), or in higher combinations (multivariable) to assess the impact on choice of preferred management strategy. In the above example, varying the incidence of endocarditis resulting from *S. aureus* bacteremia from 3% to over 50% had no impact on the choice ofTEE as the preferred strategy.

User friendly personal computer-based software packages now make the creation and analysis of decision trees much more straightforward than in the past. However, the process is still too cumbersome and time-consuming to be used on a routine basis. When medicine is practiced from a fully computerized platform, a library of prestructured decision trees with user modifiable values can be made available to support practitioners working with individual patients.

CONCLUSIONS

In this era of evidence-based medicine, it is tempting to think that all the difficult decisions practitioners face have been or soon will be solved and digested into practice guidelines and computerized reminders. For the foreseeable future, however, such is not the case. Meta-analyses cannot generate evidence where there are no adequate randomized trials, and most of what clinicians face will never be thoroughly tested in a randomized trial. Excellent clinical reasoning skills and experience supplemented by well-designed quantitative tools and a keen appreciation for individual patient preferences will continue to be of paramount importance in the professional life of medical practitioners for years to come.

(Bibliography omitted in Palm version)

Back to Table of Contents

4. ECONOMIC ISSUES IN CLINICAL MEDICINE - *Daniel B. Mark*

The United States has the distinction of having some of the best medical care of any technologically advanced country. We have many of the best hospitals and doctors in the world. The research pipeline is full of significant new therapeutic advances, with revolutionary genetic-based therapies perhaps only a decade away. Our citizens largely subscribe to the principle that excellent medical care should be available to all, regardless of ability to pay. Yet we also have over 43 million people (most of them employed and earning minimal wages) without any health insurance and many more who are inadequately insured. Since the collapse of the Clinton health care reform efforts in 1994, U.S. health policy has been directed by marketplace forces that have created powerful and sometimes perverse incentives in medicine: Health insurance companies that use every available means to avoid insuring sick people; "managed care" programs that really only manage costs; doctors who are provided incentives to provide less medical care; and pharmaceutical companies that develop powerful and expensive new drugs priced beyond the reach of many of the elderly and chronically ill who need them most.

Facing such powerful and chaotic forces, physicians tend to focus narrowly on what they are most comfortable with, taking care of individual patients and conducting academic investigations. Many doctors consider economics too arcane for them to grasp and therefore do not even try. Consequently, when presented with economic arguments and evidence they are often unable to discriminate the legitimate from the fallacious. More importantly, they are ill equipped to defend their patients' interests in the crucible of cost containment that characterizes the modern managed care era.

This chapter has two goals: first, to provide a brief introduction to some of the larger economic forces that shape modern medical practices, and second, to introduce the economic tools that are used for assessing the value of medical practices, including cost effectiveness analysis.

HEALTH CARE SPENDING AND FINANCING

HOW MUCH IS SPENT ON HEALTH CARE?

In 1997, the United States spent $1.1 trillion on its health care system, representing 13.5% of the gross domestic product (GDP) (a crude measure of national income). Most of this ($969 billion) was spent on personal health care: 34% went to hospitals, 20% to physicians, 7% to nursing homes, and 8% to outpatient pharmaceuticals. In comparison, Canada and Western European countries spend a substantially smaller portion (6 to 10%) of their national income on health care but their citizens appear to be equally healthy, at least by crude metrics such as life expectancy and infant mortality rates. Economists and politicians have for years used such data to argue that the United States spends too much on health care. The issue of how much to spend is an inherently political one, however, and the discipline of economics has little to say about it.

WHO PAYS FOR HEALTH CARE?

Two major factors are continually driving up the costs of medical care: introduction into medical practice of new medical technologies (drugs, devices, procedures) that have a high price tag, and the aging of the U.S. population (since older people require more medical care than younger ones). These costs are distributed unevenly across society. In 1997, the government paid about 47% of the total national health care bill (75% federal, 25% states), private insurance paid about 32%, and individuals paid 17%. The government, of course, gets its money from taxpayers and uses the health care segment of its budget to pay for the Medicare and Medicaid programs (discussed below). To respond to rising medical costs, the government can increase taxes or redistribute funds from other programs such as defense and education. Neither of these options are politically attractive. Alternatively, because of its size in the medical marketplace, the government can impose lower prices on providers to make the available funds go farther (see "Cost-Containment Strategies," below). Much of the private insurance bill is subsidized by employers through their employee benefits packages. As medical costs go up, health insurance costs also rise and businesses must either pass on higher premium and copayment costs to their employees, raise their prices (potentially impairing their competitive position in the marketplace), or reduce their profit margin (a very unpopular move with stockholders). Like the government, businesses may also negotiate lower prices with health care providers and/or health insurance plans.

PUBLIC FINANCING OF HEALTH CARE

The public sector (i.e., government as an agent for society) finances the Medicare and Medicaid programs as well as the Veteran's Administration Hospital system, the Department of Defense military care system, the Public Health Service, and the Indian Health Service. Of these, Medicare is by far the largest and most influential, with 39 million people receiving health insurance at a total cost in 1997 of $214.6 billion (20% of total national health expenditures). The Medicare program was enacted in 1965 by Congress as an amendment to the Social Security Act of 1935 and was envisioned by President Lyndon B. Johnson as a first step toward universal health insurance in the United States, a key part of his "great society" plan. Its impact on the evolution of the U.S. health care system has been profound. The original congressional act provided health care insurance for the elderly (defined as those 65 and older) who were eligible for social security (i.e., retired workers who had paid into the system during their working years and their dependents). Amendments in 1972 extended coverage to the disabled of all ages (currently numbering around 5 million) and to patients with chronic renal failure (who currently number about 284,000).

Medicare consists of two related insurance programs. The Medicare Hospital Insurance Trust Fund (also known as Part A) covers hospital care and skilled nursing home care and is funded by compulsory federal payroll taxes on employers and employees. Medicare Part B, the Medical Supplementary Insurance Program, covers physician fees as well as laboratory and other diagnostic tests and is funded by general federal tax revenues and patient premiums. Both programs have substantial gaps in coverage, necessitating supplemental insurance (so-called Medigap policies) for those who can afford them. Because of its compulsory income redistribution feature, taking tax money from current workers to pay for health care for elderly citizens (many of whom are on fixed income close to the poverty level), Medicare is both a health insurance program

and a social welfare program designed to combat poverty in the disabled and elderly. In exchange for their tax money, the 150 million workers funding the program are promised the same type of social security when they become elderly (paid for by future generations of workers).

Medicaid is a social insurance program for the poor that is jointly run by the federal and state governments. The federal government gives each state a grant of money for the program based on that state's per capita income (in 1997, this amount totaled $95 billion), and the states pay for the rest ($65 billion in 1997). The program, like Medicare, was enacted by Congress in 1965 as a part of President Johnson's "great society" program. It is larger than Medicare in terms of eligible beneficiaries (41 million people) but smaller in terms of budget ($160 billion, or 12% of the total national health expenditures). Because the requirements to qualify for Medicaid are stringent, many low-income individuals under age 65 (especially the working poor) do not qualify. Eligibility criteria are set by each state within general federal guidelines, and the income and asset tests individuals must meet to qualify vary widely among states. Many of the dollars in the Medicaid program actually pay for care for elderly and disabled Medicare beneficiaries who also qualify for Medicaid on the basis of poverty.

PRIVATE FINANCING OF HEALTH CARE

Approximately 70% of the non-elderly U.S. population is covered by some form of private medical insurance. The feasibility of group insurance for medical care was initially demonstrated in the 1930s by Blue Cross, a franchise of nonprofit groups providing hospitalization insurance in order to help prop up the financially strapped U.S. hospital industry. Blue Shield, a separate organization modeled after Blue Cross, started providing insurance for in-hospital physician services in 1939. During World War II, employee wages were frozen by the government and to entice workers, who were in short supply, some employers started offering health insurance as a fringe benefit. With the feasibility of employer-sponsored group health insurance demonstrated by the experience of the "Blues," commercial insurers began to enter the market. To win the support of doctors and hospitals, insurers agreed to pay "reasonable and customary charges" and to defer all medical management decisions to doctors. This "fee-for-service" reimbursement system, created in the post-World War II era, sowed the seeds of the tremendous inflation observed in the U.S. medical system during the 1970s and 1980s.

The original focus of indemnity insurance plans was to cover individuals against catastrophic financial losses from high medical care bills. Insurance is a contract for protection against specific hazards that are unpredictable for individuals but can be defined with confidence for large groups. "Major medical" health insurance was designed to provide coverage for catastrophic illness, a relatively rare event in most populations. Group coverage is less expensive than individual coverage because it allows the insurance company to diffuse the risk of a large payout among a big pool of individuals who will pay premiums but make no claims. When coverage is shifted from a focus on rare catastrophes to routine maintenance medical care (comprehensive insurance policies), health insurance becomes a means for payment of expected rather than unexpected care. The consequence is higher health insurance premiums. The early appeal of health maintenance organizations (HMOs) was that they appeared to

offer an economically efficient way to provide routine preventive care and to manage the occasional catastrophic illness.

MANAGED CARE

Managed care is a generic term that embraces a wide spectrum of systems for integrating the financing and delivery of health care. Managed care organizations (MCOs) contract with doctors and hospitals to provide comprehensive care to enrolled members for a fixed, prospectively set, premium.HMOs are a form of managed care originally organized between the 1940s and 1960s as an alternative to the prevailing fee-for-service-based private insurance. With the advent of serious medical inflation in the 1970s, the HMO model was promoted by the federal government as a way to control the growth in medical spending. Early enthusiasm for this initiative was limited; in 1984, only 5% of individuals with employer-based health insurance were in an HMO. However, by 1998 that figure had risen to 85%. The exponential growth of managed care started in the 1990s in part as an employer-driven response to the uncontrolled medical inflation of the previous two decades.

The massive increase in demand for managed care by employers and by the Medicare program produced a rapid, and sometimes bewildering, evolution in the managed care industry. One important trend has been the growth of for-profit (i.e., investor owned) managed care companies. Over half ofHMOmembers now belong to a for-profit plan. Investment dollars from Wall Street have made it easier for these plans to respond quickly to increased employer demand for managed care options. However, compared with their not-for-profit counterparts, for-profit HMOs spend a smaller proportion of each premium dollar paying for health care for members (the paradoxically named "medical loss ratio"), since stockholders also have to be paid. As a result, for-profit HMOs are less successful than not-for-profit plans in providing preventive care (a presumed strength of managed care).

Another prevalent trend of the 1990s was the move from traditionalHMOmodels to virtual HMOs, built from contractual relationships with community physicians and hospitals. The three HMO models are the staff model, the group model, and the Independent Practice Association (IPA). The staff model HMO is a vertically integrated organization. That is, it owns its own hospitals, employs all its physicians full time for a set salary, and is focused in a particular geographic area. The group model HMO, exemplified by Group Health Cooperative of Puget Sound, contracts with one or more large multispecialty group practices to care for its patients for a preset capitated reimbursement. These physicians do not care for non-HMO patients. In the IPA model, the HMO contracts with an association of self-employed physicians who maintain their own offices and see both HMO and non-HMO patients. The network model refers to a hybrid of the other three forms of HMO. IPA and network model HMOs now have the majority of HMO membership in the United States.

The other portion of the managed care industry is represented by point of service (POS) plans and preferred provider organizations (PPOs). POS plans incorporate key features of bothHMOs and traditional fee-for-service plans. A patient may choose care from a provider network or go outside the network. Care within network requires only a minimal copayment, while care outside the network requires a deductible and a large (e.g., 30%)

copayment. The goal of the plan is to offer patients a choice but to provide major financial incentives to stay within the HMO portion of the plan. PPOs use a defined provider network (physicians, hospitals) that has agreed to accept discounted fee-for-service to care for enrolled members. PPOs may incorporate various managed care features, such as physician gatekeepers and utilization review.

THE UNINSURED AND UNDERINSURED

Data from the U.S. Census Bureau indicate that 43.4 million people had no health insurance for all of 1997 and 71.5 million people were without insurance for at least part of the year. The great majority of uninsured individuals either work for small employers who do not offer a health insurance benefit or, more commonly, cannot afford the premiums of the plan(s) that are offered. Underinsurance also has a significant impact on the working poor by requiring them to pay an excessive proportion of their family's income for health insurance premiums and out-of-pocket medical costs (deductibles, copayments, and uninsured care). Outpatient prescription medications are a major source of underinsurance. Prescription drug costs are now the fastest growing segment of the national medical budget and the least likely segment to be covered by insurance. The elderly are particularly affected, since Medicare does not currently cover outpatient prescriptions and even Medigap policies have limited coverage.

Some states have experimented with expanded coverage through their Medicaid programs to help the uninsured poor (such as the Oregon Medicaid program). For the forseeable future, however, it does not appear that the federal government will address this problem comprehensively.

COST-CONTAINMENT STRATEGIES

Current projections from the federal government's Health Care Finance Administration (HCFA) are that health care expenditures will double (to $2.2 trillion, or 16.2% of theGDP) by 2008. Over the past 30 years, the U.S. health care system has experimented with a vast array of cost-containment approaches. Conceptually, there are four major ways to control medical spending: (1) control prices, (2) control volume of care provided, (3) control the total budget available to pay for care, and (4) shift costs to another payer.

Two of the most important price control initiatives in medicine have been the Medicare Hospital Prospective Payment System and the Medicare Fee Schedule for physicians. In 1983, Medicare replaced its retrospective cost-based hospital reimbursement system with a prospective payment system. In this system, all hospitalizations are classified into one of approximately 500 Diagnosis Related Groups (DRGs) based on the principal discharge diagnosis for the hospitalization and a few selected additional factors such as age, the performance of surgery, and the presence of complications. Each DRG is assigned an average reimbursement (adjusted annually). If the hospital can provide care for less than this amount, they make a profit. If they spend more than this amount, they lose money. The DRG system was designed to promote efficiency and cost containment in hospital-based care. While it has helped to control Medicare costs, it has not reduced overall U.S. health care costs, probably because of substantial cost-shifting by hospitals to the private insurance sector.

Between 1975 and 1987, Medicare payments to physicians increased at an annual rate of 18%, well above the rate of inflation. While total spending for physician services accounts for less than 25% of the Medicare budget, physicians have control over aspects of care (use of procedures, length of stay, hospital admission) that extend their direct influence to over 75% of the Medicare budget. Recognizing the importance of physicians in cost containment, Congress directed the development of a new physician payment system based on the use of a resource-based relative value scale (RBRVS). The Medicare Fee Schedule, which was first used in 1992, has three components: (1) a measure of the total work (time and complexity) involved in each physician service and standardized across all specialties, (2) a practice expense to cover the cost of running an office, and (3) an amount to cover malpractice insurance costs. The Medicare Fee Schedule classifies all physician services using the American Medical Association's Current Procedural Terminology (CPT) codes. Each CPT code has an associated relative value units (RVUs) weight. The RVU weights are multiplied by a national conversion factor to generate the actual physician fee associated with the service in question.

Price controls are attractive for cost containment because they are less expensive administratively than volume controls and don't involve micromanagement of clinical care. Price controls alone, however, don't generally achieve control of costs because of compensatory responses of providers. For example, under Medicare prospective payment, hospitals have shifted much care to the outpatient setting, whereDRGs are not used. Physicians have responded to lower fees by an increased volume and intensity of service.

Volume controls include various programs to limit the diffusion of expensive technologies (such as heart surgery) or extra hospital beds. Limits can be operationalized using either a regulatory approach [such as certificate of need (CON) programs] or a budgetary approach. Utilization review approaches attempt to discern which expensive care items are medically necessary and which are not.

Budgetary controls are simpler than either price or volume control approaches. In Canada, for example, hospitals have global annual budgets. How the money is spent is decided by each hospital. If the budget is exceeded, there are no guarantees that the shortfall will be covered.

Finally, payers can control their costs by cost-shifting to other willing payers. For example, as health insurance premiums rise, employers can choose to pass these costs on to employees. Hospitals and doctors who lose money caring for Medicare patients can try to make up their losses by charging more to private insurance patients. Insurance companies can choose to offer limited or no coverage for outpatient pharmaceuticals, shifting the full cost of expensive new medicines directly to patients.

MEDICAL ECONOMIC CONCEPTS AND TOOLS

MEDICAL COST CONCEPTS

Medical cost analysis is a field that borrows heavily from both economics and

accounting. Economics provides the theoretical structure that defines the key questions to be addressed, and accounting provides many of the measurement tools. Traditional economics has as one of its major axioms that societal resources are finite. For this reason, society must choose from among the many ways that resources can be used and not all of society's goals can be fulfilled. Economics has devised a theoretical framework and a set of tools (including cost-effectiveness analysis) to help define the major competing goals for societal resources and to assist in selecting from among the ones that most efficiently fulfill societal needs. "Cost" in economics refers not so much to money but rather to the lost opportunities that occur when the limited societal resources are expended in a particular way. For example, if our medical armamentarium is enhanced over the next decade by discovery of powerful but expensive therapies and these are incorporated into standard clinical practice, the ability of the country to invest in education, defense, or transportation may be compromised. This notion of cost as a lost opportunity to use resources in alternative ways is referred to as *opportunity cost*. While representing the purest economic notion of cost, there is no practical way to measure it.

Accountants, who are much more concerned with issues of measurement, have proposed a "gold standard" of cost measurement, *true accounting cost*, that involves enumerating all the individual resources consumed in the production of a particular medical good or service and assigning market prices for each of them. The total cost is then the sum of the dollar costs for all the component resources. Even this calculation, however, may be prohibitively difficult in "real world" applications, for several reasons. First, all medical care requires not only the easily identifiable components of personnel time and disposable supplies but also the infrastructure components such as the rent on the office building where the care is provided, the cost of utilities, and the expense of an office staff. Second, even if all the components can be identified, enumeration of exactly what is used may be prohibitively expensive. Finally, medicine does not have publicly available "market prices" that can be readily obtained for a medical cost analysis, the way one can obtain prices for automobiles or refrigerators. The reasons for this relate to the lack of a true competitive free market in medicine along with the severe price distortion created in medical charges by cost-shifting practices.

KEY COST TERMS

Several key sets of cost terms are used in medicine. As the volume of health care produced is increased or decreased, costs may exhibit either variable or fixed "behavior." *Variable costs* change with each unit shift in production volume (up or down). For example, each vaccination administered to a group of children increases costs (related to the dose of vaccine and the disposable syringe) in a predictable linear fashion. *Fixed costs* do not shift with short-term changes in the volume of care provided. For example, the rent on the clinical building and the cost of heating, lighting, and so forth do not change according to the number of individuals vaccinated per day. Some types of costs display hybrid features of both variable and fixed components. For example, clinic personnel costs (e.g., nurses, secretaries) may be fixed if these personnel are paid a salary regardless of clinic volume. If the clinic volume goes up so much that evening hours must be added, either new personnel must be hired or existing personnel must work overtime. Either of these changes would graft a variable component onto the fixed personnel costs.

Marginal cost is a concept often used by economists to refer to the cost of producing one more unit of a given health care good or service. For example, the costs of doing one more or one less diagnostic cardiac catheterization would be its marginal cost. For all practical purposes, this is the same as its variable costs (since fixed costs do not change with small changes in volume). While the concept of unit changes in volume is theoretically interesting, a more pragmatic issue is the cost effect of changing a group of patients from one strategy to another. Many experts use the term *incremental costs* to refer to this type of shift (although some use marginal and incremental synonymously). Incremental analysis is a key component of cost-effectiveness analysis (see below).

Another set of cost terms relates to the traceability of costs to the production of health care goods and services. *Direct costs*, such as nursing and physician personnel and disposable supplies, can be clearly linked to the health care provided and are under the control of the health care providers. *Indirect costs*, sometimes known as *overhead*, cannot. For example, the utility, laundry, maintenance, and administration costs of a hospital cannot be linked with the care of an individual patient and are generally not under the control of the physicians and nurses providing the medical care. The distinction of direct versus indirect is useful in cost-containment efforts, where the first step is to identify all major cost components and decide how they are to be controlled.

One common error in the evaluation of medical costs is to focus on the cost of a test or therapy in isolation. Virtually every major medical management decision creates downstream consequences. For example, if physicians order a screening diagnostic test and the result is abnormal, they will need to do a confirmatory or more definitive test. If they order a potent new antibiotic and a fraction of patients develop liver failure as an unexpected toxicity, the total cost of that course of antibiotic includes not only the cost of the drug itself but also the costs of treating the liver failure in the fraction of patients who develop it. Extra costs added as a consequence of some diagnostic or therapeutic decision are referred to as *induced costs*. Similarly, if a management decision produces downstream savings, these would be referred to as *induced savings*. For example, administration of HMG CoA reductase inhibitors to patients with hypercholesterolemia can prevent future myocardial infarctions and revascularization procedures, both of which entail expensive hospitalizations.

One final important cost concept relates to the societal costs of lost productivity (primarily lost time from work) due to illness. While economists often refer to these as indirect costs, confusion with the accounting concept of indirect costs (overhead) has led many to prefer the alternative term, *productivity costs*.

COST MEASUREMENT

Using varying degrees of simplification, medical costs can be measured using either bottom-up or top-down approaches. Bottom-up approaches build from component resources to calculate total cost for an episode or type of care. Microcosting is the gold standard approach. It involves careful enumeration of all resources consumed and detailed cost-accounting estimation of the costs for each component resource. A number of medical centers have now installed computer-based cost-accounting systems that perform a modified type of microcosting analysis. For difficult-to-obtain resource

use data (such as time required for a particular type of care by a given type of personnel), these systems use expert opinion in place of empirical data. The other extreme of the bottom-up category of approaches involves enumeration and costing for only the "big ticket" or expensive items, such as hospitalization episodes and costly tests and procedures.

The top-down methods of medical cost estimation calculate a cost estimate from aggregated data. One such approach uses hospital billing charge data and charge-to-cost conversion ratios (which each hospital produces annually in its Medicare Cost Report) to estimate hospital costs. Despite the approximations involved, this approach, which can be used for most nonfederal U.S. hospitals, has provided good agreement with bottom-up estimates in the few instances where formal comparisons have been made. The other top-down approach is the use ofDRGassignments and reimbursement rates to provide standard cost weights for hospitalization episodes.

COST-EFFECTIVENESS ANALYSIS

Given a finite budget (for health care overall or for a particular health system), how can we use the available money to provide the most health benefits for our patients? For the clinician, who is less concerned with such policy issues, a prevalent question is whether a new treatment is economically attractive. The analysis method used to address this question is dependent on how the effectiveness and costs of the new therapy compare with those of "standard care" (Fig. 4-1). *Cost-effectiveness analysis* is used when effectiveness of the new treatment is greater and its costs are higher. This analysis calculates the ratio of added (or incremental) health benefits to added costs produced by a new therapy or strategy relative to some reference standard. The general formula is:

where C = costs and E= effectiveness.

The cost-effectiveness ratio provides a quantitative statement of the amount of money required to produce a single extra unit of benefit with the new therapy relative to usual care or some other relevant reference standard. The benefit can be calculated in any meaningful clinical unit, such as added survivors or extra patients with a correct diagnosis. However, the vast majority of cost-effectiveness analyses use the epidemiologic concept of life-years to express incremental benefit. Virtually all benchmarks for cost effectiveness relate to this endpoint. Because some therapies affect quality of life but not quantity, a more generally relevant effectiveness measure combines qualify of life and life expectancy into a single composite metric, the quality-adjusted life year (QALY). Calculation of incremental dollars required to add an extra QALY is called *cost-utility analysis*. The QALY is a useful concept, but many details regarding measurement and interpretation remain controversial. The third form of economic efficiency analysis, *cost-benefit analysis*, requires conversion of health benefits into monetary equivalents. Because such conversions are controversial, this form of analysis is rarely used in medicine. In theory, the time horizon of a cost-effectiveness analysis should be long enough to capture all important cost and health consequences of the therapy or strategy being evaluated. Most often, analysts

use a lifetime time frame. Because very few empirical studies are long enough to observe lifetime outcomes (especially when chronic diseases are being studied), models are required to extrapolate from available data.

A cost-effectiveness analysis can be done from a variety of perspectives, but the most widely applicable perspective is societal. Other perspectives are often much narrower and may include unattractive qualities. For example, a managed care organization may be interested only in short-term costs and outcomes, knowing that patients tend to change their health insurance every few years.

The benchmarks for cost-effectiveness ratios are determined by comparison with other well-accepted therapies in widespread medical use. A useful benchmark is hemodialysis for chronic renal failure, since the federal government has paid for all renal failure patients to get dialysis since 1973 through the End Stage Renal Disease Program. Recent estimates are that it costs this Medicare program about $50,000 to add 1 life-year to a chronic renal failure patient. Partly for this reason, many analysts use a cost-effectiveness ratio of <$50,000 per added life-year to identify therapies that are economically attractive (i.e., have a favorable balance of extra costs to extra benefits), while therapies with ratios >$100,000 per added life-year are deemed economically unattractive and therapies between $50,000 and $100,000 per added life-year are in the economic "gray zone."

Several caveats about cost-effectiveness analysis should be noted. First, cost-effectiveness analysis is descriptive, not prescriptive. It measures value that could be produced with available health care dollars but does not mandate how these dollars are to be used. If an expensive new therapy is introduced and is found to be very economically attractive by the above benchmarks, it will still not get used if there is no money in the budget to pay for it. Second, a cost-effectiveness ratio is only as good as the data that were used to calculate it. High-quality results can be obtained if economic analysis is prospectively incorporated into the design of large-scale multicenter randomized trials. Third, although cost-effectiveness ratios are often presented as deterministic (i.e., no variability), they often incorporate large amounts of uncertainty. This should be examined either with sensitivity analyses (varying each key parameter through a plausible range to see if the results are materially changed) or calculation of confidence limits.

MEDICAL ECONOMICS AND CLINICAL PRACTICE

In evaluating new therapies, three issues must be addressed: (1) is the new therapy significantly better than what is currently available? (2) how much does it cost and is it economically attractive? and (3) how many patients will need this therapy and is it affordable? The clinician should be primarily concerned with the answer to the first question. Although cost issues are now a reality of daily clinical life and cost-containment pressures are often substantial, decisions by clinicians that are based primarily on economic rather than clinical considerations put the physician in the role of the double agent (i.e., acting on behalf of both the patient and the payer) and compromise our fiduciary obligation to patients. The second question addresses cost effectiveness and, if favorable, can be used to support an argument by clinicians for adoption of the therapy. In the ideal world, at least, therapies that have a large database

of evidence demonstrating effectiveness and economic attractiveness should be given preference over therapies that do not have such supporting data. The final question is of primary concern to payers and health policy analysts. An effective therapy that is too expensive to use is of little more value than a therapy that has yet to be discovered.

(Bibliography omitted in Palm version)

Back to Table of Contents

5. INFLUENCE OF ENVIRONMENTAL AND OCCUPATIONAL HAZARDS ON DISEASE - *Howard Hu*, *Frank E. Speizer*

Exposures to hazardous materials and processes in the home, the workplace, and the community can cause or exacerbate a multitude of diseases. Physicians commonly treat the sequelae of such diseases in the practice of medicine; however, unless the underlying connection with hazardous exposures is identified and mitigated, treatment of manifestations rather than the cause at best only ameliorates the condition. At worst, the neglect of hazardous exposures may lead to both failure of treatment and failure to recognize a public health problem with wide significance.

No existing surveillance or reporting system can estimate the total contribution of hazardous exposures to morbidity and mortality. However, careful histories have identified occupational factors as etiologic in more than 10% of all admissions to general internal medicine wards in hospitals, with even higher percentages when the primary illness is either respiratory or musculoskeletal. Estimates of the number of new cases of disease due to work in the United States range from 125,000 to 350,000 per year; these cases do not include 5.3 million work-related injuries.

Environmental exposures are increasingly associated with decrements in measures of health whose outcomes range from subclinical to clinically catastrophic. For example, exposure to lead at levels that are common in the general population has been associated with increased blood pressure and decreased creatinine clearance. Ambient air pollution with respect to levels of ozone and fine-particulate matter has been related to increased rates of hospital admission for respiratory and cardiovascular diseases and to increased mortality rates, respectively. Indoor exposure to radon and passive indoor exposure to environmental tobacco smoke have been linked with an increased risk of lung cancer. There is pressure on clinicians to be aware of and act on this type of information, which is suggestive but not necessarily conclusive with respect to causation.

Patients are becoming increasingly concerned about hazardous exposures. More than 15% of patients seen in one study conducted in a primary care clinic expressed the opinion that their health problems were work-related, and 75% of this subgroup of patients reported exposure to one or more recognized toxic agents. Patients often want answers to very specific questions, such as: Is the water in our town safe to drink? Could my breathing problem be related to the new roofing sealant used in my building at work? Physicians are consulted because they are the most trusted sources of information on health risks, including chemical risks. Unfortunately, few physicians have more than rudimentary training in environmental and occupational medicine. Therefore, it becomes important for primary care physicians to be able to recognize symptoms precipitated by exposure to environmental or occupational hazards and either to manage these cases or to make appropriate referrals.

Many manifestations of exposure-related illnesses are nonspecific (e.g., dizziness, headache) or are commonly encountered in general internal medicine (e.g., myocardial infarction, cancer). The establishment of a connection with an environmental or occupational hazard requires a high index of suspicion and the application of fundamental concepts of environmental/occupational medicine. Furthermore, early

recognition by physicians of unusual patterns of illness or of evidence of asymptomatic exposure to toxins with low-level effects (e.g., an elevated blood lead level) can alert health officials to the need for control measures. Case reports either sent to local authorities or published in the literature often prompt follow-up studies that can lead to the identification of new hazards. In many states and countries, the reporting by physicians of occupational/environmental diseases is mandatory. For instance, beginning in 1992, physicians in Massachusetts were required to report cases of pneumoconiosis, occupational asthma, carpal tunnel syndrome, and carbon monoxide poisoning, among other conditions. Identification of an environmental/occupational etiology of an illness may have important economic ramifications for the patient (e.g., the awarding of worker's compensation, which covers medical bills as well as lost wages). Finally, physicians are frequently asked to provide expert medical testimony during litigation on the causal relationship between toxic exposures and diseases. In this setting, the more knowledgeable the physician is about potential hazardous exposures, the better prepared he or she is to serve the patient.

THE ENVIRONMENTAL/OCCUPATIONAL HISTORY

For a physician, the most critical steps toward recognizing these disorders are remembering to consider them in the differential diagnosis and taking an appropriate environmental/occupational history as part of the medical workup. The level of detail that is called for depends on the clinical situation. *Information should always be obtained on current and major past occupations, and patients should be asked whether they think their health problem is related to their work or to any particular environment or exposure.* In the review of systems, patients should be asked if they have been exposed to dusts, fumes, chemicals, radiation, or loud noise. When patient and physician are confronted with an illness of uncertain etiology, these factors should be explored in more detail, with the environmental/occupational history as the point of departure. (A brief outline of a sample history is shown in Table 5-1.)

The identification of specific chemical exposures can be difficult. Household products must list chemical ingredients on their labels, and this information may prove useful. For workplace exposures, the U.S. Occupational Safety and Health Administration (OSHA) requires chemical suppliers to provide material safety data sheets with their products and requires employers to retain these sheets and make them available to employees. The data sheets can be obtained by the physician or employee by a telephoned or written request; failure of an employer to provide them within 30 days of such a request is a violation of OSHA regulations and is punishable by fines. In addition to providing information on chemical ingredients and percent composition, the material safety data sheets provide basic information on toxicity. This information is seldom adequate from a clinical perspective but may indicate the general type of toxicity to be anticipated.

EVALUATION OF POSSIBLE CHEMICAL OR ENVIRONMENTAL HAZARDS

Given the wide variety of toxic exposures that may be uncovered during a workup, a clinician should routinely consult additional reference material to evaluate whether particular hazards may be associated with the illness at hand. Many sources of information exist.OSHA and some regional poison-control centers have extensive information on hazards and brief summary documents that can be transmitted over the

Internet or by telephone or facsimile. Depending on the area, other resources may include county and state health departments; regional offices of the National Institute for Occupational Safety and Health and the Environmental Protection Agency; the Consumer Products Safety Commission in Washington, DC; academic institutions; websites of these institutions; and individual toxicologists, occupational/environmental medicine specialists, or industrial hygienists. Sophisticated computerized databases are also available, including detailed listings on CD-ROM information systems. MEDLARS, the electronic database maintained by the National Library of Medicine, is accessible by modem or the Internet and is familiar to many physicians. Files other than MEDLINE, such as the Hazardous Substances Databank, provide specific toxicity information on chemicals and include toxicologic references not covered by MEDLINE. Many of these databases can also be accessed through the Internet.

As with any other illness, laboratory investigation may be crucial. For example, tests of carboxyhemoglobin level to document carbon monoxide exposure or of serum anticholinesterase level to document organophosphate pesticide absorption should be performed within hours of exposure. As in cases of acute drug overdose, it is useful to freeze samples of urine and serum from any patient suspected of having had an acute chemical exposure; such specimens can be analyzed at a later date by sensitive methods of detection. Use of other tests must rely on knowledge of the specific hazard or illness in question.

SUSPICIOUS SCENARIOS

Some medical problems or clinical scenarios demand a particularly high degree of suspicion of occupational or environmental factors as causative or contributing agents.

Respiratory Disease The contribution of occupational/environmental factors to respiratory disease is generally underrecognized, particularly among patients who smoke and among the elderly (Chap. 254). For instance, asthma related to chemical exposure may be treated without regard to cause or may be erroneously diagnosed as acute tracheobronchitis. A study of new-onset asthma among HMO members in Massachusetts found that 21% of these individuals met criteria for clinically significant asthma attributable to occupational exposures. The types of exposures and jobs in these cases varied widely; examples include exposure to smoke in a firefighter, to welding fumes in a technical school student, to cleaning compounds in a bartender, and to epoxy in an archery repairman. No single type of job or exposure predominated. Other examples of etiologic errors include shortness of breath from asbestosis that is attributed to chronic obstructive pulmonary disease and chemical pneumonitis that is misdiagnosed as a bacterial infection.

Cancer Many cancers are thought to be causally related to occupational and environmental factors in addition to tobacco. Some are particularly likely to have a chemical etiology or another environmental cause, including cancers of the skin (solar radiation, arsenic, coal tar, soot); lung (asbestos, arsenic, nickel, radon); pleura (almost exclusively asbestos); nasal cavity and sinuses (chromium, nickel, wood and leather dusts); liver (arsenic, vinyl chloride); bone marrow (benzene, ionizing radiation); and bladder (aromatic amines).

Coronary Disease and Hypertension Carbon monoxide exposure is common, particularly in homes with malfunctioning furnaces or in workplaces close to motor vehicle exhaust. By reducing oxygen transport by hemoglobin and inhibiting mitochondrial metabolism, carbon monoxide can aggravate coronary disease. Methylene chloride, a solvent used in paint stripping, is converted to carbon monoxide and thus poses the same risk. Exposure to carbon disulfide, a chemical used in the production of rayon, accelerates the rate of atherosclerotic plaque formation. Chronic lead exposure, even at modest levels, is a risk factor for the development of hypertension as well as abnormalities of cardiac conduction.

Hepatitis/Chronic Liver Disease In the absence of evidence that a viral infection, alcohol ingestion, or drug use is the main cause of hepatitis (Chaps. 295,296, and297), the involvement of a toxin must be considered. Toxin-induced hepatic injury may be cytotoxic, cholestatic, or both. The list of hepatotoxic agents is long, including organic synthetic compounds such as carbon tetrachloride (used in solvents and cleaning fluids) and methylene diamine (a resin hardener); pesticides such as chlordecone (Kepone); metals, particularly arsenic (used in pesticides and paints and found in well water); and natural toxins such as the pyrrolidizine alkaloids.

Kidney Disease Many chemical and environmental factors can cause renal injury (Chap. 269). The etiology of much chronic kidney disease, however, remains unknown. An increasing body of evidence now links chronic renal failure with hypertension to lead exposure. One study demonstrated that chelation therapy with EDTA slowed the progression of renal insufficiency in patients with a mildly elevated body lead burden. Some studies suggest that chronic exposure to hydrocarbons (e.g., gasoline, paints, solvents) may lead to various types of glomerulonephritis, including Goodpasture's syndrome. Environmental cadmium exposure has been found to promote calcium loss via urinary excretion, which results in skeletal demineralization and thus in an increased risk of fractures.

Peripheral Neuropathy Organic solvents such as *n*-hexane, heavy metals such as lead and arsenic, and some organophosphate compounds can damage the axons of peripheral nerves. Dimethylaminopropionitrile, an industrial catalyst, causes bladder neuropathy. Nerve entrapment syndromes of the upper extremity, such as carpal tunnel syndrome, may be caused by jobs that involve repetitive motion, especially those requiring the maintenance of awkward positions.

Central Nervous System Disorders Fatigue, memory loss, difficulty in concentration, and emotional lability have been linked to chronic exposure to solvents such as toluene and perchloroethylene. Painters, metal degreasers, plastics workers, and cleaners are commonly exposed to solvents and develop these symptoms at a high rate. Among the features that distinguish these patients are characteristic patterns on formal neurobehavioral testing and stabilization of symptoms with gradual improvement after discontinuation of the exposure. Other substances associated with neurobehavioral dysfunction include metals, particularly lead, mercury, arsenic, and manganese; pesticides, such as organophosphates and organochlorines; polychlorinated biphenyls (PCBs); and gases such as carbon monoxide.

Environmental factors are also suspected of contributing to other neurologic diseases,

such as degenerative disorders, motor neuron diseases, and extrapyramidal disorders. For example, a study in monozygotic and dizygotic twin pairs found a similarity in concordance indicating that environmental (as opposed to genetic) factors play a major etiologic role in cases of typical Parkinson's disease beginning after the age of 50 years.

Teratogenesis and Reproductive Problems Toxins can impair successful reproduction at a variety of levels. Examples include insecticides and herbicides,PCBs and polybrominated biphenyls (PBBs), ethylene oxide (a sterilizing gas used in hospitals), metals (lead, arsenic, cadmium, mercury), and solvents. Dibromochloropropane, a nematocide, suppresses spermatogenesis. Some toxins, such as PCBs, PBBs, and chlorinated pesticides, are concentrated in milk. Concern has arisen over the ability of specific organic pollutants, particularly pesticides, to persist in the environment and accumulate in human tissues. Some of these chemicals may disrupt endocrine function, and these effects may be related to phenomena such as the observed increases in the incidences of testicular cancer, breast cancer, and hypospadias.

Immunosuppression, Autoimmunity, and Hypersensitivity Evidence is increasing that exposures to some chemical agents can compromise the immune system, thereby leading to a generalized increase in the incidence of tumors (e.g., exposure toPBBs) or infections (e.g., respiratory infections after exposure to common air pollutants). Mercury, dieldrin, and methylcholanthrene are known to elicit autoimmune responses. Some chemicals are potent allergic sensitizers that cause dermal and respiratory problems (Chaps. 60 and254).

BIOLOGICAL MARKERS

An increasing number of methods are available for measuring and interpreting toxic exposure, including (1) the internal dose of specific toxins and (2) markers of the biologic effects of toxins. Internal-dose markers are relevant for toxins that are sequestered in the human body, such as lead (in blood), arsenic (in hair), and other metals (Chap. 395), and for halogenated compounds (such asPCBs). Examples of markers of the biologic effects of toxins include depressed levels of acetylcholinesterase in serum after exposure to organophosphate pesticides, sister chromatid exchanges in peripheral lymphocytes after exposure to the carcinogen ethylene oxide, and DNA adducts after exposure to tobacco smoke carcinogens.

MANAGING A HAZARD-RELATED ILLNESS

Once a chemical or another environmental hazard has been identified as an important contributor to an illness, the next step is to prevent further exposure. Although for chronic diseases such as cancer this step may be irrelevant for the patient in question, prevention of further exposure may still be critical for other persons who have been similarly exposed. When prevention of further exposure is important, *the physician must be willing to become an active advocate for the patient*. This advocacy may involve writing a letter stating that the patient should no longer be exposed to a hazard or should remain out of work. Alternatively, it may involve contacting appropriate officials in government, industry, or labor or other advocates who can deal with a hazardous exposure. Treatment is dependent on the specific hazard.

In few areas of medicine does a physician deal with more scientific uncertainty. Comprehensive information on toxicants is available for only a small percentage of chemicals. In general, the physician should take a conservative approach (i.e., advise the patient to avoid a hazard likely to have contributed to illness) and should use common sense and up-to-date information to evaluate causal relationships.

LOW-LEVEL EXPOSURES AND THEIR EFFECTS

The subclinical effects of toxins that are widespread in our environment and our workplaces are of increasing concern. Given the absence of any demonstrable effect threshold, low-level exposure to carcinogens should be avoided; not only carcinogenic but also noncarcinogenic effects of chronic low-level exposure to these substances are important.

Perhaps lead provides the most important example of low-level noncarcinogenic effects that constitute a major public health problem. Multiple pathways of exposure, including the combustion of leaded gasoline, the use of lead-based paints and solder, and the presence of lead in cans containing food, have contributed to exposure of the entire population. Such low-level exposures can impair neurobehavioral development in infants and children and can raise blood pressure in adults. Furthermore, absorbed lead is stored in the skeleton and may reenter the circulation at times of heightened bone turnover (e.g., pregnancy, lactation, osteoporosis, hyperthyroidism). Subclinical toxic effects can be prevented if chronic low-level exposure is detected early and curtailed. In the case of lead, such exposure is detected by tests of blood lead level, which should be performed regularly in young children living in old housing and as a precautionary measure in adults with a history of lead exposure.

(Bibliography omitted in Palm version)

Back to Table of Contents

6. WOMEN'S HEALTH - *Anthony L. Komaroff*, *Celeste Robb-Nicholson*, *Andrea E. Dunaif*

In recent years, the medical problems and health care of women have received increasing attention. There are poorly understood differences between men and women, both in morbidity and mortality and in the expression of diseases. Many research studies of disease prevention and pathophysiology have included only male subjects; most illnesses that can affect both sexes have not been as well studied in women. It also appears that women receive different care than men for certain common health problems. Finally, an increasing number of women are seeking health care in multidisciplinary women's health units that combine the expertise of gynecology, psychiatry, and internal medicine or family medicine.

MORBIDITY AND MORTALITY IN WOMEN

Morbidity Past studies have found that women experience more days of restricted activity than men at all ages, over and above the restricted activity caused by obstetric and gynecologic conditions. However, a study in 1998 concluded differently. Women make more visits to physicians, particularly for acute self-limited illnesses.

Mortality In the developed nations, women live longer than men. In the United States, as of 1996, the projected average life expectancy from birth is 79.1 years for females, and 73.1 years for males. Although there are more male fetuses conceived than female fetuses, females have a survival advantage when compared to males, in all age groups. The longer life expectancy of women versus men in developed countries is due in large part to the difference in mortality caused by ischemic heart disease (IHD).

As shown inTable 6-1, the leading causes of death among young women in the United States are accidents, homicide, and suicide. During the middle years, breast cancer is a slightly more common cause of death thanIHD and lung cancer. In women between ages 65 and 74, IHD, lung cancer, and cerebrovascular disease supercede breast cancer as the leading causes of death. Among women of all ages, IHD is the leading cause of death by a substantial margin, with a mortality rate five to sixfold higher than the rate for either lung or breast cancer. Nevertheless, polls find that U.S. women believe breast cancer poses the greatest threat to their lives.

Social Factors Influencing Morbidity and Mortality Gender differences in morbidity and mortality may be explained in part by psychosocial factors such as socially-defined gender roles, poverty, participation in the work force, health insurance, and lifestyle.

In the past 30 years in the United States, there has been a "feminization of poverty." One-third of families headed by women currently live in poverty, and the fraction is greater than one-half for African-American and Latino women. Almost a fifth of women over age 65 live below the poverty level. People of lower socioeconomic status experience poorer health and a higher mortality rate than those in higher income groups. The poor are more likely to smoke and less likely to have recommended preventive measures, including cancer screening. Lack of adequate health insurance is a major problem for many women; in general, they are more likely than men to have low-paying, part-time, non-union jobs that do not provide health insurance. Women who

are divorced or widowed may also lose health insurance that they had through their husbands.

PREVENTION (See alsoChap. 10)

Primary prevention and screening are crucial elements in improving the health of women. Based upon available literature and the consensus of experts, various authoritative organizations have published guidelines on preventive practices in women.

Most physicians believe that a baseline history and physical examination is useful to set the stage for preventive measures appropriate to each patient. In general, authorities recommend that blood pressure be measured every other year throughout life. Counseling on diet, smoking cessation, exercise, and use of seatbelts are of demonstrated value in the primary prevention of diseases and accidents. Counseling about safe sexual practices, alcohol abuse, and violence are also recommended.

Screening for glaucoma is recommended for African-American women over age 40 and for Caucasian women over age 50. Yearly examinations to test visual acuity are recommended for women over age 70.

Regular screening for breast, cervical, and colorectal cancer is recommended, but how often tests should be performed and which tools to use are still being debated. Most authorities recommend annual clinical breast examination in all women beginning at age 35 to 40. There is strong evidence to support the efficacy of annual mammography in women age 50 to 59. For women age 60 or older, the evidence for screening is less strong. The benefits of screening for women between the ages of 40 and 49 are still being debated.

Most authorities recommend Pap smear screening beginning at age 18 or when a woman becomes sexually active. After two or three consecutive normal Pap smears, most groups recommend Pap smear testing every three years. If Pap smears have been normal for 10 years, they can be discontinued in women after age 65.

Recommendations for colorectal cancer screening vary. For patients over 50, the American Cancer Society recommends yearly fecal occult blood testing and rectal examination combined with flexible sigmoidoscopy every 5 years, colonoscopy every 10 years, or double-contrast barium enema every 5 to 10 years.

Bone mineral testing has gained rapid acceptance as a screening tool for detecting osteoporosis, as well as for predicting the likelihood of the condition in the future. With the advent of multiple preventive and therapeutic strategies for osteoporosis, many authorities now recommend bone mineral testing to screen for the condition. A bone mineral density test is recomended for all women over age 65 as well as for all postmenopausal women who are at increased risk for developing osteoporosis (Chap. 342).

Cigarette smoking, a major risk factor for cardiovascular diseases and cancers in women, has been well studied (Chap. 390). Over the past 60 years there has been a sharp decline in smoking among men, but not among women; teenage women smoke at

higher rates than their male counterparts. "Low-yield" cigarettes are marketed heavily to women. The Nurses' Health Study showed that one-third of the excess risk of ischemic heart disease was eliminated two years after smoking cessation, and that all of the excess risk was eliminated by 10 to 14 years after smoking cessation.

The National Cholesterol Education Program recommends that total cholesterol and high-density lipoprotein (HDL) levels be measured once. If both are normal, a repeat test after 5 years is recommended. A meta-analysis of several small studies of women showed an increased risk ofIHD in women with serum cholesterol greater than 265, a ratio of total cholesterol to HDL cholesterol greater than 4, or an elevated fasting triglyceride.

In various case-control and observational studies, postmenopausal estrogen therapy is associated with a 40 to 50% reduction in deaths due toIHD, but its value in a prospective, randomized trial has not yet been documented.

Calcium and estrogen, as well as alendronate and the selective estrogen receptor modulators, tamoxifen and raloxifene, slow the development of osteoporosis and reduce the frequency of hip and vertebral fracture in postmenopausal women. In randomized clinical trials, both tamoxifen and raloxifene have been shown to reduce the risk of breast cancer in postmenopausal women.

Considerable research indicates that a relatively high dietary intake of various antioxidants (including vitamins E and C) is associated with lower rates of vascular disease and malignancies. Randomized trials of supplemental antioxidants are under way. Preliminary research indicates that regular aspirin use is associated with reduced rates ofIHD and colorectal carcinoma.

GENDER DIFFERENCES IN DISEASE

Obviously, some diseases and conditions occur exclusively (or nearly exclusively) in women -- e.g., menopause and various breast and gynecological disorders. These are discussed elsewhere in this book (Chaps. 52,89,336,337). In this chapter, we seek to highlight some gender differences in diseases that occur in both women and men.

Ischemic Heart Disease (See also Chap. 244) Many persons think ofIHD as a primary problem for men rather than women, perhaps because men have more than twice the total incidence of cardiovascular morbidity and mortality between the ages of 35 and 84. However, as stated earlier, in the United States IHD is among the leading causes of death among women as well as men (Table 6-1). The curve for the IHD mortality rate in women lags behind that for men by about a decade. Nevertheless, nearly 250,000 women die annually from IHD; after age 40, one in three women will die from heart disease. Although IHD mortality has been falling in men in the United States over the past 30 years, it has been increasing in women.

Why areIHDrates lower in women? They have a more favorable risk profile in some respects: higherHDLcholesterol levels, lower triglyceride levels, and less upper-body obesity than men. But women also have a less favorable risk profile in other respects: more obesity, higher blood pressure, higher plasma cholesterol levels, higher fibrinogen

levels, and more diabetes. The simplest explanation for the sex differential in IHD is the "cardioprotective" effect of estrogen, which can be due to improvement of the lipid profile, a direct vasodilatory effect, and perhaps other factors. HDL cholesterol levels appear to be a particularly important risk factor for IHD in women. HDL levels are higher in all age groups in women compared to men, and are higher in premenopausal and estrogen-treated postmenopausal women. Smoking is the most important risk factor for IHD in women.

IHDpresents differently in men and women. In the Framingham study, angina was the most frequent initial symptom of IHD in females, occurring in 47% of women, whereas myocardial infarction was the most frequent initial symptom in males, occurring in 46% of men. The exercise electrocardiogram has a substantial false positive as well as false negative rate for women, compared to men.

Women, particularly African-American women, have a higher risk of morbidity and mortality than men following a myocardial infarction. Compared to men, women who obtain coronary artery bypass graft surgery have more advanced disease, a higher perioperative mortality rate, less relief of angina, and less graft patency; however, 5- and 10-year survival rates are similar. Women undergoing percutaneous transluminal coronary angioplasty have lower rates of clinical and angiographic success than men, but also a lower rate of restenosis and a better long-term outcome. Women may benefit less and have more frequent serious bleeding complications from thrombolytic therapy than do men. Factors such as older age, more comorbid conditions, and more severeIHD in women at the time of events or procedures appear to account for at least part of the gender differences observed. Women with IHD benefit at least as much as men, and perhaps more, from reductions in cholesterol level.

The incidence ofIHDincreases markedly at menopause, consistent with the hypothesis that estrogens are cardioprotective. A number of observational studies have supported this hypothesis by demonstrating significant decreases in IHD in women on hormone replacement therapy (HRT), both estrogen alone and estrogen-progestin combination therapy. However, the HERS, a recent clinical trial of HRT for the *secondary* prevention of IHD, showed no significant difference in cardiovascular events between therapy with combined continuous conjugated equine estrogen (0.625 mg qd) and that with medroxyprogesterone acetate (2.5 mg qd), compared to placebo over four years. Indeed, in the HRT group, there was about a 50% increase in cardiovascular events in the first year of the trial. The Women's Health Initiative is investigating directly the impact of various HRT modalities as a *primary* prevention of IHD risk. Until further data are available, caution should be exercised in prescribing HRT to women with a history of IHD, or for cardioprotection alone.

Hypertension (See alsoChap. 246) Hypertension is more common in U.S. women than men, largely owing to the high prevalence of hypertension in older age groups and the longer survival rate for women. Both the effectiveness and the adverse effects of various antihypertensive drugs appear to be comparable in women and men. Benefits of treatment for severe hypertension have been dramatic in both women and men. However, in clinical trials of the treatment of mild to moderate hypertension, women have had a smaller decrease in morbidity and mortality than men, perhaps because women have a lower risk of myocardial infarction and stroke than men to begin with.

Older women benefit at least as much as men from treatment, as demonstrated by the Systolic Hypertension in Elderly study. The incidence of hypertension (above 140/90) appears to be low (less than 5%) with the current low-dose oral contraceptives. Postmenopausal estrogen therapy is not associated with increases in blood pressure.

Immunologically Mediated Diseases Several immunologically mediated diseases -- e.g., rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, Graves' disease, and thyroiditis -- occur much more frequently in women than in men. In animal models of rheumatoid arthritis -- lupus and multiple sclerosis, for example -- it is the females of the species that are predominantly affected. On the other hand, animal studies indicate that females are less susceptible to infection.

In short, female animals appear to have more vigorous immune responses, with both beneficial and adverse consequences. Increasing evidence indicates that estrogens upregulate both cellular and humoral immunity. Also, some immunocytes contain estrogen, progestin and androgen receptors, and the uterus produces a variety of cytokines, suggesting a complex interaction between the reproductive and immune systems.

Osteoporosis (See alsoChap. 342) This condition is much more prevalent in postmenopausal women than in men of similar age. Osteoporotic hip fractures are a major cause of morbidity in elderly women. Men accumulate more bone mass and lose bone more slowly than women. Gender differences in bone mass are found as early as infancy. Calcium intake, vitamin D and estrogen all play important roles in osteoporosis; calcium intake is an important determinant of peak bone mass, particularly during adolescence. Vitamin D deficiency is surprisingly common in elderly women. Receptors for estrogens and androgens have been identified in bone. The aromatase enzyme system, which converts androgens to estrogens, is also present in bone.

Therapy withHRT, or with calcium and vitamin D, has been shown to reduce the risk of osteoporotic fractures. Newer modalities, such as bisphosphonates (alendronate), calcitonin, and raloxifene, a selective estrogen receptor modulator, prevent bone loss and reduce the risk of osteoporotic fractures.

Alzheimer's Disease (See alsoChap. 362) Alzheimer's disease (AD) affects approximately twice as many women as men, in part because women live longer. Several observational studies suggest thatHRT may decrease the risk of AD and improve cognitive function in older women. These benefits are seen in both current as well as past HRT users. In a few experimental studies, estrogen replacement has been shown to be associated with improved memory compared to placebo treatment. Estrogens enhance neuronal growth and activity, providing a biologic basis for these putative cognitive effects of HRT. Prospective clinical trials, including the Women's Health Initiative, are underway to pursue these intriguing observations.

Diabetes Mellitus (See alsoChap. 333) Estrogens enhance insulin sensitivity in women but not in men. Despite this, the prevalence of type 2 diabetes mellitus (DM) is higher in women, which is related in part to the higher prevalence of female obesity. Premenopausal women with DM lose the cardioprotective effect of female gender and have identical rates ofIHD to those in males. This is partially explained by the presence

of several IHD risk factors in women with DM: obesity, hypertension and dysplipidemia. Recent evidence suggests that vascular responses differ in women with DM, as compared to normal women. Polycystic ovary syndrome and gestational diabetes mellitus -- common conditions in premenopausal women -- are associated with a significantly increased risk for type 2 DM.

Psychological Disorders (See also Chap. 385) Depression, anxiety panic disorder and eating disorders (bulimia and anorexia nervosa) occur more often in women than in men. Epidemiologic studies from both developed and developing nations consistently find major depression to be twice as common in women as in men, with the gender disparity becoming evident in early adolescence. Depression occurs in 10% of women during pregnancy and in 10 to 15% of women during the first several months of the postpartum period. The incidence of major depression diminishes after age 45, and does not increase with the onset of menopause. Depression in women also appears to have a worse prognosis than in men; episodes of depression last longer and there is a lower rate of spontaneous remission.

Social factors may account for the greater prevalence of some disorders in women; the traditionally subordinate role of women in society may generate feelings of helplessness and frustration which contribute to psychiatric illness. In addition, it is likely that biological factors, including hormonally influenced neurochemical changes, also play a role. The limbic system and hypothalamus -- areas of the brain thought to subserve appetite, satiety and emotion -- contain estradiol and testosterone receptors.

Alcohol and Drug Abuse (See alsoChap. 387) One-third of Americans who suffer from alcoholism are women. Women alcoholics are less likely to be diagnosed than men; a greater proportion of men than women seek help for alcohol and drug abuse. Men are more likely to go to an alcohol or drug treatment facility, while women tend to approach a primary care physician or mental health professional for help under the guise of a psychosocial problem. Late-life alcoholism is more common in women than men. In 1997, an epidemiologic survey reported that, among women over age 59, an estimated 1.8 million were addicted to or abused alcohol, and over 2.8 million were addicted to or abused psychoactive or mood-altering prescription drugs.

On average, alcoholic women drink less than alcoholic men, but exhibit the same degree of impairment. Blood alcohol levels are higher in women than in men after drinking equivalent amounts of alcohol, adjusted for body weight. This greater bioavailability of alcohol in women is probably due to the higher proportion of body fat and lower total body water. Women also have a lower gastric "first-pass metabolism" of alcohol, associated with lower activity of gastric alcohol dehydrogenase. In addition, alcoholic women are more likely than alcoholic men to abuse tranquilizers, sedatives, and amphetamines. Women alcoholics have a higher mortality rate than do nonalcoholic women and alcoholic men. Compared to men, women also appear to develop alcoholic liver disease and other alcohol-related diseases with shorter drinking histories and lower levels of alcohol consumption. Alcohol abuse also poses special risks to women who are or wish to become pregnant, adversely affecting fertility and the health of the baby (fetal alcohol syndrome).

Finally, there is growing evidence that for several illicit drugs, women proceed more

rapidly to drug dependence than do men.

Human Immunodeficiency Virus Infection (See also Chap. 309) As of September 1998, the Centers for Disease Control and Prevention estimate that between 120,000 and 160,000 adolescent and adult women in the United States were living with HIV infection, including those with AIDS (Table 6-1). Between 1985 and 1998, the proportion of all U.S. AIDS cases reported among women more than tripled, from 7 to 23%. HIV infection was the fourth leading cause of death among U.S. women age 25 to 44 in 1997, and the second leading cause of death among African-American women in this age group. The CDC estimates that 30% of the approximately 40,000 new HIV infections in the United States each year are among women.

Between 1996 and 1997 the incidence of new AIDS cases in the United States decreased by 18% and that of AIDS-related deaths by 42%, largely because of advances in HIV therapies. The decline continued between 1997 and 1998, albeit at a slower rate. AIDS incidence and AIDS-related mortality fell by 11 and 20%, respectively. However, AIDS incidence and deaths are not decreasing as rapidly among women as among men. HIV and AIDS continue to affect women in racial/ethnic minorities and lower socioeconomic classes disproportionately. CDC estimates that 64% of new HIV infections in 1998 occured among African-American women, 18% among Hispanic women, and 18% among white women. Of the new HIV infections among women in the United States in 1998, CDC estimates that 75% of women were infected through heterosexual sex and 25% of women through injection drug use.

Violence Against Women Violence against women in the United States is an enormous problem. Incidents of both rape and domestic violence are vastly underreported. Sexual assault is one of the most common crimes against women. One in five adult women in the United States reports having experienced sexual assault during her lifetime. Adult women are much more likely to be raped by a spouse, ex-spouse, or acquaintance than by a stranger.

Domestic violence is defined in the American Medical Association guidelines as "an ongoing, debilitating experience of physical, psychologic, and/or sexual abuse in the home, associated with increasing isolation from the outside world and limited personal freedom and accessibility to resources." It affects women of all ages, ethnic orientations, and socioeconomic groups. Based upon national crime statistics, every year an estimated 2 million women in the United States are severely injured and more than 1000 are killed by their current or former male partner. Domestic violence is the most common cause of physical injury in women, exceeding the combined incidence of all other types of injury (such as from rape, mugging, and auto accidents). Women who are young, single, pregnant, recently separated or divorced, or who have a history of substance abuse or mental illness, or a partner with substance abuse or mental illness, are at increased risk of domestic violence.

Domestic violence and sexual assault are associated with increased rates of physical and psychologic symptoms, medical office visits, and hospitalizations. Given this indirect presentation of the consequences of violence, and the high prevalence of unreported violence, clinicians should have a low threshold for pursuing the possibility of violence in female patients, particularly those with vague symptoms and psychological disorders.

The immediate treatment of rape and domestic violence focuses on assessing and treating physical injuries; providing emotional support; assessing and dealing with the risks of sexually transmitted infection and pregnancy; evaluating the safety of the patient and other family members; and documenting the patient's history and physical examination findings. In addition to dealing with the medical and psychological issues, appropriate care includes providing information about legal services, shelters and safe houses, hotlines, support groups, and counseling services.

RESEARCH IN WOMEN'S HEALTH

The growing recognition of the importance of women's health has spawned a number of research efforts, including large observational studies and clinical trials. The U.S. National Institutes of Health has introduced guidelines to mandate the inclusion of women in clinical studies, and the reporting of gender-specific data.

Studies of Prevention Large observational studies of men and women, such as the Rancho Bernardo Study and the Framingham Study, designed to analyze data specific to women have been on the increase. The Nurses' Health Study has been following more than 200,000 women, many for more than 20 years, prospectively collecting data to study the impact of smoking, diet, physical activity, medications, prevention and screening behaviors, and some psychosocial factors on the risk of various medical disorders, including breast cancer,IHD, stroke, diabetes, and fracture, as well as causes of mortality.

These studies have set the stage for clinical trials such as the Postmenopausal Estrogens/Progestins Intervention (PEPI) Trial, the first multicenter, randomized, double-blind, placebo-control trial of the effects of three estrogen/progestin regimens on risk factors for cardiovascular disease, bone mineral density, and endometrial hyperplasia. The study found that estrogen, alone or in combination with progestin, increased serum levels ofHDL and decreased low-density lipoprotein (LDL) and fibrinogen levels. While unopposed estrogen (without progestins) resulted in the most beneficial effects on lipids, it was also associated with an increased risk of endometrial hyperplasia.

In 1992, the NIH funded the Women's Health Initiative (WHI), a study of the health of postmenopausal women. The WHI, the largest research study ever funded by the NIH, involves over 160,000 postmenopausal women participating at 45 clinical centers across the United States through the year 2002. The WHI study includes both a prospective observational study and an interventional randomized trial involving over 63,000 women, which is designed to test the effects of a low-fat diet, hormone replacement therapy, and calcium and vitamin D supplementation on the risks for cardiovascular disease, breast cancer, and osteoporotic fractures.

Many other studies currently in progress promise new insights into the health of women within the next decade.

Pharmacologic Studies Historically, women have been underrepresented in drug trials, even though the majority of pharmaceuticals sold in the United States each year

are used by women. However, this has been rapidly changing. The FDA requires information on the safety and effectiveness of experimental drugs in women, on the effects of the menstrual cycle and menopause on a drug's pharmacokinetics, and on a drug's influence on the effectiveness of oral contraceptives. The increased emphasis on entering women into drug trials is likely to yield important information. Studies that have included women indicate that there are clinically significant differences in the way women respond to a number of frequently prescribed pharmaceuticals, including sedative-hypnotics, antidepressants, antipsychotics, anticonvulsants, and b-adrenergic blocking agents. The 1992 FDA Adverse Experience Report found that women have a higher frequency of adverse drug reactions than men. Other studies suggest that the efficacy of many drugs may be different in women compared to men. For example, women require lower doses of neuroleptics to control schizophrenia than men do. Women awaken from anesthesia faster than do men who are given the same doses of anesthetics, and they have a more powerful response to certain classes of analgesics than men. The reasons for these differences are not clear. However, these observations have spurred researchers to consider separating out the effects of gender in future clinical research in an effort to define "gender-based" biologic processes.

CONCLUSION

At the same time that the health of women is undergoing more rigorous study and women's clinics are becoming increasingly common and popular, a growing fraction of health professionals are women. The number of women physicians has increased by 300% between 1970 and 1990, and more than 40% of all U.S. medical students now are women. This infusion of women into the physician work force is likely to lead to a still greater recognition of the unique aspects of health and disease in women.

(Bibliography omitted in Palm version)

Back to Table of Contents

7. MEDICAL DISORDERS DURING PREGNANCY - *Robert L. Barbieri*, *John T. Repke*

Approximately 4 million births occur in the United States each year. A significant proportion of these are complicated by one or more medical disorders. Two decades ago, many medical disorders were contraindications to pregnancy. Advances in obstetrics, neonatology, obstetric anesthesiology, and medicine have increased the expectation that pregnancy will result in an excellent outcome for both mother and fetus despite most of these conditions. Successful pregnancy requires important physiologic adaptations, such as a marked increase in cardiac output. Medical problems that interfere with the physiologic adaptations of pregnancy increase the risk for poor pregnancy outcome; conversely, in some instances pregnancy may adversely impact an underlying medical disorder.

HYPERTENSION (See also Chap. 246)

In pregnancy, cardiac output increases by 40%, most of which is due to an increase in stroke volume. Heart rate increases by approximately 10 beats per minute during the third trimester. In the second trimester of pregnancy, systemic vascular resistance decreases and this is associated with a fall in blood pressure. During pregnancy, a blood pressure of 140/90 mmHg is considered to be abnormally elevated and is associated with a marked increase in perinatal morbidity and mortality. In all pregnant women, the measurement of blood pressure should be performed in the sitting position, because for many the lateral recumbent position is associated with a blood pressure lower than that recorded in the sitting position. The diagnosis of hypertension requires the measurement of two elevated blood pressures, at least 6 h apart. Hypertension during pregnancy is usually caused by preeclampsia, chronic hypertension, gestational hypertension, or renal disease.

PREECLAMPSIA

Approximately 5 to 7% of all pregnant women develop *preeclampsia*, the new onset of hypertension (blood pressure > 140/90 mmHg), proteinuria (>300 mg per 24 h), and pathologic edema. Although the precise placental factors that cause preeclampsia are unknown, the end result is vasospasm and endothelial injury in multiple organs. Preeclampsia is associated with abnormalities of cerebral circulatory autoregulation, which increase the risk of stroke at near-normal blood pressures. Risk factors for the development of preeclampsia include nulliparity, diabetes mellitus, a history of renal disease or chronic hypertension, a prior history of preeclampsia, extremes of maternal age (>35 years or <15 years), obesity, factor V Leiden mutation, angiotensinogen gene T235, antiphospholipid antibody syndrome, and multiple gestation.

There are no well-established strategies for the prevention of preeclampsia. Clinical trials have demonstrated that low-dose aspirin treatment does *not* prevent preeclampsia in either low- or high-risk women. Two meta-analyses reported that dietary calcium supplementation appeared to be effective in reducing the risk of developing preeclampsia. Subsequently, however, a large randomized clinical trial in low-risk women did not demonstrate a protective effect of calcium supplementation. Therefore, calcium supplementation may be considered in women at high risk for preeclampsia

(see above). The observation that dietary intervention may reduce the risk of hypertension in men and nonpregnant women raises the possibility that dietary manipulations will be discovered that reduce the risk of preeclampsia.

Severe preeclampsia is the presence of new-onset hypertension and proteinuria accompanied by central nervous system dysfunction (headaches, blurred vision, seizures, coma), marked elevations of blood pressure (>160/110 mmHg), severe proteinuria (>5 g per 24 h), oliguria or renal failure, pulmonary edema, hepatocellular injury (ALT >2´ the upper limits of normal), thrombocytopenia (platelet count < 100,000/uL), or disseminated intravascular coagulation. Women with *mild preeclampsia* are those with the diagnosis of new-onset hypertension, proteinuria, and edema without evidence of severe preeclampsia. The *HELLP* (*h*emolysis, *e*levated *l*iver enzymes, *l*ow *p*latelets) syndrome is a special subgroup of severe preeclampsia and is a major cause of morbidity and mortality in this disease. The presence of platelet dysfunction and coagulation disorders further increases the risk of stroke.

TREATMENT

Preeclampsia resolves within a few weeks after delivery. For pregnant women with preeclampsia prior to 37 weeks' gestation, delivery reduces the mother's morbidity but exposes the fetus to the risk of premature delivery. The management of preeclampsia is challenging because it requires the clinician to balance the health of both mother and fetus simultaneously and to make management decisions that afford both the best opportunities for infant survival. In general, prior to term, women with *mild* preeclampsia can be managed conservatively with bed rest, close monitoring of blood pressure and renal function, and careful fetal surveillance. For women with *severe* preeclampsia, delivery is recommended after 32 weeks' gestation. This reduces maternal morbidity and slightly increases the risks associated with prematurity for the newborn. Prior to 32 weeks' gestation, the risks of prematurity for the fetus are great, and some authorities recommend conservative management to allow for continued fetal maturation. Expectant management of severe preeclampsia remote from term affords some benefits for the fetus with significant risks for the mother. Such management should be restricted to tertiary care centers where maternal-fetal medicine, neonatal medicine, and critical care medicine expertise are available.

The definitive treatment of preeclampsia is delivery of the fetus and placenta. For women with severe preeclampsia, aggressive management of blood pressures> 160/110 mmHg reduces the risk of cerebrovascular accidents.

Intravenous labetalol or hydralazine are the drugs most commonly used to manage preeclampsia. Alternative agents such as calcium channel blockers may be used. Elevated arterial pressure should be reduced slowly to avoid hypotension and a decrease in blood flow to the fetus. *Angiotensin-converting enzyme (ACE) inhibitors as well as angiotensin-receptor blockers should be avoided in the second and third trimesters of pregnancy because of their adverse effects on fetal development.* Pregnant women treated with ACE inhibitors often develop oligohydramnios, which may be caused by decreased fetal renal function.

Magnesium sulfate is the treatment of choice for the prevention and treatment of

eclamptic seizures. Two large randomized clinical trials have demonstrated the superiority of magnesium sulfate over phenytoin and diazepam. Magnesium may prevent seizures by interacting with *N*-methyl-D-asparate (NMDA) receptors in the central nervous system. Given the difficulty of predicting eclamptic seizures on the basis of disease severity, it is recommended that once the decision to proceed with delivery is made, all patients carrying a diagnosis of preeclampsia be treated with magnesium sulfate (seeGuideline).

CHRONIC ESSENTIAL HYPERTENSION

Pregnancy complicated by chronic essential hypertension is associated with intrauterine growth restriction and increased perinatal mortality. Pregnant women with chronic hypertension are at increased risk for superimposed preeclampsia and abruptio placenta. Women with chronic hypertension should have a thorough prepregnancy evaluation, both to identify remediable causes of hypertension and to ensure that the prescribed antihypertensive agents are not associated with adverse pregnancy outcome (e.g.,ACEinhibitors, angiotensin-receptor blockers).a-Methyldopa and labetalol are the most commonly used medications for the treatment of chronic hypertension in pregnancy. Baseline evaluation of renal function is necessary to help differentiate the effects of chronic hypertension versus superimposed preeclampsia should the hypertension worsen during pregnancy. There are no convincing data that demonstrate that treatment of mild chronic hypertension improves perinatal outcome.

GESTATIONAL HYPERTENSION

This is the development of elevated blood pressure during pregnancy or in the first 24 h post partum in the absence of preexisting chronic hypertension and other signs of preeclampsia. Uncomplicated gestational hypertension that does not progress to preeclampsia has not been associated with adverse pregnancy outcome or adverse long-term prognosis.

RENAL DISEASE (See also Chap. 268)

Normal pregnancy is characterized by an increase in glomerular filtration rate and creatinine clearance. This occurs secondary to a rise in renal plasma flow and increase glomerular filtration pressures. Patients with underlying renal disease and hypertension may expect a worsening of hypertension during pregnancy. If superimposed preeclampsia develops, the additional endothelial injury results in a capillary leak syndrome that may make the management of these patients challenging. In general, patients with underlying renal disease and hypertension benefit from more aggressive management of blood pressure than do those with gestational hypertension. Preconception counseling is also essential for these patients so that accurate risk assessment can occur prior to the establishment of pregnancy and important medication changes and adjustments be made. In general, a prepregnancy serum creatinine level <133 umol/L (<1.5 mg/dL) is associated with a favorable prognosis. When renal disease worsens during pregnancy, close collaboration between the nephrologist and the maternal-fetal medicine specialist is essential so that decisions regarding delivery can be weighed in the context of sequelae of prematurity for the neonate versus long-term sequelae for the mother with respect to future renal function.

Successful pregnancy after renal transplantation has been reported increasingly. Predictors for success include a normal-functioning transplanted kidney, absence of rejection for at least 2 years prior to the pregnancy, absence of hypertension, and preferably minimal doses of immunosuppressant medications. Pregnancies in women using cyclosporine are more likely to be complicated by renal insufficiency and/or the development of hypertension. Such patients require very careful maternal and fetal surveillance. Nearly half of these pregnancies deliver preterm, and 20% of neonates are small for their gestational age. Rejection occurs in approximately 10% of pregnancies, and approximately 15% of patients will have deterioration in their renal function that persists after delivery. While pregnancy is generally well tolerated in renal transplant recipients, controversy remains as to whether or not deterioration of graft function is accelerated by pregnancy. More aggressive management of blood pressure has been suggested in this group of patients in an effort to protect the grafted kidney.

Another subset of patients with chronic renal disease and hypertension are those patients whose pregnancies are complicated by systemic lupus erythematosus (SLE) (Chap. 311). In the past, SLE was considered to be a contraindication to pregnancy. With improved understanding of the effects of SLE on pregnancy, and vice versa, and with improved pharmacologic methods for managing SLE, successful pregnancy outcome is likely. Good prognostic factors for establishment of pregnancy in the presence of SLE are as follows:

1. Disease quiescence> 6 months

2. Normal blood pressure (with or without medication)

3. Normal renal function [creatinine < 133 umol/L (< 1.5 mg/dL)]

4. Absence of antiphospholipid antibodies

5. Minimal or no need for immunosuppressive drugs

6. Absence of prior adverse reproductive outcome

Previously a point of controversy, there is now increasing consensus that pregnancy and the postpartum period are times of increased lupus activity. In severe flares early in gestation, pregnancy termination is often recommended. If pregnancy termination is not an option, then medical therapy to manage the lupus flare should not be influenced by the pregnancy, provided informed consent for treatment is obtained from the patient. Pulsed glucocorticoid therapy, azathioprine, hydroxychloroquine, and cyclophosphamide have all been used successfully in pregnancy.

CARDIAC DISEASE

VALVULAR HEART DISEASE (See alsoChap. 236)

This is the most common cardiac problem complicating pregnancy.

Mitral Stenosis This is the valvular disease most likely to cause death during pregnancy. The pregnancy-induced increase in blood volume and cardiac output can cause pulmonary edema in women with mitral stenosis. Pregnancy associated with long-standing mitral stenosis may result in pulmonary hypertension. Sudden death has been reported when hypovolemia has been allowed to occur in this condition. Careful control of heart rate, especially during labor and delivery, minimizes the impact of tachycardia and reduced ventricular filling times on cardiac function. Pregnant women with mitral stenosis are at increased risk for the development of atrial fibrillation and other tachyarrythmias. Medical management of severe mitral stenosis and atrial fibrillation with digoxin and beta blockers is recommended. Balloon valvulotomy can be carried out during pregnancy.

Mitral Regurgitation and Aortic Regurgitation These are both generally well tolerated during pregnancy. The pregnancy-induced decrease in systemic vascular resistance reduces the risk of cardiac failure with these conditions. As a rule, mitral valve prolapse does not present problems for the pregnant patient and aortic stenosis, unless very severe, is also well tolerated. In the most severe cases of aortic stenosis, limitation of activity or balloon valvuloplasty may be indicated.

For women with artificial valves contemplating pregnancy, it is important that warfarin be stopped and heparin initiated prior to conception. Warfarin therapy during the first trimester of pregnancy has been associated with fetal chondrodysplasia punctata. In the second and third trimester of pregnancy, warfarin may cause fetal optic atrophy and mental retardation.

CONGENITAL HEART DISEASE (See also Chap. 234)

The presence of a congenital cardiac lesion in the mother increases the risk of congenital cardiac disease in the newborn. Prenatal screening of the fetus for congenital cardiac disease with ultrasound is recommended. Atrial or ventricular septal defect is usually well tolerated during pregnancy in the absence of pulmonary hypertension, provided that the woman's prepregnancy cardiac status is favorable. Use of air filters on intravenous sets during labor and delivery in patients with intracardiac shunts is generally recommended.

OTHER CARDIAC DISORDERS

Supraventricular tachycardia (Chap. 230) is a common cardiac complication of pregnancy. Treatment is the same as in the nonpregnant patient, and fetal tolerance of medications such as adenosine and calcium channel blockers is acceptable. When necessary, electrocardioversion may be performed and is generally well tolerated by mother and fetus.

Peripartum cardiomyopathy (Chap. 238) is a rare disorder of pregnancy associated with myocarditis, and its etiology remains unknown. Treatment is directed toward symptomatic relief and improvement of cardiac function. Many patients recover completely; others are left with a progressive dilated cardiomyopathy. Recurrence in a subsequent pregnancy has been reported, and women should be counseled to avoid pregnancy after a diagnosis of peripartum cardiomyopathy.

SPECIFIC HIGH RISK CARDIAC LESIONS

Marfan Syndrome (See alsoChap. 351) This is an autosomal dominant disease, associated with a high risk of maternal morbidity. Approximately 15% of pregnant women with Marfan syndrome develop a major cardiovascular manifestation during pregnancy, with almost all women surviving. An aortic root diameter<40 mm is considered to be associated with a favorable outcome of pregnancy. Prophylactic therapy with beta blockers has been advocated, although large-scale clinical trials in pregnancy have not been performed.

Pulmonary Hypertension (See alsoChap. 260) Maternal mortality in the setting of severe pulmonary hypertension is high, and primary pulmonary hypertension is a contraindication to pregnancy. Termination of pregnancy may be advisable in these circumstances to preserve the life of the mother. In the Eisenmenger syndrome, i.e., the combination of pulmonary hypertension with right-to-left shunting due to congenital abnormalities (Chap. 234), maternal and fetal death occur frequently. Systemic hypotension may occur after blood loss, prolonged Valsalva maneuver, or regional anesthesia; sudden death secondary to hypotension is a dreaded complication. Management of these patients is challenging, and invasive hemodynamic monitoring during labor and delivery is generally recommended.

In patients with pulmonary hypertension, vaginal delivery is less stressful hemodynamically than Cesarean section, which should be reserved for accepted obstetric indications.

DEEP VENOUS THROMBOSIS AND PULMONARY EMBOLISM (See also Chaps. 248 and 261)

A hypercoagulable state is characteristic of pregnancy, and deep venous thrombosis (DVT) is a common complication. Indeed, pulmonary embolism is the most common cause of maternal death in the United States. Activated protein C resistance caused by the factor V Leiden mutation increases the risk for DVT and pulmonary embolism during pregnancy. Approximately 25% of women with DVT during pregnancy carry the factor V Leiden allele. The presence of the factor V Leiden mutation also increases the risk for severe preeclampsia. If the fetus carries a factor V Leiden mutation, the risk of extensive placental infarction is very high. Additional genetic mutations associated with DVT during pregnancy include the prothrombin G20210A mutation (heterozygotes and homozygotes) and the methylenetetrahydrofolate reductase C677T mutation (homozygotes).

TREATMENT

Aggressive diagnosis and management ofDVT and suspected pulmonary embolism optimize the outcome for mother and fetus. In general, all diagnostic and therapeutic modalities afforded the nonpregnant patient should be utilized in pregnancy. Anticoagulant therapy with heparin is indicated in pregnant women with DVT. Warfarin therapy is contraindicated in the first trimester due to its association with fetal chondrodysplasia punctata. In the second and third trimesters, warfarin may cause fetal

optic atrophy and mental retardation. In the initial treatment of DVT, heparin, which does not cross the placenta, may be administered as an intravenous bolus of approximately 100 IU per kilogram of body weight. Continuous heparin infusion is generally initiated at 1000 IU/h and then titrated to achieve a target activated partial thromboplastin time of 50 to 80 s. After initial intravenous anticoagulation, intermittent subcutaneous heparin therapy with 10,000 IU two or three times daily may be employed. When deep venous thromboembolism occurs in the postpartum period, heparin therapy for 7 to 10 days may be followed by warfarin therapy for 3 to 6 months. Warfarin is not contraindicated in breast-feeding women.

Low-molecular-weight heparins are of sufficient size and charge that they do not cross the placenta and may be substituted for unfractionated heparin in the pregnant patient. Recent concerns about low-molecular-weight heparin use and epidural hematoma suggest that caution be used in the anesthetic management of patients who had been receiving low-molecular-weight heparin near the onset of labor.

ENDOCRINE DISORDERS

DIABETES MELLITUS (See alsoChap. 333)

In pregnancy, the fetoplacental unit induces major metabolic changes, the purpose of which is to shunt glucose and amino acids to the fetus while the mother uses ketones and triglycerides to fuel her metabolic needs. These metabolic changes are accompanied by maternal insulin resistance, caused in part by placental production of steroids, a growth hormone variant, and placental lactogen. Although pregnancy has been referred to as a state of accelerated starvation, it is better characterized as accelerated ketosis. In pregnancy, after an overnight fast, plasma glucose is lower by 0.8 to 1.1 mmol/L (15 to 20 mg/dL) than in the nonpregnant state. This is due to the use of glucose by the fetus. In early pregnancy, fasting may result in circulating glucose concentrations in the range of 2.2 mmol/L (40 mg/dL) and may be associated with symptoms of hypoglycemia. In contrast to the decrease in maternal glucose concentration, plasma hydroxybutyrate and acetoacetate levels rise to two to four times normal after a fast.

TREATMENT

Pregnancy complicated by diabetes mellitus is associated with higher maternal and perinatal morbidity and mortality rates. Preconception counseling and treatment are important for the diabetic patient contemplating pregnancy. Optimizing preconception glucose control and attention to other dietary needs such as appropriate levels of folate can significantly reduce the risk of congenital fetal malformations. Folate supplementation reduces the incidence of fetal neural tube defects, which occur with greater frequency in fetuses of diabetic mothers. In addition, optimizing glucose control during key periods of organogenesis reduces other congenital anomalies including sacral agenesis, caudal dysplasia, renal agenesis, and ventricular septal defect.

Once pregnancy is established, glucose control should be managed more aggressively than in the nonpregnant state. In addition to dietary changes, this requires more frequent blood glucose monitoring and often involves additional injections of insulin or

conversion to an insulin pump. Fasting blood glucose levels should be maintained at<5.8 mmol/L (<105 mg/dL) with no values exceeding 7.8 mmol/L (140 mg/dL). Commencing in the third trimester, regular surveillance of maternal glucose control as well as assessment of fetal growth (obstetric sonography) and fetoplacental oxygenation (fetal heart rate monitoring or biophysical profile) optimize pregnancy outcome. Pregnant diabetic patients without vascular disease are at greater risk for delivering a macrosomic fetus, and attention to fetal growth via clinical and ultrasound examinations is important. Fetal macrosomia is associated with an increased risk of maternal and fetal birth trauma. Pregnant women with diabetes have an increased risk of developing preeclampsia, and those with vascular disease are at greater risk for developing intrauterine growth restriction, which is associated with an increased risk of fetal and neonatal death. Excellent pregnancy outcomes in patients with diabetic nephropathy and proliferative retinopathy have been reported with aggressive glucose control and intensive maternal and fetal surveillance.

Glycemic control may become more difficult to achieve as pregnancy progresses. Because of delayed pulmonary maturation of the fetuses of diabetic mothers, early delivery should be avoided unless there is biochemical evidence of fetal lung maturity. In general, efforts to control glucose and maintain the pregnancy until the estimated date of delivery result in the best overall outcome for both mother and newborn.

GESTATIONAL DIABETES

All pregnant women should be screened for gestational diabetes unless they are in a low-risk group. Women at low risk for gestational diabetes are those <25 years of age; those with a body mass index < 25 kg/m_2, no maternal history of macrosomia or gestational diabetes, and no diabetes in a first-degree relative; and those not members of a high-risk ethnic group (African American, Hispanic, Native American). A typical two-step strategy for establishing the diagnosis of gestational diabetes involves administration of a 50-g oral glucose challenge with a single serum glucose measurement at 60 min. If the serum glucose is < 7.8 mmol/L (<140 mg/dL), the test is considered normal. Serum glucose > 7.8 mmol/L (>140 mg/dL) warrants administration of a 100-g oral glucose challenge with serum glucose measurements obtained in the fasting state, and at 1, 2, and 3 h. Normal values are serum glucose concentrations<5.8 mmol/L (<105 mg/dL), 10.5 mmol/L (190 mg/dL), 9.1 mmol/L (165 mg/dL), and 8.0 mmol/L (145 mg/dL), respectively.

Pregnant women with gestational diabetes are at increased risk of preeclampsia, delivering infants who are large for their gestational age, and birth lacerations. Their fetuses are at risk of hypoglycemia and birth trauma (brachial plexus) injury.

TREATMENT

Gestational diabetes is first treated with dietary measures. Inability to maintain fasting glucose concentrations <5.8 mmol/L (<105 mg/dL) or 2-h postprandial glucose concentrations <6.7 mmol/L (<120 mg/dL) should prompt initiation of insulin therapy. Oral agents should not be used to treat diabetes in pregnancy. Patients with a diagnosis of gestational diabetes will benefit from postpartum follow-up as they are at increased risk for developing type 2 diabetes.

THYROID DISEASE (See alsoChap. 330)

In pregnancy, the estrogen-induced increase in thyroxine-binding globulin causes an increase in circulating levels of total T_3 and total T_4. The normal range of circulating levels of free T_4, free T_3, and thyroid stimulating hormone (TSH) remain unaltered by pregnancy.

The thyroid gland normally enlarges during pregnancy. Maternal hyperthyroidism occurs at a rate of approximately 2 per 1000 pregnancies and is generally well tolerated by pregnant women. Clinical signs and symptoms should alert the physician to the occurrence of this disease. Many of the physiologic adaptations to pregnancy may mimic subtle signs of hyperthyroidism. Although pregnant women are able to tolerate mild hyperthyroidism without adverse sequelae, more severe hyperthyroidism can cause spontaneous abortion or premature labor, and thyroid storm is associated with a significant risk of maternal mortality.

TREATMENT

Hyperthyroidism in pregnancy should be aggressively evaluated and treated. The treatment of choice is propylthiouracil. Because it crosses the placenta, the minimum effective dose should be used to maintain free T_4 in the upper normal range. Methimazole crosses the placenta to a greater degree than propylthiouracil and has been associated with fetal aplasia cutis. Radioiodine should not be used during pregnancy, either for scanning or treatment, because of effects on the fetal thyroid. In emergent circumstances, additional treatment with beta blockers and a saturated solution of potassium iodide may be necessary. Hyperthyroidism is most difficult to control in the first trimester of pregnancy and easiest to control in the third trimester.

The goal of therapy for *hypothyroidism* is to maintain the serumTSH in the normal range, and thyroxine is the drug of choice. Children born to women with an elevated serum TSH (and a normal total thyroxine) during pregnancy have impaired performance on neuropsychologic tests. During pregnancy, the dose of thyroxine required to keep the TSH in the normal range rises. In one study, the mean replacement dose of thyroxine required to maintain the TSH in the normal range was 0.1 mg daily before pregnancy, and it increased to 0.15 mg daily during pregnancy.

DISORDERS OF CALCIUM METABOLISM (See also Chap. 340)

Serum *total* calcium concentration decreases throughout gestation due to a reduction in serum albumin concentration, while serum *ionized* calcium remains unchanged during pregnancy. Circulating parathyroid hormone concentration is slightly reduced throughout the course of pregnancy. Pregnancy has been described as a state of physiologic absorptive hypercalciuria. Estrogen and increased production of 1,25-dihydroxyvitamin D by both the kidney and the placenta mediate the increased absorption of calcium during pregnancy. Due to the fetal requirements for calcium, the National Institutes of Health has recommended that pregnant women receive 1500 mg/d of elemental calcium, slightly higher than the recommended daily intake of 1200 mg/d for nonpregnant adults.

HEMATOLOGIC DISORDERS

Pregnancy has been described as a state of physiologic anemia. Part of the reduction in hemoglobin concentration is dilutional, but iron and folate deficiencies are the major causes of correctable anemia during pregnancy. Folic acid food supplementation implemented in 1998 has reduced the risk of fetal neural tube defects.

In populations at high risk for hemoglobinopathies (Chap. 106), hemoglobin electrophoresis should be performed as part of the prenatal screen. Hemoglobinopathies can be associated with increased maternal and fetal morbidity and mortality. Management is tailored to the specific hemoglobinopathy and is generally the same for both pregnant and nonpregnant women. Prenatal diagnosis of hemoglobinopathies in the fetus is readily available and should be discussed with prospective parents either prior to or early in pregnancy.

Thrombocytopenia occurs commonly during pregnancy. The majority of cases are benign gestational thrombocytopenias, but the differential diagnosis should include immune thrombocytopenia (Chap. 116) and preeclampsia. Maternal thrombocytopenia may also be caused by catastrophic obstetric events such as retention of a dead fetus, sepsis, abruptio placenta, and amniotic fluid embolism.

NEOPLASTIC DISEASES

Maternal neoplasms are rarely, if ever, transmitted to the fetus. The three most common cancers in pregnant women are cervical cancer (~1 case per 1000 pregnancies, depending on the country), breast cancer (~2 cases per 10,000 pregnancies), and lymphomas (Hodgkin's disease or non-Hodgkin's lymphomas). Cervical cancer may be missed when its early sign, vaginal bleeding, is attributed to the pregnancy. Pregnant women with vaginal bleeding should be examined, and suspicious cervical lesions biopsied. Conization is generally performed only after the first trimester because of the abortion risk.

Breast lumps may also be attributed to change associated with pregnancy. However, women with a dominant mass should undergo diagnostic evaluation (mammogram, ultrasound, biopsy). Resection of the primary lesion is safe, but radiation therapy is unsafe at any time during pregnancy. The fetus cannot be shielded from internal scattering of radiation; therapeutic doses are associated with spontaneous abortion, increased perinatal death, and defects in central nervous system and/or cognitive function. Tamoxifen is not safe for pregnant women.

Lymphoma is usually diagnosed on the basis of adenopathy or constitutional symptoms (fever, sweats, or weight loss). Staging evaluation is not undertaken during the first trimester; women in the first trimester should be counseled about termination of the pregnancy. Single-agent chemotherapy can be used in the second or third trimester as a temporizing measure. Vinblastine or doxorubicin have been used most commonly. Early induction of labor may permit the physician to maximize the survival chances of both the fetus and the mother. Survival rates for 28-week-old fetuses are about 75% and about 90% for 32-week-old fetuses.

Cancer survivors of reproductive age may desire children. Pregnancy may increase the risk of melanoma recurrence but does not influence breast cancer recurrence. Cancer treatment may deplete oocytes. Oocyte retrieval and storage of fertilized or nonfertilized eggs before cancer treatment may permit conception after the cancer has been treated successfully.

GASTROINTESTINAL AND LIVER DISEASE

Up to 90% of pregnant women experience nausea and vomiting during the first trimester of pregnancy. Occasionally, hyperemesis gravidarum requires hospitalization to prevent dehydration, and sometimes parenteral nutrition is required.

Crohn's disease may be associated with exacerbations in the second and third trimesters. Ulcerative colitis is associated with disease exacerbations in the first trimester and during the early postpartum period. Medical management of these diseases during pregnancy is identical to the management in the nonpregnant state (Chap. 287).

Exacerbation of gall bladder disease is commonly observed during pregnancy. In part this may be due to pregnancy-induced alteration in the metabolism of bile and fatty acids. Intrahepatic cholestasis of pregnancy is generally a third-trimester event. Profound pruritus may accompany this condition and may be associated with increased fetal mortality. It has been suggested that placental bile salt deposition may contribute to progressive uteroplacental insufficiency. Therefore, regular fetal surveillance should be undertaken once the diagnosis of intrahepatic cholestasis is made. Favorable results with ursodiol have been reported.

Acute fatty liver is a rare complication of pregnancy. Frequently confused with theHELLPsyndrome (see "Preeclampsia," above) and severe preeclampsia, the diagnosis of acute fatty liver of pregnancy may be facilitated by imaging studies and laboratory evaluation. Acute fatty liver of pregnancy is generally characterized by markedly increased levels of bilirubin and ammonia and by hypoglycemia. Management of acute fatty liver of pregnancy is supportive; recurrence in subsequent pregnancies has been reported.

All pregnant women should be screened for hepatitis B. This information is important for pediatricians after delivery of the infant. All infants receive hepatitis B vaccine. Infants born to mothers who are carriers of hepatitis B surface antigen should also receive hepatitis B immune globulin as soon after birth as possible and preferably within the first 72 h.

INFECTIONS

BACTERIAL INFECTIONS

Other than bacterial vaginosis, the most common bacterial infections during pregnancy involve the urinary tract (Chap. 280). Many pregnant women have asymptomatic bacteriuria, most likely due to stasis caused by progestational effects on ureteral and

bladder smooth muscle and to compression effects of the enlarging uterus. In itself, this condition is not associated with an adverse outcome of pregnancy. However, if asymptomatic bacteriuria is left untreated, symptomatic pyelonephritis may occur. Indeed, approximately 75% of cases of pregnancy-associated pyelonephritis are the result of untreated asymptomatic bacteriuria. All pregnant women should be screened with a urine culture for asymptomatic bacteriuria at the first prenatal visit. Subsequent screening with nitrite/leukocyte esterase strips is indicated for high-risk women, such as those with sickle cell trait or a history of urinary tract infections. All women with positive screens should be treated.

Because of the association between bacterial vaginosis and preterm delivery, screening for bacterial vaginosis has been used in an effort to reduce risk. However, standard treatment for bacterial vaginosis does not reduce the risk of preterm delivery.

Abdominal pain and fever during pregnancy create a clinical dilemma. The diagnosis of greatest concern is intrauterine amniotic infection. While amniotic infection most commonly follows rupture of the membranes, this is not always the case. In general, antibiotic therapy is not recommended as a temporizing measure in these circumstances. If intrauterine infection is suspected, induced delivery with concomitant antibiotic therapy is generally indicated. Intrauterine amniotic infection is most often caused by pathogens such as *Escherichia coli* and group B streptococcus. In high-risk patients at term or in preterm patients, routine intrapartum prophylaxis of group B streptococcal disease is recommended. Penicillin G and ampicillin are the drugs of choice. In penicillin-allergic patients, clindamycin is recommended.

Postpartum infection is a significant cause of maternal morbidity and mortality. While rare after vaginal delivery, postpartum endomyometritis develops in 5% of patients having elective repeat cesarean section and in 25% of patients after emergency cesarean section following prolonged labor. Prophylactic antibiotics should be given to all patients undergoing cesarean section. As most cases of postpartum endomyometritis are polymicrobial, broad-spectrum antibiotic coverage with a penicillin, aminoglycoside, and metronidazole is recommended (Chap. 167). Most cases resolve within 72 h. Women who do not respond to antibiotic treatment for postpartum endomyometritis should be evaluated for septic pelvic thrombophlebitis. Imaging studies may be helpful in establishing the diagnosis, which is primarily a clinical diagnosis of exclusion. Patients with septic pelvic thrombophlebitis generally have tachycardia out of proportion to their fever and respond rapidly to intravenous administration of heparin.

All patients are screened prenatally for gonorrhea and chlamydial infections, and the detection of either should result in prompt treatment. Ceftriaxone and azithromycin are the agents of choice (Chaps. 147 and179).

VIRAL INFECTIONS

Cytomegalovirus Infection Viral infection in pregnancy presents a significant challenge. The most common cause of congenital viral infection in the United States is cytomegalovirus (CMV) (Chap. 185). As many as 50 to 90% of women of childbearing age have antibodies to CMV, but only rarely does CMV reactivation result in neonatal infection. More commonly, primary CMV infection during pregnancy creates a risk of

congenital CMV. No currently accepted treatment of CMV during pregnancy has been demonstrated to protect the fetus effectively. Moreover, it is impossible to predict which fetus will sustain life-threatening CMV infection. Severe CMV disease in the newborn is characterized most often by petechiae, hepatosplenomegaly, and jaundice. Chorioretinitis, microcephaly, intracranial calcifications, hepatitis, hemolytic anemia, and purpura may also develop. Central nervous system involvement resulting in the development of psychomotor, ocular, auditory, and dental abnormalities over time have been described.

Rubella (See alsoChap. 195) Rubella virus is a known teratogen; first-trimester rubella carries a high risk of fetal anomalies, though the risk decreases significantly later in pregnancy. Congenital rubella may be diagnosed by percutaneous umbilical blood sampling with the detection of IgM antibodies in fetal blood. All pregnant women should be screened for their immune status to rubella. Indeed, all women of childbearing age, regardless of pregnancy status, should have their immune status for rubella verified and be immunized if necessary. The incidence of congenital rubella in the United States is extremely low.

Herpesvirus (See alsoChap. 182) The acquisition of genital herpes during pregnancy is associated with spontaneous abortion, prematurity, and congenital and neonatal herpes. A recent cohort study of pregnant women without evidence of previous herpes infection demonstrated that approximately 2% of the women acquired a new herpes infection during the pregnancy. Approximately 60% of the newly infected women had no clinical symptoms. Infection occurred equally in all three trimesters. If herpes seroconversion occurred early in pregnancy, the risk of transmission to the newborn was very low. In women who acquired genital herpes shortly before delivery, the risk of transmission was high. The risk of active genital herpes lesions at term can be reduced by prescribing acyclovir for the last 4 weeks of pregnancy to women who have had their first episode of genital herpes during the pregnancy. However, whether or not this strategy results in less viral shedding or enhanced fetal protection at delivery remains to be determined.

Herpesvirus infection in the newborn can be devastating. Disseminated neonatal herpes carries with it high mortality and morbidity rates from central nervous system involvement. It is recommended that pregnant women with active genital herpes lesions at the time of presentation in labor be delivered by cesarean section.

Parvovirus (See alsoChap. 187) Parvovirus infection (human parvovirus B19) may occur during pregnancy. It rarely causes sequelae, but susceptible women infected during pregnancy may be at risk for fetal hydrops secondary to erythroid aplasia and profound anemia.

Toxoplasmosis (See alsoChap. 217) In the United States, approximately 70% of women of childbearing age are susceptible to *Toxoplasma*. Most primary infections of toxoplasmosis in the United States come from eating undercooked meat. The diagnosis of congenital toxoplasmosis is possible through sampling of fetal umbilical blood. If there is no evidence of placental/fetal infection, single-drug treatment with spiramycin is recommended. Triple-drug therapy with spiramycin, pyrimethamine, and sulfa is recommended if there is evidence of fetal infection and the woman does not wish to

terminate the pregnancy or cannot terminate it because of advanced gestational age. Prenatal treatment has been shown to reduce the number of infants with severe infection.

Human Immunodeficiency Virus (See also Chap. 309) The predominant cause of HIV infection in children is transmission of the virus from the mother to the newborn during the perinatal period. Exposures, which increase the risk of mother-to-child transmission, include vaginal delivery, preterm delivery, trauma to the fetal skin, and maternal bleeding. Additionally, recent infection with high maternal viral load, low maternal CD4+T cell count, prolonged labor, prolonged length of membrane rupture, and the presence of other genital tract infections, such as syphilis or herpes, increase the risk of transmission. Breast feeding may also transmit HIV to the newborn and is therefore contraindicated in most developed countries for HIV-infected mothers. There is no clear evidence to suggest that the course of HIV disease is altered by pregnancy. There is also no clear evidence to suggest that uncomplicated HIV disease adversely impacts pregnancy other than by its inherent infection risk.

TREATMENT

The majority of cases of mother-to-child (vertical) transmission of HIV-1 occur during the intrapartum period. Mechanisms of vertical transmission include infection after rupture of the membranes and direct contact of the fetus with infected secretions or blood from the maternal genital tract. In women with HIV infection who are not receiving antiretroviral therapy, the rate of vertical transmission is approximately 25%. Cesarean section and treatment with zidovudine, administered both before and during delivery, decrease the rate of vertical transmission. In a meta-analysis, zidovudine treatment of both the mother during the prenatal and intrapartum periods and of the neonate at birth reduced the risk of vertical transmission to 7.3%. The combination of elective cesarean section plus zidovudine treatment reduced the risk of vertical transmission to 2%. The role of multiple drug therapy during pregnancy has not yet been established, pending safety data for the neonate.

SUMMARY

Maternal mortality has decreased steadily during the past 60 years. The maternal death rate has decreased from nearly 600/100,000 live births in 1935 to 8.5/100,000 live births in 1996. The most common causes of maternal death in the United States today are, in decreasing order of frequency, thromboembolic disease, hypertension, ectopic pregnancy, and hemorrhage. With improved diagnostic and therapeutic modalities as well as with advances in the treatment of infertility, more patients with medical complications will be seeking, and be in need of, complex obstetric care. Improving outcome of pregnancy in these women will be best obtained by assembling a team of internists and specialists in maternal-fetal medicine (high-risk obstetrics) to counsel these patients about the risks of pregnancy and to plan their treatment prior to conception. The importance of preconception counseling cannot be overstated. It is the responsibility of all physicians caring for women in the reproductive age group to assess their patient's reproductive plans as part of their overall health evaluation.

(Bibliography omitted in Palm version)

Back to Table of Contents

8. ADOLESCENT HEALTH PROBLEMS - *Mehul T. Dattani*, *Charles G. D. Brook*

Adolescence marks the transition from childhood to adulthood. It is a time of dramatic physical and psychological change. Adolescents are particularly prone to risk-taking behaviors. In the United States, 73% of deaths among adolescents and young adults result from motor vehicle and other accidents, homicide, and suicide. As a result of sexual maturation, adolescents begin to experiment sexually and, consequently, are susceptible to sexually transmitted diseases (STDs) and unwanted pregnancies. For adolescents with underlying disease, the psychological consequences can be as important as the physical disabilities; denial or resentment of disease is common and can hamper treatment. Adolescence is also a time when many lifelong health-relevant behaviors are established, including dietary habits, exercise patterns, tobacco and alcohol use, and interactions with the health care system. The physician, working together with parents, can help to guide adolescents through this dynamic period of life.

PUBERTY

Puberty encompasses (1) the adolescent growth spurt, (2) development of secondary sexual characteristics, (3) attainment of fertility, and (4) establishment of individual sexual identity.

There is wide variation in the timing of puberty. Signs of puberty are first evident between 9 and 14 years of age (mean 11.5) in 95% of American boys (Fig. 8-1*A*). In girls, puberty begins earlier, with 95% of American girls entering puberty between 8 and 12 years of age (mean 10.5) (Fig. 8-1*B*). The age of menarche in girls from developed countries has decreased by approximately 2 to 3 months per decade over the past 100 to 150 years. This trend is likely the result of improvements in socioeconomic conditions, nutritional status, and general health and well-being. Currently in the United States, the average age of menarche is 12.8 years. Genetic factors also influence the course of puberty. Data from twin studies indicate that the average age of menarche is more similar in identical twin sisters than in nonidentical twins. Secondary sexual development occurs earlier in girls of Asian and African-Caribbean heritage than in girls of European heritage. Recognition of the progressively earlier onset of puberty, the ethnic variations, and wide age distribution in the timing of puberty is important for identifying precocious or delayed puberty and for counseling adolescents and parents about the natural course of physiologic changes.

HORMONAL CHANGES

Puberty is accompanied by dramatic changes in multiple hormonal systems including alterations in adrenal steroid production, maturation of the reproductive axis, and increased production and action of growth hormone (GH). Serum GH levels increase early in puberty as a consequence of the rise in gonadal steroids. GH in turn increases the level of insulin-like growth factor 1 (IGF-1), which enhances linear bone growth (Chap. 328). The prolonged pubertal exposure to gonadal steroids ultimately causes epiphyseal closure and limits further bone growth. Appetite increases in association with the growth spurt, and sleep patterns change with a tendency to stay up later and a desire to sleep later into the morning.

The development of secondary sexual characteristics is initiated by *adrenarche*, which usually occurs between 6 and 8 years of age and marks the time when the adrenal gland begins to produce greater amounts of androgens. Despite the search for hormonal mediators, the mechanism that controls adrenarche is unknown. It may well result from adrenal cell differentiation, characterized by the growth of the innermost zone of the adrenal cortex, the zona reticularis, which is the principal site of dehydroepiandrosterone (DHEA) production. The increase in adrenal androgen (DHEA, androstenedione) secretion precedes activation of the reproductive axis. Nonetheless, adrenarche and gonadarche are independent events: children with Addison's disease enter puberty at the normal time, and children with premature adrenarche achieve gonadarche at the expected age.

The sexual maturation process is greatly accelerated by the activation of the hypothalamic-pituitary axis, leading to gonadal stimulation and the production of sex steroids. The hypothalamic-pituitary-gonadal axis is controlled predominantly by gonadotropin-releasing hormone (GnRH), a decapeptide produced by the arcuate nucleus of the mediobasal hypothalamus. GnRH is released in a pulsatile fashion, leading in turn to the pulsatile secretion of the pituitary gonadotropins, luteinizing hormone (LH) and follicle-stimulating hormone (FSH). The so-called GnRH pulse-generator in the hypothalamus is active during fetal life and early infancy but is then quiescent during early childhood. In the early stages of puberty, the sensitivity to steroid inhibition is gradually lost, causing reactivation of GnRH secretion. Leptin, a hormone produced by adipose cells, may play a permissive role in this process, as leptin-deficient individuals fail to enter puberty. Early puberty is characterized by nocturnal surges of LH and FSH. As the reproductive axis matures, the characteristic patterns of feedback regulation are acquired. In males, testosterone inhibits hypothalamic GnRH and pituitary gonadotropin production (Chap. 335); in females, estrogen and progesterone feed back to generate the characteristic hormonal patterns of the menstrual cycle (Chap. 336).

FEMALE PUBERTY

The first sign of ovarian estradiol secretion is breast development, or *thelarche*. Pubic and axillary hair growth and the onset of apocrine sweat production result from adrenal androgen secretion, though they may be facilitated by estrogen. The progression of puberty is classified according to Tanner stages for breast and pubic hair development: stage 1 represents the preadolescent appearance, and stage 5 represents the adult appearance (Fig. 8-1*A*). Each of these aspects of puberty should be staged separately, as they are controlled by different underlying endocrine mechanisms. Concurrent with these outward signs of puberty are the changes in the size and shape of the uterus. The pubertal growth spurt is dependent on estradiol secretion, which leads to increasedGHsecretion. This in turn results in a doubling of the growth rate, and peak height velocity is usually coincident with breast stage 3. After menarche, a girl usually grows only an additional 5 cm.

MALE PUBERTY

In boys, growth of the testes is usually the first sign of puberty, reflecting the effects of pulsatile gonadotropin secretion on seminiferous tubule volume and, to some degree,

Leydig cell mass. Testosterone is converted to dihydrotestosterone by 5-a reductase. Both hormones act via the androgen receptor to induce growth of the external genitalia and pubic hair (Fig. 8-1;Chap. 338). The growth spurt in boys occurs at a testicular volume of about 10 to 12 mL (as measured by a Prader orchidometer). Testosterone also deepens the voice and increases muscle growth. Dihydrotestosterone stimulates prostate growth and beard growth and initiates recession of the temporal hairline. Although boys enter puberty approximately 6 to 12 months later than girls, they are potentially fertile at an earlier stage of puberty. Aromatization of testosterone to estradiol increasesGHsecretion, which acts synergistically with testosterone to induce a greater peak height velocity in boys than in girls.

DISORDERS OF PUBERTY

Precocious puberty is usually defined as an early onset of puberty in boys younger than 9 years or in girls younger than 8 years of age. Some authorities suggest that the lower limits of normal in girls be revised downward to age 7 for Caucasians and age 6 for African-American girls. *Premature thelarche* refers to breast development in the absence of other signs of puberty. It occurs most commonly in girls between infancy and 3 years of age and usually resolves spontaneously. Causes of precocious puberty are divided into central *gonadotropin-dependent* forms and peripheral *gonadotropin-independent* forms (Table 8-1). Central precocious puberty is much more common in girls than in boys, and the majority of these cases involve idiopathic activation of spontaneousGnRHpulses. It can also be caused by a variety of central nervous system tumors, structural lesions, and inflammatory conditions.

Delayed puberty is defined as the 3% of girls and boys who have not developed the first signs of puberty by 13.2 and 14.2 years, respectively. It is most commonly due to delayed activation of the hypothalamic-pituitary-gonadal axis (Table 8-1). Most individuals who meet this definition will progress through puberty normally, but at a later age. Short stature and delayed skeletal maturation are commonly seen in association with delayed puberty. Growth delay may have been evident earlier in childhood; the diagnosis of *constitutional delay of growth and puberty* can be suspected from a delayed bone age in a short child who is otherwise well. Individuals who experience delays in puberty may be emotionally as well as physically immature relative to their peers.

The main diagnostic challenge in delayed puberty is to distinguish those with constitutional delay, who will progress through puberty at a later age, from those with an underlying pathologic process.LH andFSHresponses toGnRH do not differentiate constitutional delay from pathologic causes of *hypogonadotropic hypogonadism* (Chap. 335). Thus, constitutional delay is a diagnosis of exclusion and requires ongoing evaluation during development to assure that normal growth and development occur at a later time. Reassurance without hormonal treatment is appropriate for most individuals with presumed constitutional delay of puberty. Alternatively, an anabolic steroid (e.g., 50 to 100 mg per month testosterone enanthate, intramuscularly) in boys or estrogen (5 to 10 mg/d ethinyl estradiol, orally) in girls may be useful to induce growth and secondary sexual characteristics appropriate for age. Low-dose oral oxandrolone (2.5 mg/d), an anabolic steroid that is not aromatized to estrogen, is also used for boys because it does not accelerate skeletal maturation when used for short periods. After treatment for a year or more, hormonal treatments can be stopped and the function of the

reproductive axis can be reassessed.

PSYCHOLOGICAL CHANGES AND SOCIAL FACTORS

The adolescent years are characterized by a multitude of psychological changes, including (1) the development of abstract thinking, (2) greater independence from family, (3) the formation of a personal and sexual identity, (4) the establishment of a system of values, and (5) an increase in socialization. For most adolescents, these transitions occur relatively smoothly. For others, however, these years can be frustrating and tumultuous; parents and clinicians must be attuned to the needs of those who show signs of struggling with emotional, sexual, and social issues.

Young people tend to share their feelings openly, one of which is ambivalence. These contradictory feelings most often involve both a desire for greater autonomy and, at the same time, a need to cling to the emotional and physical security provided by the family. Adolescents are granted increasing responsibilities but still lack some of the social and legal privileges of adults. This feature of adolescence can lead to conflict and challenges to parental authority.

Adolescents have a strong desire to establish an identity that is increasingly independent of the family. This new identity is strongly influenced by peer groups, some of which are institutionalized (e.g., team sports). Role confusion is quite common in adolescence, and some young people move from one intense allegiance to another with alarming speed. These transitional arrangements are eventually replaced by more permanent attachments to individuals.

In an attempt to alleviate some of the transitions associated with adolescence, many cultures have traditionally used "rites of passage" to acknowledge and accelerate an adolescent's evolution to adulthood. Among Native American Great Plains cultures, for example, a boy was sent away from the village at the time of puberty to fast and receive a vision from a spirit; upon returning to the community, he took his place among the adult men. Similarly, it was traditional in many societies for girls to be secluded at the time of the first menstruation before returning a "full-grown woman." These rites provide a public recognition of the end of childhood, and the ritual leaves the young person with the conviction that he or she has undergone a personal transformation. A relative lack of these coming-of-age rituals in western cultures may contribute to the sense of alienation experienced by some adolescents in this part of the world.

During adolescence, gender identity must be renegotiated. Though prepubertal children have a relatively secure view of themselves as either a boy or a girl, experimentation with gender roles is a common feature of adolescence. For example, adolescents may explore, at least in fantasy, alternative gender roles (e.g., cross-dressing), homosexuality, or relationships with older men or women.

The hormonal changes of puberty influence behavior as well as causing physical changes. Rising levels of testosterone in boys and the increase in adrenal and ovarian androgens in girls increase libido. The mean age of sexual intercourse varies widely within and among cultures, but ranges between ages 15 and 18 for most groups. Boys generally report sexual intercourse about 1 year earlier than girls.

ADOLESCENT VIOLENCE

Adolescents and young adults are subject to much greater rates of violence, both as victims and perpetrators. Males are involved in violence much more commonly than females and account for>90% of homicides involving those 10 to 17 years of age. Ethnic and racial differences in rates of adolescent violence have been noted consistently. African Americans, Hispanics, and Native Americans are much more likely to be victims and perpetrators of lethal violence than are people of Asian or European ancestry. The origins of different rates of violence are complex. Higher rates of lethal aggression are associated with low socioeconomic status, high housing density, increased population turnover in neighborhoods, single-parent households, and socially disorganized communities. In many cases, these factors interact; increased violence leads to high population turnover and social disorganization.

Gangs represent a potentially volatile environment that is characterized by power struggles, initiation and detachment rituals, battles over territory, and escalating violence associated with retaliation. The increase in lethal violence has been attributed in part to easier access to firearms and a greater willingness to use firearms. A Centers for Disease Control and Prevention study in 1995 found that about one-fourth of students had carried a weapon to school during the preceding month and 8 to 10% had carried a gun. Many adolescents lack the abstract reasoning skills required to understand social mores and the consequence of gun use. Though firearms do not cause violence, handguns in particular provide a facile means to a lethal outcome; widespread reduction in access to handguns is essential to curb the current trend in adolescent homicide and serious injury.

Aggressive behavior can often be recognized in early childhood; bullying is a precursor to later antisocial behavior. Child abuse, antisocial parents, inadequate child-rearing practices, and dysfunctional interpersonal interactions between parents or among siblings are associated with aggressive behavior. The physician, along with teachers, clergy, and others in positions of authority, should be alert to a pattern of aggressive behavior or problems in the home. Though these issues are not easily remedied, appropriate interventions to improve family functioning and parenting may interrupt a pattern of violence, which is all too often perpetuated by the adolescent.

HEALTH PROBLEMS

Adolescence is generally a healthy period and is often accompanied by a feeling of immortality, which leads to risk-taking. When diseases of childhood or the consequences of their treatment extend into adolescence, or when disease strikes during adolescence, the sense of unfairness may be overwhelming. Anger and denial can lead to poor compliance with therapeutic regimens.

Relatively few diseases are unique to adolescents. Rather, diseases of childhood, including many inherited disorders and infectious diseases, extend into the adolescent period. Similarly, many of the disorders that affect teenagers are also seen in the adult population. The presentation and management of asthma, for example, is similar in adolescents and adults. Some of the diseases with relatively increased prevalence

during adolescence are summarized inTable 8-2. These diseases should be borne in mind when considering the differential diagnosis. For example, when an adolescent presents with exertional chest pain, dyspnea, and syncope, hypertrophic cardiomyopathy or congenital heart disease should be considered as likely diagnoses, whereas coronary artery disease would be more likely in an adult.

SEXUALLY TRANSMITTED DISEASES

Sexually active adolescents are at greater risk of acquiringSTDsthan their adult counterparts (Chap. 132). Prevention of STDs in adolescence depends on adequate sexual education coupled with access to appropriate clinical services. Early age of first sexual intercourse is associated with (1) an increased number of lifetime sexual partners; (2) an increased risk of acquiring chronic STDs, such as herpes simplex, HIV, and hepatitis B; and (3) cervical cancer in women. In addition, pelvic inflammatory disease in adolescent females increases the likelihood of future ectopic pregnancy, tubal infertility, and chronic pelvic inflammation. A low rate of barrier contraceptive use, combined with ignorance about the acquisition and prevention of infectious diseases, also contributes to the increased risk of STDs among adolescents. Screening for STDs is recommended in sexually active teens (Table 8-3). Adolescents with sexually transmitted infections, particularly those who deny sexual activity, may be victims of sexual abuse.

CHILD SEXUAL ABUSE

Child sexual abuse is defined as the involvement of developmentally immature children and adolescents in sexual activities they do not comprehend, to which they are unable to give consent, or that violate social taboos or family roles. In a U.S. study in 1985, sexual abuse during childhood was reported by 27% of adult females and 16% of adult males. Females are more likely than males to have been sexually abused by a family member. Although there is a paucity of literature on male sexual abuse, it is probably more common than generally recognized. The psychological trauma appears to be similar for boys and girls. Sexual abuse during adolescence may merge with peer sexual assault, or "date rape." Sexual abuse in adolescent girls can be associated with a constant fear of pregnancy. Teenage pregnancy orSTD may, in fact, be the first indication of ongoing abuse.

Psychological consequences of child sexual abuse often involve behavioral problems, psychiatric disturbances, or adjustment difficulties at the onset of adolescence, even though the actual abuse may have taken place at a younger age. Child sexual abuse may lead to low self-esteem and/or a degree of sexual disinhibition. The cognitive maturation that occurs with adolescence may bring about the realization and expression of these feelings. Young women who have been sexually abused have significantly higher rates of early-onset consensual sexual activity, teenage pregnancy, multiple sexual partners, unprotected intercourse,STDs, and later sexual assault. Poor psychological outcome is related to the duration of abuse, the extent to which the abuse involves violence or coercion, and the perception that the child has cooperated with the abuser, with ensuing feelings of guilt. The impact of these sequelae can be reduced by supportive peer and family relationships. Disclosure of the abuse may help to ameliorate some of the psychological traumas associated with abuse.

SUBSTANCE ABUSE

Substance abuse and drug misuse among adolescents is a significant cause of morbidity and mortality (Chaps. 386 to 389). The prevalence rates vary widely by region, ethnic group, age, and gender. The age of initiation into substance abuse has gradually declined. In 1997, rates among American teenagers for substance use or abuse, at some stage during their lifetimes, were: cigarettes smoking (70%), alcohol use (79%), marijuana use (47%), cocaine use (8%), anabolic steroids (4%), injected illegal drugs (2%), and other illegal drugs (17%), e.g., lysergic acid (LSD), phencyclidine (PCP), methylenedioxymethamphetamine (ecstasy), methamphetamine (ice), or heroin.

The forms of substance abuse change continuously. Anabolic steroids, for example, are now used by 3 to 5% of male high school seniors, with a 10% prevalence rate among male adolescent athletes. In addition to their use by athletes in an effort to increase muscle strength, nonathletes use anabolic steroids with a goal of achieving a more virile appearance. In contrast to popular views, anabolic steroids do not appear to enhance performance except at very high doses, which are associated with significant side effects (Chap. 335). Other performance-enhancing agents include human growth hormone and erythropoietin (EPO), but the high cost of these hormones limits their use.

In addition to the direct effect on health, substance abuse is associated with other risk-taking behaviors. The relationship of alcohol use and motor vehicle accidents, for example, is well documented. However, drug and alcohol use are also correlated with many other problems during adolescence including violence, suicide, depression,STD, and unwanted pregnancies. Therefore, the presence of one form of risky behavior should prompt consideration of others.

SUICIDAL BEHAVIOR AND DEPRESSION

After motor vehicle accidents and homicide, suicide is the third leading cause of death in adolescents, and the rate has risen almost fourfold over the past 50 years. In 1988, the suicide rate among 15- to 19-year olds was 11.3 in 100,000. The causes for increased rates of suicide are not well understood, but one theory holds that modern society fosters increased social isolation and alienation. Nearly one-fourth of adolescents acknowledge seriously considering suicide, and 8% have actually attempted it. Attempted suicide is three times more common in females than males, with drug overdose or wrist-cutting being the most common means of suicide attempt. Completed suicide is three to five times more common in teenage boys than girls and usually involves firearms, hanging, or jumping from heights. Suicide is rare before puberty. Risk factors for suicide among adolescents include prior attempt of suicide, a history of depression or other major psychiatric disorder, history of substance abuse, medical illness, family history of suicidal behavior, and knowing someone who has committed suicide. Unfortunately, these and other risk factors are relatively common among nonsuicidal youth as well, making suicide difficult to predict in individual cases. Stressful events can precipitate depression and increase risk of suicide; these can include the death of a relative or friend, disciplinary crisis, rejection or humiliation, school difficulty, and anxiety about homosexuality. Apparently impulsive actions may be harbingers of more serious underlying mood disturbances, personality disorders, or substance abuse.

Major depression occurs in 4 to 6% of adolescents, and the *DSM-IV* criteria for diagnosis are the same as in adults (Chap. 385). Every depressed or suicidal adolescent should undergo psychiatric examination, whether hospitalized or not. Comprehensive evaluation requires exploration of the adolescent's history of mental health problems, symptoms of depression, level of functioning in school, interactions with friends and family, and evaluation for comorbid disorders. Indications for hospitalization include imminent risk of suicide as evidenced by an identified plan and access to lethal means, recurrent suicide attempts, the presence of severe depression or psychosis, substance abuse, and the need to remove the individual from an overwhelmingly stressful environment.

ADOLESCENT EATING DISORDERS

Many adolescents have voracious appetites in response to the increased energy and caloric requirements generated by the growth spurt. The unique physical, psychological, and social transitions of adolescence provide a context for the development and perpetuation of eating patterns. Adolescents with a body mass index (BMI), measured as weight (kg)/height (m_2), greater than the 95th percentile for age and gender are overweight, and those between the 85th and 94th percentiles are at risk for becoming overweight. Based on the NHANES III survey for 1988 to 1994, there was evidence for a 6% increase in the prevalence of overweight adolescents compared to the previous decade. The increasing prevalence of obesity is multifactorial and involves patterns of eating behavior as well as alterations in activity level (Chap. 77). Physical activity among both girls and boys tends to decline steadily during adolescence. Regular involvement in enjoyable forms of exercise should be encouraged to help promote lifelong habits that involve physical activity.

Eating disorders such as anorexia nervosa or bulimia nervosa often have their onset during adolescence (Chap. 78). Control over dietary intake is perhaps one of the first mechanisms that adolescents use to establish autonomy and achieve independence from family. The majority of female adolescents and young adults in western cultures report feeling discontented with their body shape. Surveys of normal adolescent populations disclose a surprisingly high frequency of dieting and abnormal eating patterns. For instance, up to 79% binge, 70% consider themselves fat, 11% induce vomiting, 5% abuse laxatives, and about 3% meet diagnostic criteria for anorexia or bulimia nervosa. Eating disorders also occur in males, but much less frequently than in females.

PHYSICIAN-ADOLESCENT RELATIONSHIP

The transition from the pediatrician to an adult medical practice can be difficult for adolescents, their parents, and their physicians. The emergence of adolescent medicine as a specialty practice has helped to facilitate this transition and to focus on the special needs of this group. When adolescents transfer to an adult-based practice, the physician should first establish a relationship with the patient and his or her parents. Previous medical history should be reviewed and medical records obtained. The need for the teenager to be seen alone, and office policies concerning confidentiality, should be discussed and agreed to with the parent(s) and adolescent together.

Legal issues related to the medical care of minors arise frequently, and laws vary in different countries and from state to state (Chap. 2). As a general rule, anyone who has reached the age of majority (usually 18 years) may consent to treatment. Under this age, a parent or legal guardian must consent for medical intervention. However, there are several exceptions to this requirement. The delivery of medical care is generally accepted in an emergency, but it is important to document the nature of the emergency and any efforts to notify parents. Emancipated minors may also provide consent. This group includes those fulfilling adult roles (e.g., military service), married teens, and those who are financially independent and living separately from their parents. In addition, when the health of a minor is potentially endangered by disorders for which they may be reluctant to seek parental consent, such as substance abuse, pregnancy, orSTD, mature minors may generally provide consent. In these circumstances, the caregiver must assess the minor's maturity, ability to understand the risks and benefits of treatment, and capacity to provide informed consent. Mature or emancipated minors do not need to reveal consent or treatment to their parents.

Obtaining a medical history from an adolescent includes many elements that are distinct from an adult history. It should include, for example, schoolwork, home environment, and relationships with parents, siblings, and peers. Adolescents often lack knowledge about medical issues and may be reluctant to discuss sensitive topics with authority figures. Most will be nervous, even when these issues do not pertain. It can be useful, therefore, to provide printed forms or questionnaires. These not only serve to gather information in a relatively nonthreatening manner but also provide an indication of the kinds of issues that might be discussed with the physician. It is difficult to predict the topics that are paramount to the adolescent. Some may be preoccupied with concerns about the onset of acne, whereas others fear HIV or pregnancy. Adolescents may harbor guilt about sexual abuse or feel overwhelmed by peer pressure to engage in certain activities. The physician is well positioned to assist with many of these issues, if there is trust and an indication of interest and understanding. Because of these types of questions, it is important to interview the adolescent in private. Some parents will resist this approach, but it is necessary if the adolescent is to volunteer information that he or she is unwilling to discuss in the presence of parents. It is useful to reinforce the fact that conversations will be kept confidential. Adolescents are sometimes willing to bring concerns to the attention of nurses or other caregivers before raising these issues with a physician. It is helpful, therefore, to have another health care provider interact with the patient, if only briefly. In addition to direct questioning, general conversation about topical issues or inquiries about school or peers may provide insight into an adolescent's interests, activities, and potential risk factors. Because of the prevalence of substance abuse, risk-taking behavior, suicide, sexual orientation crises,STDs, unwanted pregnancies, sexual abuse, depression, and eating disorders in the teenage years, these topics warrant specific inquiry as part of routine health assessment. The interview should also include adequate time for education, health care guidance, and counseling. It is also useful to provide written information about topics that are pertinent to the care of the adolescent.

The physical examination of the adolescent, while incorporating many elements of the adult examination, has several unique features. Foremost among these is the assessment of growth and sexual development. In addition to questionnaires that allow

the adolescent an opportunity to self-assess stages of pubertal development, the examination can be made less stressful by using it as opportunity to explain normal physiology. The issue of when to perform a pelvic examination as part of the routine health maintenance is controversial. Some advocate pelvic examinations in all sexually active young women as a means to detectSTDs and for Pap smears. With the advent of urinary screening tests for chlamydia and gonorrhea, others suggest that pelvic examinations are not routinely necessary in the absence of specific indications. When a pelvic examination is performed, the patient should be asked whether she prefers her mother or a member of the health care team as an observer. The physical examination should also focus on diseases that tend to present during adolescence (Table 8-2).

Disorders such as hypertension, hyperlipidemia, and obesity are often first detected during adolescence. Strategies for disease prevention also include immunization, avoidance of cigarette smoking or excessive alcohol use, establishing good dietary habits, and engaging in regular exercise. General guidelines for adolescent preventive services (GAPS) are summarized inTable 8-3.

SUMMARY

The term *adolescent* is derived from a Latin phrase meaning "to grow up." Adolescence is, in many ways, the culmination of development, with the achievement of identity and reproductive competence. Though these processes are triggered by internal physiologic events, they are intimately intertwined with the family and social environment. Physicians have an important role to facilitate these transitions by providing information and managing the diseases of adolescents. Moreover, it should be remembered that many adolescents view physicians as role models and will seek objective and informed advice about issues that reach beyond medicine.

(Bibliography omitted in Palm version)

Back to Table of Contents

9. GERIATRIC MEDICINE - *Neil M. Resnick*

Of all the people who have ever lived to age 65, more than half are now alive. This statistic has important demographic and economic implications, and its impact on medical care is also substantial.

BIOLOGY OF AGING

Numerous molecular concomitants of aging have been described. For instance, there is an increase in chromosome structural abnormalities, DNA cross-linking, and frequency of single-strand breaks; a decline in DNA methylation; and loss of DNA telomeric sequences. The primary structure of proteins is unaltered, but posttranslational changes, such as deamidation, oxidation, cross-linking, and nonenzymatic glycation, increase. Mitochondrial structure also deteriorates, albeit not universally.

However, the biologic changes are clearer than the mechanisms that mediate them. In fact, although the senescent phenotype appears to be ubiquitous, biologists disagree about whether senescence even exists beyond zoos and civilized societies and whether it occurs at all in many species. There is little evolutionary rationale for a process that happens after reproduction is complete, particularly one associated with such a long and complex course. In nature, senescence is most notable for its absence; nearly all animals die of predation, disease, or environmental hazards rather than aging. The argument that different species have different maximum life spans can be explained without invoking a specific aging process: while growth and development are based on a genetic template, aging may reflect merely the accumulation of random damage rather than a specific mechanism.

If aging exists as a distinct process, there is consensus that the mechanisms are likely multifactorial, environmentally influenced, and species-specific, if not organ- and cell-specific, making the paucity of available human data particularly problematic. As a result, there are nearly as many theories of aging as investigators. Most theories overlap or are not mutually exclusive, and none is completely compatible with the dearth of data. As a group, the theories can be divided into two broad categories, based on whether they attribute aging to a genetic program or to progressive and random damage to homeostatic systems.

Enthusiasm for genetic theories of aging is fueled by several observations, including the dramatic species-specific differences in maximal life span, the strong correlation with survival among monozygotic compared to dizygotic twins, and the fact that single mutations can prolong life span by more than 50% in some nematodes and mice. However, all genetic theories must account for the fact that evolutionary selection pressure is minimal following completion of reproduction. Three genetic theories have recently been advanced, but few relevant data have yet been accrued. The first theory suggests that, since animals usually succumb to natural forces long before reaching their maximal life span, aging might reflect mutations that impair long-term survival. These mutations would accumulate in the genome because there is no selection pressure to delete them. A second theory, "pleiotropic antagonism," proposes that aging may be caused by the late and deleterious effects of genes that are conserved because of the survival advantages they confer prior to reproduction. The third theory applies to

ecological niches where extrinsic hazards are relatively low. In such an environment, evolution might select for mutations that retard the aging process since these might allow an animal to produce and protect many more litters. In support of this theory, the rate of aging in an isolated clan of Virginia opossums was calculated to be roughly half of that seen in their less fortunate cousins.

The "random damage" theories are based on the possibility that the balance between ongoing damage and repair is disrupted. The theories differ in the emphasis placed on increased damage (e.g., by free radicals, oxidation, or glycation) versus deficient repair, as well as in the mechanisms that might mediate each. However, all share the observation that cell and organ repair capacity declines with age. Some 40 years ago, Hayflick and Moorehead observed that the number of replications among cultured cells is finite. Subsequent research revealed that this replicative senescence was due to arrest of the cell cycle at the G_1/S phase, the point at which DNA synthesis begins. Recently, cell replication has also been linked to the length of telomeric DNA. Present at the termini of chromosomes, telomeric DNA prevents chromosomal instability, fragmentation, and rearrangement; anchors chromosomes to nuclear matrix; and provides a buffer between coding regions of DNA and the ends of the chromosomes. In addition, telomeric DNA is necessary for cell division. With each cell division, however, roughly 50 of the total 2000 base pairs of the telomere are lost. Telomeric shortening might thus result in loss of gene accessibility, which is necessary to repair ongoing cell damage caused by metabolism. Together with cytoplasmic factors mediating arrest of DNA synthesis, telomeric shortening could also limit the cell's ability to divide and thereby replace cells lost to apoptosis.

Many mechanisms previously postulated to mediate aging have not been borne out, including the somatic mutation theory (in which aging would result from cumulative spontaneous mutations), the error catastrophe theory (in which aging would result from errors in the synthesis of proteins critical to the synthesis of genetic material or protein-synthesizing machinery), and the intrinsic mutagenesis theory (in which aging is the result of ongoing intrinsic DNA rearrangements).

To date, the only intervention known to delay aging is caloric restriction. The salutary effect of restricting caloric intake by 30 to 40% has been documented in multiple species, from single-cell organisms to rodents. In rodents, it not only increases average life expectancy and maximum life span but also delays the onset of some typical age-associated diseases as well as deterioration of physiologic systems (e.g., immune responsiveness, glucose metabolism, muscle atrophy). Moreover, its impact is evident in both mitotic and postmitotic cells, in gene expression, and in protein turnover and cross-linking. Although the mechanism is still not determined, it is specific to caloric restriction rather than to reduction of any dietary component (e.g., fat intake) or supplements with vitamins or antioxidants. Unfortunately, adequate data from primates are not yet available, and the effect of caloric restriction in humans is still unknown.

PRINCIPLES OF GERIATRIC MEDICINE

Despite the biologic controversy, from a physiologic standpoint human aging is characterized by progressive constriction of the homeostatic reserve of every organ system. This decline, often referred to as *homeostenosis*, is evident by the third decade

and is gradual and progressive, although the rate and extent of decline vary. The decline of each organ system (Table 9-1) appears to occur independently of changes in other organ systems and is influenced by diet, environment, and personal habits as well as by genetic factors.

Several important principles follow from these facts: (1) Individuals become more dissimilar as they age, belying any stereotype of aging; (2) an *abrupt* decline in any system or function is always due to disease and not to "normal aging"; (3) "normal aging" can be attenuated by modification of risk factors (e.g., increased blood pressure, smoking, sedentary lifestyle); and (4) "healthy old age" is not an oxymoron. In fact, *in the absence of disease, the decline in homeostatic reserve causes no symptoms and imposes few restrictions on activities of daily living regardless of age*.

Appreciation of these facts may make it easier to understand the striking increases that have occurred in life expectancy. Average life expectancy is now 17 years at age 65, 11 years at age 75, 6 years at age 85, 4 years at age 90, and 2 years at age 100. Moreover, the bulk of these years is characterized by a lack of significant impairment (Table 9-2). Even beyond age 85, only 30% of people are impaired in any activity required for daily living and only 20% reside in a nursing home. Yet, as individuals age they are more likely to suffer from disease, disability, and the side effects of drugs, all of which, when combined with the decrease in physiologic reserve, make the older person more vulnerable to environmental, pathologic, and pharmacologic challenges.

The following concepts underlie the remainder of the chapter:

1. Disease presentation is often atypical in the elderly, especially in those more than 75 to 80 years old. Homeostatic strain caused by onset of a new disease often leads to symptoms associated with a different organ system, particularly one compromised by preexisting disease. For example, fewer than one-fourth of older patients with hyperthyroidism present with goiter, tremor, and exophthalmos; more likely are atrial fibrillation, confusion, depression, syncope, and weakness. Significantly, because the "weakest link" is so often the brain, the lower urinary tract, or the cardiovascular or musculoskeletal system, a limited number of presenting symptoms predominate -- acute confusion, depression, incontinence, falling, and syncope -- no matter what the underlying disease. Thus for the most common geriatric syndromes, regardless of the presenting symptom, the differential diagnosis is often largely similar. The corollary is equally important: The organ system usually associated with a particular symptom is less likely to be the source of that symptom in older individuals than in younger ones. Compared with middle-aged individuals, for example, acute confusion in older patients is less often due to a new brain lesion, depression to a psychiatric disorder, incontinence to bladder dysfunction, falling to a neuropathy, or syncope to heart disease.

2. Because of decreased physiologic reserve, older patients often develop symptoms at an earlier stage of their disease (Fig. 9-1). For example, heart failure may be precipitated by mild hyperthyroidism, cognitive dysfunction by mild hyperparathyroidism, urinary retention by mild prostatic enlargement, and nonketotic hyperosmolar coma by mild glucose intolerance. Paradoxically, therefore, treatment of the underlying disease may be easier because it is frequently less advanced at the time of presentation. A

corollary is that drug side effects can occur with drugs and drug doses unlikely to produce side effects in younger people (Chap. 71). For instance, an antihistamine (e.g., diphenhydramine) may cause confusion, loop diuretics may precipitate urinary incontinence, digoxin may induce depression even with normal serum levels, and over-the-counter sympathomimetics may precipitate urinary retention in men with mild prostatic obstruction.

Unfortunately, the predisposition to develop symptoms at an earlier stage of disease is often offset by the change in illness behavior that occurs with age. Raised at a time when symptoms and debility were accepted as normal consequences of aging, the elderly are less likely to seek attention until symptoms become disabling. Thus, any symptom, particularly those associated with a change in functional status, must be taken seriously and evaluated promptly.

3. Since many homeostatic mechanisms may be compromised concurrently, there are usually multiple abnormalities amenable to treatment, and small improvements in each may yield dramatic benefits overall. For instance, cognitive impairment in patients with Alzheimer's disease may respond much better to interventions that alleviate comorbidity than to prescription of donepezil (Fig. 9-2). Similar approaches apply to most other geriatric syndromes, including falls, incontinence, depression, delirium, syncope, and fracture. In each case, substantial functional improvement can result from treating the contributing factors even if -- as in Alzheimer's disease -- the disease itself is largely untreatable.

4. Many findings that are abnormal in younger patients are relatively common in older people -- e.g., bacteriuria, premature ventricular contractions, low bone mineral density, impaired glucose tolerance, and uninhibited bladder contractions. However, they may not be responsible for a particular symptom but only be incidental findings that result in missed diagnoses and misdirected therapy. For instance, the finding of bacteriuria should not end the search for a source of fever in an acutely ill older patient, nor should an elevated random blood sugar -- especially in an acutely ill patient -- be incriminated as the cause of neuropathy. On the other hand, certain other abnormalities must not be dismissed as due to old age -- e.g., there is no anemia, impotence, depression, or confusion of old age.

5. Because symptoms in older people are often due to multiple causes, the diagnostic "law of parsimony" often does not apply. For instance, fever, anemia, retinal embolus, and a heart murmur prompt almost a reflex diagnosis of infective endocarditis in a younger patient but may reflect aspirin-induced blood loss, a cholesterol embolus, insignificant aortic sclerosis, and a viral illness in an older patient. Moreover, even when the diagnosis is correct, treatment of a single disease in an older patient is unlikely to result in cure. For instance, in a younger patient, incontinence due to involuntary bladder contractions is treated effectively with a bladder relaxant medication. However, in an older patient with the same condition but who also has fecal impaction, takes medications that cloud the sensorium, and suffers from arthritis-associated impairments of mobility and manual dexterity, treatment of the bladder spasms alone is unlikely to restore continence. On the other hand, disimpaction, discontinuation of the offending medications, and treatment of the arthritis are likely to restore continence without the need for a bladder relaxant. Failure to recognize these principles often leads to

prescribing "ineffective" therapy and to unjustified therapeutic nihilism towards older patients.

6. Because the older patient is more likely to suffer the adverse consequences of disease, treatment -- and even prevention -- may be equally or even more effective. For instance, the survival benefits of exercise, as well as thrombolysis and beta-blocker therapy after a myocardial infarction, are as impressive in older patients as in younger ones; and treatment of hypertension and transient ischemic attacks, as well as immunization against influenza and pneumococcal pneumonia, are more effective in older patients. In addition, prevention in older patients must often be seen in a broader context. For instance, although interventions to increase bone density may be limited in older patients, fracture may still be prevented by efforts to improve balance, strengthen legs, reduce peripheral edema, treat other contributing medical conditions, replete nutritional deficits, eliminate environmental hazards, and remove adverse medications -- not so much those that affect bone metabolism, but rather those that induce orthostasis, confusion, and extrapyramidal stiffness.

In summary, optimal treatment of the older patient generally requires treating much more than the organ system usually associated with the disease or symptom, and often permits ignoring that system entirely.

EVALUATION

Evaluation of the older patient can be time-consuming, even when it is tailored to the problem. Yet, such initial investment can reduce subsequent morbidity and resource utilization and enhance patient and physician satisfaction. Additionally, the assessment can often be accomplished over several visits. Moreover, much can be gleaned from questionnaires filled out by the patient or caregiver in advance as well as from observation. For instance, greeting the patient in the waiting room allows the physician to note affective and cognitive response, the strength of the handshake, the ease of rising from a chair without using the arms, the length and steadiness of the stride, and the ability to follow directions to the examining room and to sit down safely in the examining room chair. Observing the patient dress or undress can also enhance detection of impaired cognition, fine motor skills, balance, and judgment. Such observations often provide more information than standard examinations and can shorten the clinical evaluation.

HISTORY TAKING IN ELDERLY PATIENTS

Most older patients are able to provide a reliable medical history; however, a multitude of complaints may make obtaining a history more difficult. If the patient is unable to comprehend or communicate, data should be sought from family, friends, and caregivers. The history should also include drug ingestion; dietary patterns; falling, incontinence, sexual dysfunction, depression and anxiety.

Advance Directives All older patients should be asked whether they have drafted advance health care directives, and, if they have, a copy should be placed in the record. Such directives may consist of a health care proxy or durable power of attorney for health care, in which patients designate a surrogate decision-maker who makes health

care decisions if the patient cannot, and/or a living will or medical directive, in which patients specify their desires for treatment in specific situations if they cannot communicate at the critical time.

Whether or not the patient has formally drafted these directives, it is useful to indicate in the record who should make health care decisions if the patient is no longer able to do so. Patients should then be encouraged to discuss their thoughts with the physician as well as the designated proxy. It is not feasible to cover all possible future complications in such discussions. Ascertaining patients' perspectives on specific interventions, such as resuscitation or intubation, is also difficult because preferences will likely differ depending on prognosis. For instance, a patient may not be interested in feeding tube placement following a massive stroke with little chance of recovery but would prefer the same intervention if it is short-term and helps ensure more rapid and complete recovery from an intercurrent illness such as pneumonia. More useful is a discussion that uses open-ended questions and empathic comments to elicit the patient's values and goals. Moreover, for any given condition, preferences may differ depending on baseline clinical status. For robust elderly individuals, recovery is a realistic goal, albeit the odds of complications are higher than for younger individuals. For the frail elderly patient with comorbidity that impairs functional status, reduction or alleviation of symptoms may be the goal. For patients with advanced dementia or terminal illness, palliation may be the most appropriate strategy. In each situation, however, early elicitation of a patient's preferences and values -- when the patient can still state them -- can often help both physicians and families in subsequent difficult decisions by giving surrogate decision-makers the sense that they are doing as the patient would have wanted.

PHYSICAL EXAMINATION

Certain features of the examination should receive special attention, depending in part on clues from the history. Weight and postural blood pressure should be measured at most visits. Vision and hearing should be checked; if hearing is impaired, excess cerumen should be removed from the external auditory canals prior to audiologic referral. Denture fit should be assessed, and the oral cavity should be inspected with the dentures removed. Although thyroid disease becomes more common with age, the sensitivity and specificity of related findings are substantially lower than in younger individuals; consequently, the physical examination can rarely corroborate or exclude thyroid dysfunction in older patients. The breasts should not be overlooked, since older women are more likely to have breast cancer and less likely to do breast self-examination. The systolic murmur of aortic sclerosis is common and may be difficult to differentiate from aortic stenosis, especially since the presence of a fourth heart sound in an elderly person does not imply significant cardiac disease, and the carotid upstroke normally increases owing to age-related arterial stiffening.

In inactive patients and those with fecal or urinary incontinence, one should check for fecal impaction. In patients with urinary incontinence -- especially men -- a distended bladder must be looked for, since it may be the only finding in urinary retention; perineal sensation and the bulbocavernosus reflex should also be tested. Patients who fall should be observed standing up from a chair, bending down, reaching up, walking 10 feet, turning, returning, and sitting again; abnormalities of gait and balance should be evaluated with the patient's eyes open and closed and in response to a sternal push. It

should be appreciated that "frontal release signs" (e.g., "snout," "glabellar," or palmomental reflexes) and absent ankle jerks and vibratory sense in the feet may be normal in the elderly.

MENTAL STATUS EXAMINATION

In addition to evaluating mood and affect, some form of cognitive testing is essential in all elderly patients, even if it involves only checking different components of the history for consistency. People with mild degrees of dementia usually retain their social graces and may mask intellectual impairment by a cheerful and cooperative manner. Thus, the examiner should always probe for content. For patients who follow the news, one can ask what stories they are particularly interested in and why; the same applies to reading, social events -- even the soap operas on television.

If there is any suspicion of a cognitive deficit after this kind of conversational probing, further questioning is indicated. An examination that tests only orientation as to person, place, and time is insufficient to detect mild or moderate intellectual impairment. As a quick screen, simply assessing orientation and asking the patient to draw a clock with the hands at a set time (e.g., 10 min before 2:00) can be very informative regarding cognitive status, visuospatial deficits, ability to comprehend and execute instructions in logical sequence, and presence or absence of perseveration. For slightly more detailed examinations, many practical mental status tests are available. The most widely used is the Mini-Mental Status Examination of Folstein (Chap. 24), which provides a numerical score that can be obtained in 5 to 10 min. Regardless of the test employed, the total score is less useful diagnostically than is knowledge of the specific domain of the deficit. As a general rule, disproportionate difficulty with immediate recall (e.g., of a list of three items) suggests depression, while predominant difficulty with recalling the items 5 min later suggests dementia. For patients with deficits of attention -- recognized by inability to spell simple words backwards, repeat five digits, or recite the months of the year backwards -- delirium is probably present, and the accuracy of the remainder of the test is dubious. However, the test can be interpreted accurately only in the context of a comprehensive evaluation.

EVALUATION OF FUNCTIONAL CAPACITY

Medical problem lists, a standard tool for assessing and following younger patients, often prove inadequate for older patients. Heart failure, stroke, and prostate cancer can describe a bedbound institutionalized person as well as a Supreme Court justice. Thus, it is essential to ascertain the patient's degree of functional incapacity owing to both medical and psychosocial problems. The functional assessment includes determination of the patient's ability to perform basic activities of daily life (ADL), which are those needed for personal self-care, as well as the ability to perform more complex tasks required for independent living, the instrumental activities of daily living (IADL). ADLs include bathing, dressing, toileting, feeding, getting in and out of chairs and bed, and walking. IADLs include shopping, cooking, money management, housework, using a telephone, and traveling outside the home. For frail patients, an assessment in the home by a trained observer may be required, but for most patients a questionnaire dealing with these activities can be completed by the family or patient. In either case, the physician must determine the cause of any impairment and whether it can be

treated. Assessment should conclude with determination of the socioeconomic circumstances and social support systems.

MANAGEMENT OF COMMON GERIATRIC CONDITIONS

Diseases more common in the elderly are covered elsewhere in the text. The medical problems discussed below do not usually present as clear-cut organ-specific diagnoses and are most common in the frail elderly, especially those over 80 years of age.

INTELLECTUAL IMPAIRMENT

The predominant causes of impaired mentation in older patients are delirium, dementia, and depression. Each condition is covered elsewhere in the text in detail (Chaps. 24 and362), but their management in the elderly is discussed here.

Differentiating the causes of impaired mentation is important, but in older patients they frequently coexist. Thus, the most important first step is to search for and correct all factors that may contribute to cognitive impairment, even in patients with dementia (Fig. 9-2). Evidence of dangerous behavior should also be sought (e.g., leaving the stove on, wandering, and getting lost), and plans should be devised to deal with it. Although there is no specific pharmacologic treatment for Alzheimer's disease and agents such as donepezil are of limited efficacy, this does not mean that the physician has no further role in treating the patient and family. In addition to discontinuing all nonessential medications and treating new intercurrent illness, the physician should help the family and patient predict and deal with the disease; indeed, the family often needs the physician's support more than the patient does.

TREATMENT

Community services should be suggested as needed, including a visiting nurse, a home health aide to assist with personal hygiene, a homemaker to assist with housework, meal delivery, transportation services, day health centers, and respite care to ease the burden on family members. Support groups such as the Alzheimer's Association are often of value to the family and help them to anticipate problems. Signs of patient abuse by an overstressed caregiver should be watched for. Legal counsel should be recommended to help the patient and family devise plans for ongoing management and ultimate disposition of assets not already obtained; advance directives should be sought as soon as possible while the patient can still participate.

Finally, abrupt worsening of mentation or the onset of disruptive behavior should always prompt a search for new illness or medication. Exacerbation of cognitive dysfunction may occur with mild infections (e.g., subungual toe abscess, vaginitis, or pressure ulcer); with "therapeutic" levels of many drugs; with use of nonprescribed drugs or alcohol; with modest abnormalities of serum sodium, calcium, glucose, or thyroxine; with mild hypoxia; with borderline nutritional deficiencies; with subdural hematoma or "minor" stroke; and with the development of fecal impaction, urinary retention, pain, or change in environment, particularly in frail older patients. However, if a cause is not found and behavior does not respond to environmental manipulation (e.g., ignoring the behavior, distracting the patient, addressing situational "triggers," and providing a calm

environment), low doses of an antipsychotic medication may be helpful (e.g., haloperidol 0.25 to 2 mg/d orally; see below).

DEPRESSION

Depression of significant degree occurs in 5 to 10% of community-dwelling elderly but is often overlooked. At highest risk are individuals with recent medical illness (e.g., stroke or fracture), bereavement, lack of social supports, recent nursing home admission, or psychiatric history (including alcohol abuse). The diagnosis requires the presence of a depressed mood for at least two consecutive weeks plus at least four of the following eight symptoms: sleep disturbance, lack of interest, feelings of guilt, decreased energy, decreased concentration, decreased appetite, psychomotor agitation/retardation, and suicidal ideation. Also helpful diagnostically are a personal or family history of depression, anhedonia (loss of pleasure), and past response to an antidepressant. It is essential to bear in mind that depression in older patients is often caused or contributed to by drugs or a systemic illness. Although "subsyndromal" depression (fewer than four of the above symptoms) also causes substantial morbidity and health resource utilization, it appears to be less responsive than major depression to therapy.

TREATMENT

For the hospitalized patient in whom acute depression delays recovery or rehabilitation -- when correction of medical and pharmacologic contributing factors is ineffective and there is no prior history of mania or major depression -- methylphenidate, 5 to 10 mg at 8 A.M. and noon (to avoid insomnia) is often very effective, with benefits discernible within a few days. For patients with major depression, there is no ideal antidepressant drug. All are about equally effective, but the side effects differ (see below and Chap. 385). Consequently, one should become familiar with one or two agents for patients with psychomotor retardation (e.g., sertaline, desipramine) and for those with agitation (e.g., nortriptyline or nefazodone). Because of its potent anticholinergic and orthostatic side effects, amitriptyline should be avoided whenever possible in older patients. Initial low dosages should be increased slowly to avoid serious side effects; low doses of each medication (e.g., nortriptyline, 10 to 50 mg daily; desipramine, 25 to 75 mg daily; or sertraline 50 to 150 mg daily) are often effective in the elderly. Careful follow-up is required to anticipate and minimize anticholinergic side effects, orthostatic hypotension, sedating effects, confusion, bizarre mental symptoms, cardiovascular complications, and drug overdose with suicidal intent. Adverse drug reactions should not be assumed to be due to the aging process.

Cautious use of the monoamine oxidase inhibitors is sometimes of benefit when other antidepressants are ineffective. Neither monoamine oxidase inhibitors nor selective serotonin reuptake inhibitors should be used in combination with the cyclic compounds. Electroconvulsive therapy has been successful and is usually well tolerated by elderly patients who remain severely depressed despite drug treatment, particularly if they also have delusions.

URINARY INCONTINENCE

Transient Incontinence (Table 9-3) Because urinary continence requires adequate

mobility, mentation, motivation, and manual dexterity -- in addition to integrated control of the lower urinary tract -- problems outside the bladder can result in incontinence.

1. *Delirium*. A clouded sensorium impedes recognition of both the need to void and the location of the nearest toilet; once delirium clears, incontinence resolves.

2. *Infection*. Symptomatic urinary tract infection commonly causes or contributes to incontinence; asymptomatic infection does not.

3. *Atrophic urethritis/vaginitis*. Atrophic urethritis/vaginitis, characterized by the presence of vaginal telangiectasia, petechiae, erythema, or friability, commonly contributes to incontinence in women and responds to a several-month course of low-dose estrogen or vaginal estrogen creams.

4. *Pharmaceutical*. The drugs most commonly causing transient incontinence are listed inTable 9-4.

5. *Psychologic*. Depression and psychosis are uncommon but treatable causes.

6. *Excess urine output*. Excess urine output may overwhelm the ability to reach a toilet in time. Causes include diuretics, alcohol, excess fluid intake, and metabolic abnormalities (e.g., hyperglycemia, hypercalcemia, diabetes insipidus); nocturnal incontinence may also result from mobilization of peripheral edema.

7. *Restricted mobility*. If mobility cannot be improved, access to a urinal or commode may restore continence. (See "Immobility," below.)

8. *Stool impaction*. This is a common cause of urinary incontinence, especially in hospitalized or immobile patients. Although the mechanism is unknown, a clue to its presence is the coexistence of both urinary and fecal incontinence. Disimpaction restores continence.

Established Incontinence (Table 9-3) The causes of established incontinence include irreversible functional deficits, such as *end-stage* Alzheimer's disease, and intrinsic lower urinary tract dysfunction. Lower urinary tract dysfunction should be sought after transient causes have been excluded.

Detrusor Overactivity This disorder (involuntary bladder contraction) accounts for two-thirds of geriatric incontinence in both sexes, regardless of whether patients are demented. Detrusor overactivity can be diagnosed presumptively in a woman when leakage occurs in the absence of stress maneuvers or urinary retention and is preceded by the abrupt onset of an intense urge to urinate that cannot be forestalled. In men, the symptoms are similar, but since detrusor overactivity often coexists with urethral obstruction, urodynamic testing should be done if prescription of a bladder relaxant is planned. Because detrusor overactivity may also be due to bladder stones or tumor, the abrupt onset of otherwise unexplained urge incontinence -- especially if accompanied by perineal/suprapubic discomfort or sterile hematuria -- should prompt cystoscopy and cytologic examination.

TREATMENT

The cornerstone of treatment is behavioral therapy with or without biofeedback. Patients without dementia are instructed to void every 1 to 2 h (while awake only) and to suppress urgency in between; once daytime continence is restored, the interval between voiding can be progressively increased. Demented patients are "prompted" to void at similar intervals. When drugs are necessary, they should be added to these regimens and monitored to avoid inducing urinary retention. Effective drugs include oxybutynin (2.5 to 5 mg three or four times daily, or sustained release, 5 to 20 mg once daily), dicyclomine (10 to 30 mg three times daily), tolterodine (1 to 2 mg twice daily), and imipramine or doxepin (25 to 100 mg at bedtime). If prescribed for older patients, DDAVP should be used cautiously -- especially in the setting of renal insufficiency or heart failure -- and it probably should not be given to patients with hyponatremia or urine output >2500 mL/d. Alternative treatments, such as neuromodulation, are under investigation.

Indwelling catheterization is rarely indicated for detrusor overactivity. If all measures fail, an external collection device or protective pad or undergarment may be required.

Stress Incontinence This disorder, the second most common cause of established incontinence in older women (it is rare in men), is characterized by symptoms and evidence of *instantaneous* leakage of urine in response to stress. Leakage is worse or occurs only during the day unless another abnormality (e.g., detrusor overactivity) is also present. On examination, with the bladder full and the perineum relaxed, instantaneous leakage upon coughing strongly suggests stress incontinence, especially if it reproduces symptoms and if urinary retention has been excluded by a postvoiding residual determination; a several-second delay suggests that leakage is instead caused by an involuntary bladder contraction induced by coughing.

TREATMENT

Surgery is the most effective treatment. For women who can comply indefinitely, pelvic muscle exercises are an option for mild to moderate stress incontinence, but they often require specialized training using vaginal cones or biofeedback. If not contraindicated, an a-adrenergic agonist (e.g., phenylpropanolamine) is also helpful in such cases, especially if combined with estrogen. Occasionally, a pessary or even a tampon (for women with vaginal stenosis) provides some relief.

Urethral Obstruction Rarely present in women, urethral obstruction (due to prostatic enlargement, urethral stricture, bladder neck contracture, or prostate cancer) is the second most common cause of established incontinence in older men. It can present as dribbling incontinence after voiding, urge incontinence due to detrusor overactivity (which coexists in two-thirds of cases), or overflow incontinence due to urinary retention. Renal ultrasound is recommended to exclude hydronephrosis in men whose postvoiding residual volume exceeds 100 to 200 mL; in older men for whom surgery is planned, urodynamic confirmation of obstruction is strongly advised.

TREATMENT

Surgical decompression is the most effective treatment for obstruction, especially if there is urinary retention. For a nonoperative candidate, intermittent or indwelling catheterization is used; a condom catheter is contraindicated when urinary retention is present. For a man with prostatic obstruction who is not in retention, treatment with an a-adrenergic antagonist (e.g., terazosin 5 to 10 mg daily) may lessen symptoms in a few weeks. The 5a-reductase inhibitor finasteride may also ameliorate symptoms in a third or more of patients, but its impact is modest and not apparent for many months. Combined treatment with both agents has proved no better than treatment with an alpha blocker alone in most men.

Detrusor Underactivity Whether idiopathic or due to sacral lower motor nerve dysfunction, this is the least common cause of incontinence (<10% of cases). When it causes incontinence, detrusor underactivity is associated with urinary frequency, nocturia, and frequent leakage of small amounts. The elevated postvoiding residual volume (generally >450 mL) distinguishes it from detrusor overactivity and stress incontinence, but only urodynamic testing (rather than cystoscopy or intravenous urography) differentiates it from urethral obstruction in men; such testing is not usually required in women, in whom obstruction is rare.

TREATMENT

For the patient with a poorly contractile bladder, augmented voiding techniques (e.g., double voiding or applying suprapubic pressure) are often effective; pharmacologic agents (e.g., bethanechol) are rarely effective. If further emptying is needed or for the patient with an acontractile bladder, intermittent or indwelling catheterization is the only option. Antibiotics should be used for symptomatic upper tract infection, or as prophylaxis for recurrent symptomatic infections only in a patient using intermittent catheterization; they should not be used as prophylaxis with an indwelling catheter.

FALLS

Falls are a major problem for elderly people, especially women. Some 30% of community-dwelling elderly individuals fall each year, and the proportion increases with age. Nonetheless, falling must *not* be viewed as accidental, inevitable, or untreatable.

Causes of Falls Balance and ambulation require a complex interplay of cognitive, neuromuscular, and cardiovascular function and the ability to adapt rapidly to an environmental challenge. With age, balance becomes impaired and sway increases. The resulting vulnerability predisposes the older person to fall when challenged by an additional insult to *any* of these systems. Thus, a seemingly minor fall may be due to a serious problem, such as pneumonia or a myocardial infarction.

Much more commonly, however, falls are due to the complex interaction between a variably impaired patient and an environmental challenge. While a warped floorboard may pose little problem for a vigorous, unmedicated, alert person, it may be sufficient to precipitate a fall and hip fracture in the patient with impaired vision, strength, balance, or cognition. Thus, falls in older people are rarely due to a single cause, and effective prevention entails a comprehensive assessment of the patient's intrinsic deficits (usually diseases and medications), the routine activities, and the environmental obstacles.

Intrinsic deficits are those that impair sensory input, judgment, blood pressure regulation, reaction time, and balance and gait (Table 9-5). Medications and alcohol use are among the most common, significant, and reversible causes of falling. Other treatable contributors include postprandial hypotension (which peaks 30 to 60 min after a meal), insomnia, urinary urgency, foot problems, and peripheral edema [which can burden impaired leg strength and gait with an additional 2 to 5 kg (5 to 10 lb)].

Environmental obstacles are listed inTable 9-6. Since most falls occur in or around the home, a visit by a visiting nurse, physical therapist, or physician often reaps substantial dividends.

Complications of Falls and Treatment One out of four people who fall suffers serious injury. About 5% of falls result in fractures, and an equal proportion cause serious soft tissue damage. Falls are the sixth leading cause of death for older people and a contributing factor in 40% of admissions to nursing homes. Resultant hip problems and fear of falls are major causes of loss of independence.

Subdural hematoma is a treatable but easily overlooked complication of falls that must be considered in any elderly patient presenting with new neurologic signs, including confusion alone, even in the absence of a headache. Dehydration, electrolyte imbalance, pressure sores, rhabdomyolysis, and hypothermia may also occur and endanger the patient's life following a fall.

The risk of falling is related to the number of contributory conditions. Because the relationship is multiplicative rather than additive, however, even minor improvement in a number of these factors will reduce the risk substantially. In addition, gait training by a physical therapist often alleviates fear of falling. Ensuring the availability of phones at floor level, a portable phone, or a lightweight radio call system is also important, as is detection and treatment of osteoporosis.

IMMOBILITY

The main causes of immobility are weakness, stiffness, pain, imbalance, and psychological problems. Weakness may result from disuse of muscles, malnutrition, electrolyte disturbances, anemia, neurologic disorders, or myopathies. The most common cause of stiffness in the elderly is osteoarthritis; however, Parkinson's disease, rheumatoid arthritis, gout, pseudogout, and antipsychotic drugs such as haloperidol may also contribute. Pain, whether from bone (e.g., osteoporosis, osteomalacia, Paget's disease, metastatic bone cancer, trauma), joints (e.g., osteoarthritis, rheumatoid arthritis, gout), bursa, muscle (e.g., polymyalgia rheumatica, intermittent claudication, or "pseudoclaudication"), or foot problems may immobilize the patient.

Imbalance and fear of falling are major causes of immobilization. Imbalance may result from general debility, neurologic causes (e.g., stroke; loss of postural reflexes; peripheral neuropathy due to diabetes mellitus, alcohol, or malnutrition; and vestibulocerebellar abnormalities), orthostatic or postprandial hypotension, or drugs (e.g., diuretics, antihypertensives, neuroleptics, and antidepressants) or may occur following prolonged bed rest. Psychological conditions such as severe anxiety or

depression may also contribute to immobilization.

Consequences In addition to thrombophlebitis and pulmonary embolus, there are multiple hazards of bed rest in the elderly. Deconditioning of the cardiovascular system occurs within days and involves fluid shifts, fluid loss, decreased cardiac output, decreased peak oxygen uptake, and increased resting heart rate. Striking changes also occur in skeletal muscle. At the cellular level, intracellular ATP and glycogen concentrations decrease, rates of protein degradation increase, and contractile velocity and strength decline, while at the whole-muscle level, atrophy, weakness, and shortening are seen. Pressure sores are another serious complication; mechanical pressure, moisture, friction, and shearing forces all predispose to their development. As a result, within days of being confined to bed, the risk of postural hypotension, falls, and skin breakdown rises. Moreover, these changes usually take weeks to months to reverse.

TREATMENT

The most important step is preventive -- to avoid bedrest whenever possible. When it cannot be avoided, several measures can be employed to minimize its consequences. Patients should be positioned as close to the upright position as possible several times daily. Range-of-motion exercises should begin immediately, and the skin over pressure points should be inspected frequently. Isometric and isotonic exercises should be performed while the patient is in bed, and whenever possible patients should assist their own positioning, transferring, and self-care. As mobility becomes feasible, graduated ambulation should begin. For individuals confined to a wheelchair, ring-shaped devices ("donuts") should not be used to prevent pressure ulcers since they cause venous congestion and edema and actually increase the risk.

If a pressure ulcer develops, therapy depends on its stage. Stage 1 ulcers are characterized by nonblanchable erythema of intact skin; stage 2 lesions involve an ulcer of the epidermis, dermis, or both; stage 3 ulcers extend to the subcutaneous tissue; and stage 4 lesions involve muscle, bone, and/or the supporting tissues. For stage 1 lesions, eliminating excess pressure and ensuring adequate nutrition and hygiene are sufficient. For the remaining types, the caregiver must also ensure that the wound stays clean and moist; thus, if saline dressings are used they should be changed when they are damp rather than dry. Synthetic dressings are more expensive than saline but are more effective because they require fewer changes (with less disruption of reepithelialization) and protect against contamination. Because bacterial colonization of pressure ulcers is universal, swab cultures should not be performed and topical treatment should be considered only for patients whose ulcers have not healed after 2 weeks of therapy. By contrast, associated cellulitis, osteomyelitis, or sepsis requires systemic therapy after cultures of blood and the wound border (by needle aspiration or biopsy) have been obtained. Surgical or enzymatic debridement is required for stage 3 and 4 lesions. In addition to a daily multivitamin, prescribing vitamin C (500 mg twice daily) is also useful. For debilitated patients, special mattresses are beneficial, including those that reduce pressure (e.g., static air mattress or foam) and those that relieve it (e.g., dynamic units that sequentially inflate and deflate).

In addition to treating all identified factors that contribute to immobility, consultation with

a physical therapist should be sought. Installing handrails, lowering the bed, and providing chairs of proper height with arms and rubber skid guards may allow the patient to be safely mobile in the home. A properly fitted cane or walker may be helpful.

IATROGENIC DRUG REACTIONS

For several reasons, older patients are two or three times more likely to have adverse drug reactions (Chap. 71). Drug clearance is often markedly reduced. This is due to a decrease in renal plasma flow and glomerular filtration rate and a reduced hepatic clearance. The last is due to a decrease in activity of the drug-metabolizing microsomal enzymes and an overall decline in blood flow to the liver with aging. The volume of distribution of drugs is also affected, since the elderly have a decrease in total-body water and a relative increase in body fat. Thus, water-soluble drugs become more concentrated, and fat-soluble drugs have longer half-lives. In addition, serum albumin levels decline, particularly in sick patients, so that there is a decrease in protein binding of some drugs (e.g., warfarin, phenytoin), leaving more free (active) drug available.

In addition to impaired drug clearance, which alters pharmacokinetics, older patients have altered responses to similar serum drug levels, a phenomenon known as *altered pharmacodynamics*. They are more sensitive to some drugs (e.g., opiates, anticoagulants) and less sensitive to others (e.g.,b-adrenergic agents). Finally, the older patient with multiple chronic conditions is likely to be taking several drugs, including nonprescribed agents. Thus, adverse drug reactions and dosage errors are more likely to occur, especially if the patient has visual, hearing, or memory deficits.

Precautions to Avoid Drug Toxicity

Drug Selection and Administration Before initiating treatment, the physician should first ensure that the symptom requiring treatment is not itself due to another drug. For example, antipsychotic agents can cause symptoms that mimic depression (flat affect, restlessness, and pacing); such symptoms should prompt lowering of the dose rather than initiation of an antidepressant. In addition, drug therapy should be employed only after nonpharmacologic means have been considered or tried and only when the benefit clearly outweighs the risk.

Once pharmacotherapy has been decided upon, it should begin at less than the usual adult dosage and the dose should be increased slowly. However, given the marked variability in pharmacokinetics and pharmacodynamics in the elderly, dose escalation should continue until either a successful endpoint is reached or an intolerable side effect is encountered. The final dosage schedule should be kept as simple as possible, and the number of pills should be kept as low as possible. Serum drug levels are often useful in older patients, especially for monitoring drugs with narrow therapeutic indices such as phenytoin, theophylline, quinidine, aminoglycosides, lithium, and psychotropic agents such as nortriptyline. However, toxicity can occur even with "normal" therapeutic levels of some drugs (e.g., digoxin, phenytoin).

Over-the-Counter Agents Nearly three-quarters of the elderly regularly use nonprescribed drugs, many of which cause significant symptoms and/or interact with other medications. Frequent offenders include nonprescribed agents for insomnia (all of

which are anticholinergics), and nonsteroidal anti-inflammatory drugs (NSAIDs), which can hamper control of hypertension in addition to causing renal dysfunction and gastrointestinal bleeding. Gingko biloba, increasingly used as a "memory booster," may interfere with previously stable anticoagulation regimens. Because older patients often consider such agents "nostrums" rather than drugs, the physician must ask about them directly.

Sedative-Hypnotics If nonpharmacologic treatment of insomnia is unsuccessful, low-dose and short-term or intermittent use of an intermediate-acting agent whose metabolism is not affected by age (e.g., oxazepam, 10 to 30 mg/d) may be useful. Because of the increased risk of confusion and other adverse effects, benzodiazepines with either short (e.g., triazolam) or long duration of action (e.g., flurazepam and diazepam) should be avoided. Barbiturates should be avoided for the same reasons. An antidepressant should not be prescribed for insomnia unless the patient is depressed.

Antibiotics Serum creatinine is not a good index of renal function in old people; however, when it is elevated, special care must be taken with the administration of drugs normally excreted by the kidneys. Concentrations of relevant antibiotics should be measured directly.

Cardiac Drugs In older patients, digitalis, procainamide, and quinidine have prolonged half-lives and narrow therapeutic windows; toxicity is common at the usual dosages. For example, digoxin toxicity -- especially anorexia, confusion, or depression -- can occur even with therapeutic digoxin levels.

H_2Receptor Antagonists Most of these agents interfere with hepatic metabolism of other drugs, and all can produce confusion in the elderly. Because they are renally excreted, lower doses should be used to minimize the risk of toxicity in older individuals.

Antipsychotics and Tricyclic Antidepressants These drugs can produce anticholinergic side effects in old people (e.g., confusion, urinary retention, constipation, dry mouth). These can be minimized by switching to a nonanticholinergic agent (e.g., sertraline or nefazodone) or one with less anticholinergic effect (e.g., olanzapine, desipramine). In general, the least potent agents for psychosis (e.g., chlorpromazine) have the most sedating and anticholinergic effects and are the most likely to induce postural hypotension. By contrast, the most potent antipsychotic agents (e.g., haloperidol) have the least sedating, anticholinergic, and hypotensive side effects but cause extrapyramidal side effects, including dystonia, akathisia, rigidity, and tardive dyskinesia. The newer potent antipsychotics (e.g., risperidone, olanzapine, quetiapine, and clozapine) are relative exceptions to this rule. More specific for serotonin than dopamine D_2receptors, these medications may be safer for older demented patients, especially those with hallucinations associated with Lewy body dementia or in those receiving therapy for Parkinson's disease. Unfortunately, even these newer drugs lose their specificity at the higher doses that are commonly required in clinical practice. Thus all of these agents are potentially toxic. Moreover, since both depression and agitation often remit spontaneously, cautious discontinuation of these drugs should be considered periodically.

Glaucoma Medications Both topical beta blockers and carbonic anhydrase inhibitors can

cause systemic side effects. The latter can cause malaise and anorexia independent of the induced metabolic acidosis.

Anticoagulants Elderly patients benefit from anticoagulation as much as do younger individuals but are more vulnerable to serious bleeding and drug interactions. Hence, more careful monitoring and less aggressive anticoagulation are advisable.

Analgesics Both propoxyphene and meperidine are associated with a disproportionate risk of delirium, and propoxyphene also increases the risk of hip fracture. Of theNSAIDs, indomethacin is most likely to induce confusion, fluid retention, and gastrointestinal bleeding. Each of these agents should be avoided in the elderly.

Avoidance of Overtreatment Drugs are frequently not indicated in some common clinical situations. For instance, antibiotics need not be given for asymptomatic bacteriuria unless obstructive uropathy, other anatomic abnormalities, or stones are also present. Ankle edema is often due to venous insufficiency, drugs such asNSAIDsor some calcium antagonists, or even inactivity or malnutrition in chairbound patients. Diuretics are usually not indicated unless edema is associated with heart failure. Fitted, pressure gradient stockings are often helpful. Regular exercise is much more useful for claudication than is pentoxifylline. Finally, since older patients generally tolerate aspirin and other NSAIDs less well than do younger patients, localized pain should be treated when possible with local measures such as injection, physical therapy, heat, ultrasound, or transcutaneous electrical stimulation (Chap. 12).

PREVENTION

Much can be done to prevent the progression and even the onset of disease in older people. Dietary inadequacies should be corrected. Daily calcium intake should approximate 1500 mg, and most elderly people should take 400 to 800 IU of vitamin D daily (contained in one to two multivitamin tablets). Tobacco and alcohol use should be minimized, since the benefits of discontinuing these accrue even to individuals over age 65. The importance of reviewing all of a patient's medications and discontinuing them whenever feasible cannot be overemphasized.

Hypertension, whether isolated systolic hypertension or combined systolic and diastolic hypertension, should be treated. Treatment reduces the risk of stroke and the risk of death due to cardiovascular causes substantially in this age group and may also reduce the risk of cognitive impairment. These benefits have been achieved using *low doses* of a thiazide-like diuretic (e.g., chlorthalidone, 12.5 to 25 mg/d) as the first step (alone effective in almost half of patients) and adding low-dose reserpine (0.05 to 0.1 mg/d) or atenolol (25 to 50 mg/d) only as needed. Benefits are dramatic, side effects are minimal, cost is trivial, and concerns about potential toxicity have not been borne out.

Because of the prevalence, functional impact, and ease of treatment, glaucoma should be screened for, and visual and auditory impairment should be corrected. Dentures should be assessed for their fit, and oral lesions beneath them should be detected.

Because thyroid dysfunction is more prevalent in the elderly, difficult to detect clinically, and treatable, serum levels of thyroid-stimulating hormone should be measured at least