I-3

Fletcher RH, Fletcher SW, Wagner EH.
Clinical Epidemiology. The essentials,
2nd ed., Williams & Wilkins,
Baltimore, 1988.

chapter 1

INTRODUCTION

A 51-year-old man sees a physician because of chest pain. He had been well until 2 weeks ago, when he noticed tightness in the center of his chest when he was walking uphill. The tightness stopped after 2 to 3 minutes of rest. A similar discomfort occurred several times since then, sometimes during exercise and sometimes at rest. He smokes one pack of cigarettes per day and has been told in the past that his blood pressure is "a little high." He is otherwise well and takes no medications. However, he is worried about his health, particularly about coronary disease. A complete physical examination and resting electrocardiogram are normal except for a blood pressure of 150/96.

This patient is likely to have many questions. Am I sick? How sure are you? What is causing my illness? How will it affect me? What can be done about it?

The clinician caring for this patient must respond to these questions, considering them in their full complexity. Is the probability of serious, treatable disease high enough to proceed immediately beyond simple explanation and reassurance? How well might various diagnostic tests distinguish among the possible causes of chest pain: angina pectoris, esophageal spasm, muscle strain, anxiety, and the like? Specifically, how helpful will an exercise electrocardiogram be in either confirming or ruling out coronary artery disease? If coronary disease is found, how long can the patient expect to have the pain, and to live? How likely is it that other complications—congestive heart failure, myocardial infarction, or atherosclerotic disease of other organs—will occur? Will reduction of his risk factors for coronary disease—cigarette smoking and hypertension—reduce his risk? If medications control the pain, should the patient have coronary artery bypass surgery anyway to prevent untimely death?

As clinicians, we use various sources of information to answer these questions: our own experiences, the experiences of our colleagues, and the medical literature. In general, we depend on past observations on other

patients to predict what will happen to the patient at hand. The manner in which such observations are made and interpreted frequently determines whether the conclusions we reach are valid.

CLINICAL EPIDEMIOLOGY

Clinical epidemiology is one approach to making and interpreting scientific observations in medicine. *Clinical epidemiology* is the application of epidemiologic principles and methods to problems encountered in clinical medicine. It is a science concerned with counting clinical events occurring in intact human beings, and it uses epidemiologic methods to carry out and analyze the count.

Types of questions addressed by clinical epidemiology are listed in Table 1.1. By and large, these are the same questions confronting the doctor and patient in the example presented at the beginning of this chapter. They are at issue in most doctor-patient encounters.

The clinical events of primary interest in clinical epidemiology are the health outcomes of particular concern to patients and those caring for them (Table 1.2). They are the events doctors try to understand, predict, interpret, and change when caring for patients. Thus, an important distinction between clinical epidemiology and other sciences contributing to medicine is that the events of interest in clinical epidemiology can be studied directly only in intact humans and not in animals or parts of humans, such as tissue cultures, pituitary hormones, or red cell membranes.

Table 1.1
Clinical Issues and Questions in the Practice of Medicine

ISSUE	QUESTION
Normality/abnormality	Is a person sick or well? What abnormalities are associated with having a disease?
Diagnosis	How accurate are diagnostic tests or strategies used to find a disease?
Frequency	How often does a disease occur?
Risk	What factors are associated with an increased likelihood of disease?
Prognosis	What are the consequences of having a disease?
Treatment	How does treatment change the future course of a disease?
Prevention	Does intervention on people without disease keep disease from arising? Does early detection and treatment improve the course of disease?
Cause	What conditions result in disease? What are the pathogenetic mechanisms of disease?

Table 1.2
Health Outcomes (The Five D's)[a]

Death	A universal health outcome, the timeliness of the event being the issue.
Disease	A combination of symptoms, physical signs, and laboratory test results.
Disability	The functional status of patients in terms of ability to live independently and go about their daily lives at home, work, or recreation.
Discomfort	Uncomfortable symptoms, such as pain, nausea, vertigo, tinnitus, or fatigue.
Dissatisfaction	Emotional and mental states, such as agitation, sadness, or anger.

[a] Some suggest a sixth "D"—destitution—because physicians should be concerned with financial consequences of health care to their patients. Others have pointed out that the five D's emphasize the negative side of health outcomes. Nevertheless, the five D's do remind physicians that clinical events other than death and disease are important.

Many of the methods used to count the clinical events were developed in the field of epidemiology, which has been defined as "the study of the distribution and determinants of disease frequency in man" (1). These methods are the subject of this book.

The basic purpose of clinical epidemiology, then, is to develop and apply methods of clinical observation that will lead to valid clinical conclusions. For clinicians who intend to make up their own minds about the soundness of clinical information, some understanding of this field is as necessary as an understanding of anatomy, pathology, biochemistry, and pharmacology. Indeed, clinical epidemiology is one of the "basic sciences," a foundation on which modern medicine is practiced.

Even so, the principles of clinical epidemiology are not yet second nature to most clinicians. This is partly related to underlying differences between the parent disciplines, clinical medicine and epidemiology. Both are scientific approaches to the causes and consequences of disease in humans. Both have a practical bent, by and large seeking to discover information that can be useful in the control of disease and alleviation of suffering. However, there are some major differences between them as well. When we have explored these differences, we will consider how they can be brought together for the purpose of enriching our understanding of clinical medicine.

Clinical Medicine

As most readers of this book know, clinicians are, by and large, concerned with individual patients. They meet all of their patients personally, take their own histories, and do their own physical examinations. Usually they are not particularly interested in how patients happen to be found in their practices, as opposed to some other medical setting. As a result, clinicians may not feel particularly responsible for other patients, who may be just as sick but have not come to their attention.

Because clinicians bear intense personal responsibility for individual patients, they tend to see what is special about each patient. Thus, it is not surprising that clinicians are often reluctant to lump patients into crude categories of risk, diagnosis, or treatment. Many clinicians are uneasy about uncertainty and reluctant to express it through probabilities. In their work, the stakes commonly are high so that both they and their patients want as much certainty as possible.

Clinical training is largely oriented toward the mechanisms of disease through study of biochemistry, anatomy, physiology, and other basic sciences. These traditional basic sciences are powerful influences in a medical student's formative years, and go on to become the predominant forces in clinical research and publications. This fosters the belief that to understand the detailed processes of disease in individual patients is to understand medicine. The implication is that one can predict the course of disease and select appropriate treatments through knowledge of the mechanisms of disease.

Table 1.3 summarizes both biologic and clinical outcomes for the modern treatment of a patient with acute myocardial infarction. Studies of the biology and pathophysiology of coronary artery disease have greatly increased our understanding of the process. However, for clinical practice such understanding is important only to the extent that these processes are known to be related to the kinds of health outcomes listed in the table.

Epidemiology

Epidemiology is the research discipline concerned with the distribution and determinants of disease in populations. Epidemiologists differ from clinicians in several ways. In contrast to clinicians, epidemiologists are more concerned that their observations represent fairly some defined group of people, or "population." They wish to record the experiences of all members of the group, whether or not they are sick, and whether or not they have come to medical attention. Epidemiologists usually do not personally collect all their data themselves nor do they usually meet the people they study. Because they work with groups rather than with individual patients, epidemiologists are more comfortable with probability.

When epidemiologists study disease, their categories are often crude by clinical standards. They work with such variables as "cigarette smokers" and "nonsmokers" or "sudden death" and "myocardial infarction," even though a myriad of special circumstances is hidden within these classes.

Table 1.3
Biologic and Clinical Outcomes of Medical Care: Treatment of Acute Myocardial Infarction

DISEASE	INTERVENTIONS	OUTCOMES	
		BIOLOGIC	CLINICAL
Acute myocardial infarction	Coronary care Streptokinase Angioplasty	Coronary patency Infarction size Ejection fraction Serum enzymes ECG changes	Death Congestive heart failure Reinfarction Angina
		⟵ Association known or assumed? ⟶	

Epidemiologists are generally more interested in whether something occurs than in how it occurs at a pathogenetic, mechanistic level. They rely in part on an understanding of the mechanism of disease to form hypotheses and then test these hypotheses in human populations. However, if a pathogenetic mechanism is not known, epidemiologists do not necessarily discard a hypothesis. For example, if it can be shown that cigarette smoking in itself is related to cardiovascular disease, and the risk of heart disease decreases when smoking ceases, everything else being equal, the epidemiologist might consider the role of smoking in heart disease largely settled. The more mechanistically inclined biomedical researcher might remain dissatisfied until the causative agent in cigarette smoke is isolated and the pathway by which it causes heart disease is specified.

Because epidemiologists work with groups of individuals, each of which displays unique genetic and environmental characteristics, their stock-in-trade is to know how to deal with variables that are extraneous to the main question at hand. Along the way, they work closely with biostatisticians, who use statistics to summarize the experience of groups, adjust for extraneous differences between groups being compared, and assess whether chance could have determined the findings.

Clinical Epidemiology

Clinical medicine and epidemiology began together. The founders of epidemiology were, for the most part, clinicians. It is only during the past several decades that the two drifted apart, with separate schools, training, journals, and opportunities for employment.

In recent years, clinicians and epidemiologists have become increasingly aware that their fields interrelate and clinical epidemiology began to develop, in recognition of the following:

- In most clinical situations the diagnosis, prognosis and results of treatment are uncertain for individual patients and therefore must be expressed as probabilities.

- Probability for an individual patient is best estimated by referring to past experience with groups of similar patients.
- Because clinical observations are made on people who are free to do as they please, by clinicians with variable skills and prejudices, the observations may be influenced by a variety of systematic errors that can distort the true nature of events and thereby be misleading.
- To deal with these misleading effects, clinical observations should be based on sound scientific principles, including ways to reduce bias and estimate the role of chance.
- These principles are as important to clinicians who wish to be self-sufficient in judging clinical information as they are to researchers who will produce research.

Although the principles of scientific methods are implicit in a great deal of current clinical teaching, they are not a formal part of most medical curricula. But this is changing. A growing understanding of the relevance of epidemiology to clinical medicine is reflected in the recent development of clinical epidemiology courses for medical students, the increased numbers of articles in clinical journals dealing with precepts of clinical epidemiology, and the development of fellowship programs for training clinicians in clinical epidemiology (2).

We hope that this book will contribute to the growth of clinical epidemiology by bringing together, in a few pages, the basic rules of evidence that are important in the clinician's world. We also hope that this book will help clinicians evaluate the constant flow of new medical information. This means that clinicians must learn about the various research designs used in medical research and something about their strengths and pitfalls. With such knowledge, conscientious clinicians can develop skills to help them decide whether new information is worthy of the effort to master it in the first place. For those of us with limited, already oversaturated medical memories, such a skill is indeed useful.

BASIC PRINCIPLES

The basic purpose of clinical epidemiology is to foster methods of clinical observation and interpretation that lead to valid conclusions. This activity is grounded in basic, scientific principles. The remainder of this chapter will introduce these principles. The rest of the book will build on this introduction.

Populations and Samples

In general, *populations* are large groups of people in a defined setting. These include relatively unselected people in the community, the usual population for epidemiologic studies of cause; and groups of people selected because of their attendance in a clinic or hospital or because of a characteristic such as age, race, or the presence of disease. Thus, one speaks of

the general population, a hospitalized population, or a population of patients with a specific disease.

A *sample* is a subset of a population and is selected from it. Clinical research is ordinarily carried out on samples. One is interested in the characteristics of the defined population, but must, for practical reasons, describe them through a sample. In doing so, two fundamental questions arise. First, are the conclusions of the research correct for the people in the sample? Second, if so, does the sample represent fairly the population of interest?

Whenever a clinical question is answered by observing people, there are three possible explanations for the answer (Table 1.4). The observation may be incorrect because of bias or chance, or it may be correct.

Bias

Bias is "a process at any stage of inference tending to produce results that depart systematically from the true values" (3).

Suppose, for example, that treatment A is found to work better than treatment B. What kinds of biases might explain the observation? Perhaps A is given to healthier patients than B: then the results could be due to the systematic difference in health between the groups of patients rather than to differences in treatment. Or A might taste better than B so that patients take the drug more regularly. Or A might be a new, very popular drug and B an old one, so that researchers and patients might be more inclined to think that the new drug works better whether or not it really does. All of these are examples of potential biases.

Compared to basic science research, observations on patients (whether for patient care or research) are particularly susceptible to bias. The process tends to be just plain untidy. As participants in a study, human beings have the disconcerting habit of doing as they please and not necessarily what would be required for scientific rigor. When one attempts to conduct

Table 1.4
Possible Explanations for Clinical Observations

BIAS (systematic error)
Selection bias occurs when comparisons are made between groups of patients that differ with respect to determinants of the outcome other than those under study.
Measurement bias occurs when the methods of measurement are consistently dissimilar among groups of patients.
Confounding bias occurs when two factors or processes are associated or "travel together," and the effect of one is confused with or distorted by the effect of the other.
CHANCE (random error)
Because of random variation, the characteristics of people in a particular sample are different from others in the population from which it was taken.
TRUTH
The observation is correct. (Accept this explanation only after excluding the others!)

an experiment with them, as one might in a laboratory, things tend to go wrong. Some people refuse to participate, while others drop out or choose another treatment. What is more, some of the most important things about humans—feelings, comfort, performance—are generally more difficult to measure than physical properties, such as blood pressure or serum sodium. The methods of measuring these phenomena are less direct. Then, too, there is the normal inclination of clinicians to believe that their therapies succeed. (Most patients would not want a physician who felt otherwise.) This attitude, so important in the practice of medicine, makes clinical observations particularly vulnerable to bias.

Although dozens of particular biases have been defined, most fall into one of three broad categories (Table 1.4):

1. ***Selection bias*** **occurs when comparisons are made between groups of patients that differ with respect to determinants of the outcome other than those under study.**

Groups of patients often differ in many ways. If we compare the experience of two groups to decide about the effects of one of their characteristics, free of the others, and the compared groups differ in important ways (other than the factor of interest) at the time of selection, the comparison is biased. As a result little can be concluded about the independent effects of the characteristic of interest.

2. ***Measurement bias*** **occurs when the methods of measurement are consistently dissimilar among groups of patients.**

An example of a potential measurement bias would be in the use of information taken from medical records to determine if women on birth control pills were more at risk for thromboembolism than those not on "the pill." Suppose a study were made comparing the frequency of oral contraceptive use among women admitted to a hospital because of thrombophlebitis and a group of women admitted for other reasons. It is entirely possible that women, aware of the reported association between estrogens and thrombotic events, might report use of oral contraceptions more completely if they had thrombophlebitis. For the same reasons, clinicians might obtain and record information about oral contraceptive use more thoroughly for women with phlebitis than for those without phlebitis. If so, an observed association between oral contraceptives and thrombophlebitis could be due to the biased way in which the history of exposure was reported.

3. ***Confounding bias*** **occurs when two factors or processes are associated or "travel together," and the effect of one is confused with or distorted by the effect of the other.**

Example—Is diabetes mellitus, in and of itself, a risk factor for atherosclerotic disease? The clinical manifestations of atherosclerosis—myocardial infarction, sudden death, peripheral vascular disease, stroke, and others—are all more likely to occur in patients with diabetes. But other risk factors for cardiovascular disease—blood pressure, weight, and serum cholesterol—are all higher in diabetics. Perhaps the increased risk in diabetics is because of these other factors, rather than the diabetes itself. However, even after the effects of these other factors are taken into account, diabetics are at increased risk (4).

Example—Several studies have shown that serum triglyceride levels (TG) are associated with risk for coronary heart disease (CHD): the higher the TG, the higher the risk. Because of this, clinicians have screened for TG and attempted to lower TG when it was elevated. This might be helpful if TG were an independent cause of CHD. But other known causes of CHD, particularly elevated serum cholesterol levels and reduced levels of high density lipoprotein, are related to both serum triglycerides and CHD. When the contribution of these other factors is removed, the relationship between TG and CHD no longer is present. It seems unlikely, therefore, that TG is an independent cause of CHD; its relationship to CHD is confounded with other factors that are independent causes (5) (Fig. 1.1).

It should be apparent that selection bias and confounding bias are not mutually exclusive. They are described separately, however, because they present problems at different points in a clinical observation or study. Selection bias is at issue primarily when patients are chosen for observation, and so it is important in the design of a study. Confounding bias comes to the fore in analysis of the data, once the observations have been made.

Often in the same study more than one bias operates, as in the following hypothetical example.

Example—A study was done to determine whether regular exercise lowers the risk of CHD. An exercise program was offered to employees of a plant, and the rate of subsequent coronary events was compared among employees who volunteered for the program and those who did not volunteer. Coronary events were determined by means of regular voluntary checkups, including a careful history and electrocardiogram, and review of routine health records. The group that exercised had lower rates of CHD. However, they also smoked cigarettes less.

In this example, selection bias could be present if volunteers for the exercise program were, for any reason, at lower risk for coronary disease even before the program began—e.g., because they had lower serum lipids or less family history of coronary disease. Measurement bias might have occurred because the exercise group stood a better chance of having a coronary event detected, inasmuch as more of them were likely to be

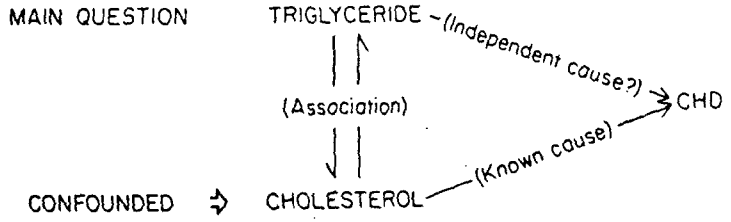

Figure 1.1. Confounding bias. Serum triglyceride is a risk factor for coronary heart disease (CHD) but not independently of serum cholesterol.

examined routinely. Finally, the conclusion that exercise lowered the risk of coronary disease might be the result of a confounding bias if the association between exercise and coronary events in this particular study resulted from the fact that smoking cigarettes is a risk factor for coronary disease and the exercise group smoked less.

A distinction must be made between the potential for bias and the actual presence of bias in a particular study. It is first necessary to recognize the importance of bias and to know where and how to look for it and what can be done about it. But one should not stop there. It is also necessary to determine whether bias is in fact present and how large it is likely to be, in order to decide whether it is important enough to change the conclusions of the study in a clinically meaningful way.

Chance

Observations about disease are ordinarily made on a sample of patients rather than all those with the disease in question. A single set of observations, even if selected in an unbiased way, may misrepresent the truth because of error arising from random variation. In fact, actual observations on a single sample are unlikely to correspond exactly to the true state of affairs in the larger group of all patients. However, if the observations were repeated on many such samples, they would be found to vary about the true value. The divergence of an observation on a sample from the true population value, due to chance alone, is called *random variation*.

We are all familiar with chance as an explanation for why a coin does not come up heads exactly 50% of the time when it is flipped, say, 100 times. The same influence, random variation, applies when assessing the effects of treatments A and B, discussed earlier. Suppose all biases were removed from a study of the relative effects of two treatments. Suppose, further, that the two treatments were, in reality, equally effective, each improving about 50% of the patients treated. Even so, because of chance alone, a single study involving small numbers of patients in each treatment group might easily find A improving a much larger proportion of patients than B, or vice versa.

Chance can affect all of the steps involved in clinical observations. In the assessment of treatments A and B, random variation can occur in the sampling of patients for the study, the selection of treatment groups, and the measurements made on the groups.

Unlike bias, which deflects values in one direction or another, random variation is as likely to result in observations above the true value as below. As a consequence, the mean of many unbiased observations on samples tends to correspond to the true value in the population, even though the results of individual small samples may not.

Statistics can be used to estimate the probability of chance or random variation accounting for clinical results. A knowledge of statistics can also help reduce that probability. It is important, however, to understand that random variation cannot be totally eliminated. Chance should always be considered when assessing the results of clinical observations.

Relationship between Bias and Chance

The relationship between bias and chance is illustrated in Figure 1.2. The measurement of diastolic blood pressure on a single patient is taken as an example. True blood pressure can be obtained by an intra-arterial cannula, and multiple readings are illustrated as all being 80 mm Hg. But this method is not possible for routine measurements. Blood pressure is ordinarily measured indirectly, using a sphygmomanometer. The simpler instrument is prone to error, or deviations from the true value. In the example, this is represented by all of the sphygmomanometer readings falling to the right of the true value. The deviation of sphygmomanometer readings to the right (bias) may have several explanations—e.g., a poorly calibrated sphygmomanometer, a wrong cuff size, or a deaf clinician. Bias could also result if different sounds were chosen to represent diastolic blood pressure. The usual end points—phase IV and phase V Korotkoff sounds—tend to be above and below the true diastolic pressure, respectively; and even that is unpredictable in obese people. Individual blood pressure readings would also be subject to error because of random variation in measurement, as illustrated by the spread of the sphygmomanometer readings around the mean value (90 mm Hg).

The two sources of error—bias and chance—are not mutually exclusive. In most situations, both are present. The main reason for distinguishing between the two is that they are handled differently.

Bias, in theory, can be prevented by conducting clinical investigations

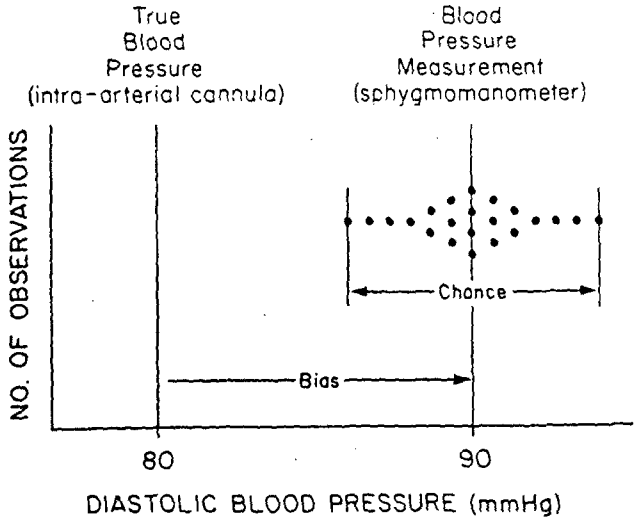

Figure 1.2. Relationship between bias and chance. Blood pressure measurements by intra-arterial cannula and sphygmomanometer.

properly or corrected through proper data analysis. If not eliminated, bias often can be detected by the discerning reader. Most of this book is about how to recognize, avoid, or minimize bias.

Chance cannot be eliminated, but its influence can be reduced by proper design of research, and the remaining error estimated by statistics. Statistics can also help remove the effects of known biases. However, no amount of statistical treatment can correct for unknown biases in data. Some would go so far as to prefer that statistics not be applied to data vulnerable to bias because of poor research design, for fear of giving false respectability to misleading work.

Validity

A clinical observation is valid if it corresponds to the true state of the phenomenon being measured. For the observation to be valid, it must be neither biased nor incorrect due to chance. There are two general kinds of validity—internal validity and external validity, or generalizability.

Internal validity is the degree to which the results of an observation are correct for the patients being studied. It is "internal" because it applies to the particular conditions of the particular group of patients being observed, and not necessarily to others. The internal validity of clinical observations is determined by how well they are carried out, and is threatened by all of the biases and random variation discussed previously. For a clinical observation to be useful, internal validity is a necessary but not sufficient condition.

External validity (*generalizability*) is the degree to which the results of an observation hold true in other settings. For an individual physician, it is an answer to the question: "Assuming that the results of a study are true, do they apply to my patient as well?" Generalizability expresses the validity of assuming that patients in a study are comparable to other patients.

An unimpeachable study, with high internal validity, may be totally misleading when the results are generalized to certain other patients. This is because of yet another bias, sampling bias. *Sampling bias* occurs when observations and conclusions based on a sample of people are generalized to other groups who are not similar.

Example—Sampling bias was demonstrated in studies of febrile seizures in children. Because febrile seizures commonly occur in childhood (reportedly in 2–4% of all children) it is important to know if such seizures recur. If they frequently recur, treatment with anticonvulsant therapy may be worth considering. On the other hand, if febrile seizures are usually one-time phenomena, reassurance of the parents is in order.

Figure 1.3 shows the recurrence rates of seizures among children reported in various studies, according to how the children were chosen for the study. On the left, population-based studies followed up all children who had febrile seizures in the defined population. The recurrence rates reported in these studies were all very low. Clinic-based studies, shown on the right of Figure 1.3, described recurrence

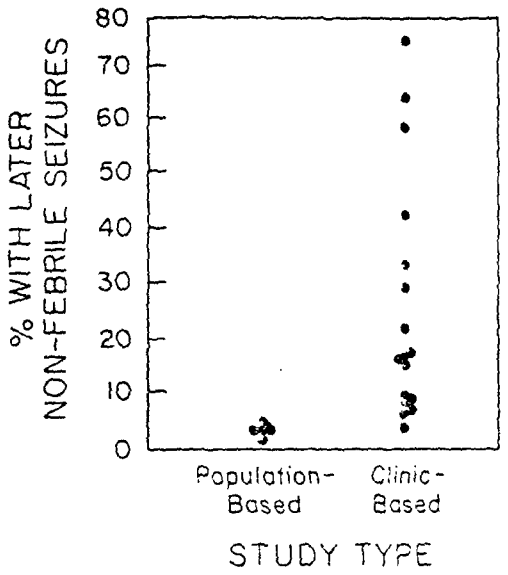

Figure 1.3. Example of sampling bias. Recurrent seizures in infants with febrile seizures in population-based and clinic-based studies. (Redrawn from Ellenberg JH, Nelson KB: Sample selection and the natural history of disease. Studies of febrile seizures. *JAMA* 243:1337–1340, 1980.)

rates in children attending hospital clinics or specialty referral units, and reported much higher recurrence rates.

Applying the results of referral hospital studies to community practices would result in falsely high estimates of the likelihood of recurring seizures. On the other hand, using the conclusion of the population-based studies when discussing the prognosis of febrile seizures with the parents of a child referred to a pediatric neurology unit might also be inaccurate, this time in the opposite direction (6).

The generalizability of clinical observations, even those with internal validity, is a matter of opinion about which reasonable people might disagree. For example, the Veterans Administration (VA) study of hypertension treatment was carried out with special attention to potential biases and chance. As a result, the internal validity of this study, with its conclusion that lowering blood pressure decreased risk of death and severe morbid events in the particular patients in the study, is generally accepted (7). But the study was confined to men. Some clinicians were willing to use the results of this study to guide decisions about treatment of women. Others, more skeptical because of known differences in the frequency of cardiovascular disease between men and women, were unwilling to do so. The VA study itself held no solution to this disagreement. The only way

to resolve the dispute was to collect data on women with hypertension. Such studies have subsequently been done and found similar results for women.

Generalizability can rarely be dealt with satisfactorily in any one study. Even a defined, geographically based population is a biased sample of larger populations; for example, hospital patients are biased samples of county residents, counties of states, states of regions, and so on. Adding additional centers may improve generalizability, but not settle the issue.

Usually, the best a researcher can do about generalizability is to ensure internal validity, have the study population fit the research question, and avoid studying groups so unusual that experience with them generalizes to few other settings. It then remains for other studies, in other settings, to extend generalizability.

Because most clinical studies take place in medical centers and because patients in such centers usually overrepresent the serious end of the disease spectrum, sampling bias in clinical research tends to exaggerate the serious nature of disease.

The relationships among internal and external validity, bias, and chance are shown in Figure 1.4.

DESCRIPTION AND COMPARISON

Clinical research can simply describe the frequency of a clinical phenom-

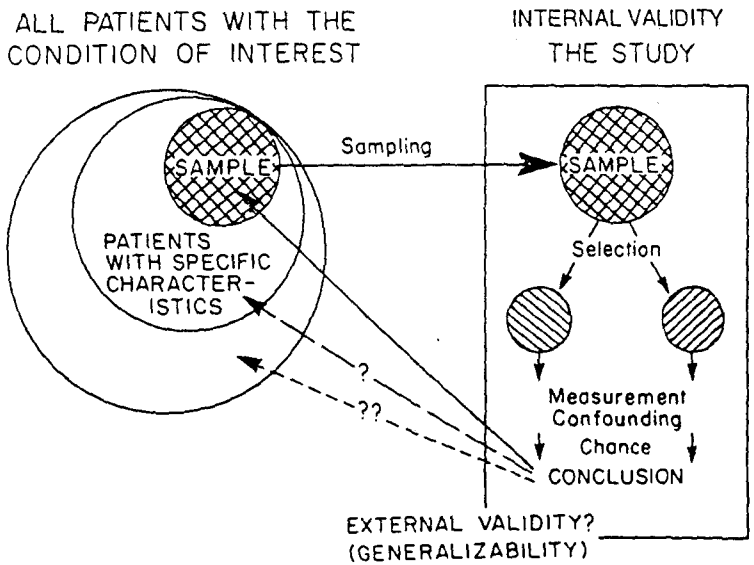

Figure 1.4. Relationships among internal and external validity, bias, and chance.

enon. For example, one might ask: What proportion of young women with dysuria have bacterial cystitis?; or what is the rate of stroke in people with carotid bruits? Simple descriptions are useful in their own right. They guide clinical predictions.

More often, a comparison is also made: Is cystitis more common in women with, as opposed to without, pyuria?; or are strokes more common in people with carotid bruits, compared to people without? In this process, one is looking for an association (or possibly a cause-and-effect relationship) between two or more factors.

When comparisons are made, bias and chance are acknowledged by two modifiers to the basic question: Is X associated with Y (*a*) everything else being equal? (*bias*); and (*b*) more often than would have been expected by chance alone? (*chance*)

INFORMATION AND DECISIONS

The primary concern of this book is the quality of information and its correct interpretation. Making decisions is another matter. True, good decisions depend on good information; but they involve a great deal more as well, including value judgments and the weighing of competing risks and benefits.

In recent years, various methods for "quantitative" decision making have become popular. Among these are decision analysis, cost-benefit analysis, and cost-effectiveness analysis. They involve presenting the decision-making process in an explicit way, so that the components of the decision and the consequences of assigning various values to them can be examined. Some aspects of decision analysis, such as evaluation of diagnostic tests, are included in this book. However, we have elected not to go deeply into medical decision making itself. Our justification is that ultimately decisions are only as good as the information used to make them, and we have found enough to say about the essentials of collecting and interpreting clinical information to fill a book. Readers who wish to delve more deeply into decision analysis can begin with some of the readings suggested at the end of this chapter.

ORGANIZATION OF BOOK

In most textbooks of clinical medicine, information about disease is presented as answers to traditional clinical questions, as outlined in Table 1.1. On the other hand, most books about clinical investigation are organized around research strategies: clinical trials, surveys, case-control studies, etc. This way of organizing a book may serve those who would perform clinical research, but it is often awkward for clinicians. As a result, clinicians do not have a comprehensive source of information about the basic structure of clinical observations as it relates to clinical practice, as they do for the clinical relevance of the basic sciences.

This book is written for clinicians who wish to understand the methods

of clinical observation and research. We have not written primarily for those who would perform clinical research, but for all the rest who depend on it. However, we believe that the basic needs of those who create and those who use clinical research findings are similar.

We have organized the book primarily according to the clinical questions surrounding doctor-patient encounters. Figure 1.5 illustrates how these questions correspond to the book chapters, taking lung cancer as an example. The questions relate to the entire natural history of disease, from the time nondiseased people are first exposed to risk, through when some acquire the disease and emerge as patients, until the end results of disease are manifest.

In each chapter, strategies used to answer the clinical questions are described. Occasionally, a given strategy, such as a cohort study, may be

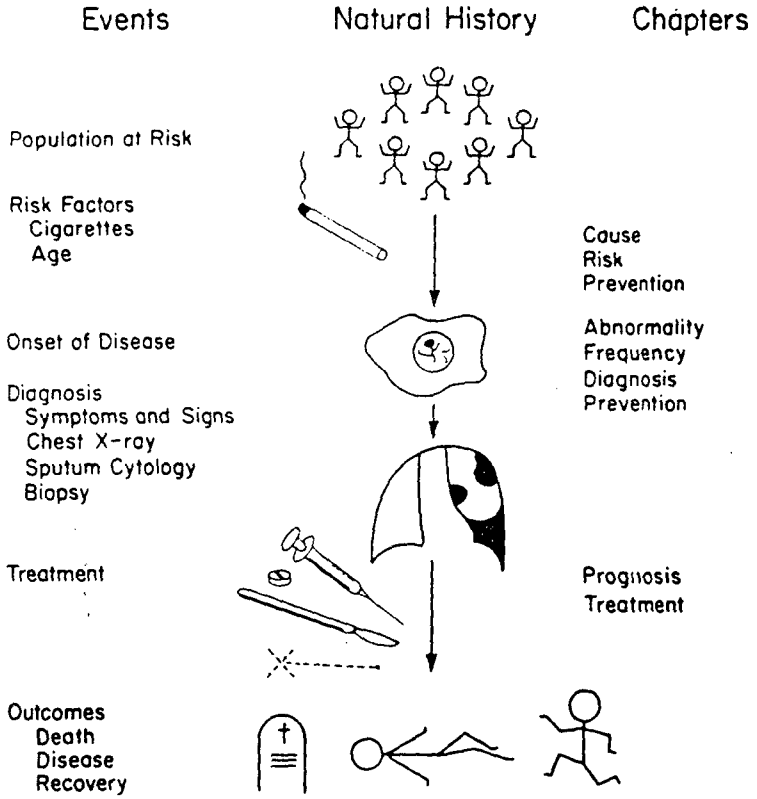

Figure 1.5. Organization of the book, illustrated for the disease lung cancer. Issues covered in the chapters on chance and cause relate to the entire progression of disease.

useful in answering several clinical questions. For the purposes of presentation, we have arbitrarily discussed these strategies primarily in one chapter. But it is important to keep in mind that given clinical epidemiologic methods may be applicable to more than one clinical question. When this is so, we have attempted to refer to these methods in other relevant chapters.

REFERENCES

1. Sackett DL, Haynes RB, Tugwell P: *Clinical Epidemiology: A Basic Science for Clinical Medicine*. Boston, Little, Brown, and Co, 1985.
2. Fletcher RH, Fletcher SW: Clinical epidemiology. A new discipline for an old art. *Ann Intern Med* 99:401–403, 1983.
3. Murphy EA. *The Logic of Medicine*. Baltimore, Johns Hopkins University Press, 1976.
4. Dawber TR: *The Framingham Study. The Epidemiology of Atherosclerotic Disease*. Cambridge, MA, Harvard University Press, 1980.
5. Hulley SB, Rosenman RH, Bawol RD, Brand RJ: Epidemiology as a guide to clinical decisions: The association between triglyceride and coronary heart disease. *N Engl J Med* 302:1383–1389, 1980.
6. Ellenberg JH, Nelson KB: Sample selection and the natural history of disease: Studies of febrile seizures. *JAMA* 243:1337–1340, 1980.
7. Veterans Administration Cooperative Study Group on Antihypertensive Agents. Effects of treatment on morbidity in hypertension. *JAMA* 202:1028–1034, 1967 and 213:1143–1152, 1970.

SUGGESTED READINGS

Cebul RD, Beck LH (eds): *Teaching Clinical Decision Making*. New York, Praeger, 1985.
Chalmers AF: *What is This Thing Called Science? An Assessment of the Nature and Status of Science and Its Methods*, ed 2. Queensland, Queensland University Press, 1982.
Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre. Clinical Epidemiology Rounds: How to read clinical journals I–V. *Can Med Assoc J* 124:555–558, 703–710, 869–872, 985–990, 1156–1162, 1981.
Feinstein AR: *Clinical Judgment*. Baltimore, Williams & Wilkins, 1967.
Feinstein AR: Why clinical epidemiology *Clin Res* 20:821–825, 1972.
Feinstein AR: *Clinical Biostatistics*. St. Louis, CV Mosby, 1977.
Feinstein AR: *Clinical Epidemiology. The Architecture of Clinical Research*. Boston, Little, Brown, and Co, 1985.
Fletcher RH, Fletcher SW: Clinical epidemiology, A new discipline for an old art. *Ann Intern Med* 99:401–403, 1983.
Hulley SB, Cummings SR: *Designing Clinical Research. An Epidemiologic Approach*. Baltimore: Williams & Wilkins, in press.
Jenicek M, Clèroux R: *Épidémiologie Clinique. Clinimétrie*. St-Hyacinthe, Quebec, Edisem, 1985.
Michael M III, Boyce WT, Wilcox AJ: *Biomedical Bestiary: An Epidemiologic Guide to Flaws and Fallacies in the Medical Literature*. Boston: Little, Brown and Co, 1984.
Research Development Committee, Society for Research and Education in Primary Care Internal Medicine: Clinical research methods: An annotated bibliography. *Ann Intern Med* 99:419–424, 1983.
Riegelman RK: *Studying a Study and Testing a Test. How to Read the Medical Literature*. Boston, Little, Brown and Co, 1981.
Sackett DL: Clinical epidemiology. *Am J Epidemiol* 89:125–128, 1969.
Sackett DL: Bias in analytic research. *J Chron Dis* 32:51–63, 1979.
Sackett DL: Three cheers for clinical epidemiology. *Int J Epidemiol* 13:117–119, 1984.

Sackett DL, Haynes RB, Tugwell P: *Clinical Epidemiology. A Basic Science for Clinical Medicine.* Boston, Little, Brown, and Co, 1985.
Sox HC Jr, Blatt M, Higgins MC, Marton KI: *Medical Decision Making.* London, Butterworths, in press.
Spitzer WO. Clinical epidemiology. *J Chron Dis* 39:411–415, 1986.
Weinstein MC, Fineberg HV, Elstein AS, Frazier HS, Neuhauser D, Neutra RR, McNeil BJ: *Clinical Decision Analysis.* Philadelphia, WB Saunders, 1980.
Weiss NS. *Clinical Epidemiology: The Study of the Outcome of Illness.* New York, Oxford University Press, 1986.

APPENDIX 1.1. QUESTIONS THAT APPLY TO MOST CLINICAL RESEARCH

1. The research question
   a. What general *kind of clinical question* is the research intended to answer: abnormality, diagnosis, frequency, risk, prognosis, treatment, or prevention? (Appropriate research design depends on the question.)
   b. What is the *specific question*?
2. Generalizability/importance
   a. Were the *patients* in the study like the ones I see?
   b. If *predictive factors* (risk or prognostic factors, prevention, or therapeutic interventions) were studied, are they present in my setting?
   c. Were the *outcomes* important to my patients and me?
3. Bias
   a. Could the observed findings be the result of bias? (See appendices at the end of Chapters 3, 4, 6, and 10.)
   b. If bias is present, is it large enough to affect the conclusions?
4. Chance
   a. If the study found a difference, how likely is it to have occurred by chance alone (alpha error), either for single comparison (probability of error estimated by a *p* value) or as a result of multiple comparisons?
   b. If the study found no difference, how likely was it to detect a clinically important difference? (*statistical power*)
   c. If the study reports a rate, or difference in rates, what is the *precision* of the observation—the range that is likely to include the true rate (confidence intervals).

chapter 2

ABNORMALITY

Clinicians spend a great deal of time distinguishing "normal" from "abnormal" biology. When confronted with something grossly different from the usual, there is of course little difficulty telling the two apart. We are all familiar with pictures in classic textbooks of physical diagnosis showing obvious examples of massive hepatosplenomegaly, goiter, or elephantiasis. We can take no particular pride in recognizing this degree of abnormality. More often, however, clinicians must make subtler distinctions between normal and abnormal. Is fleeting chest pain pleurisy or inconsequential? Is a soft systolic heart sound a sign of valvular heart disease or an innocent murmur? Is a slightly elevated serum alkaline phosphatase evidence for liver disease, asymptomatic Paget's disease, or nothing at all?

Decisions about what is abnormal are most difficult among unselected patients usually found outside the hospital. When patients have already been screened and selected for special attention, as is the case in most referral centers, it is usually clear that something is wrong. The task is then to refine a description of the problem and treat it. In primary care settings, however, subtle manifestations of disease are freely mixed with more everyday complaints of healthy people, and it is not possible to pursue all of those complaints aggressively. Which of many patients with abdominal pain have self-limited gastroenteritis and which have early appendicitis? Which patients with sore throat and hoarseness have a "garden variety" pharyngitis and which the rare but potentially lethal *Haemophilus* epiglottitis? These are examples of how difficult, and important, distinguishing various kinds of abnormality can be.

The point of distinguishing normal from abnormal is to separate out those observations that should be considered for action from those that can be discounted. Observations considered normal are usually described as "within normal limits," "unremarkable," or "noncontributory" and remain buried in the body of a medical history. The abnormal findings are set out under a problem list, "impressions," or "diagnoses" and are the basis for action.