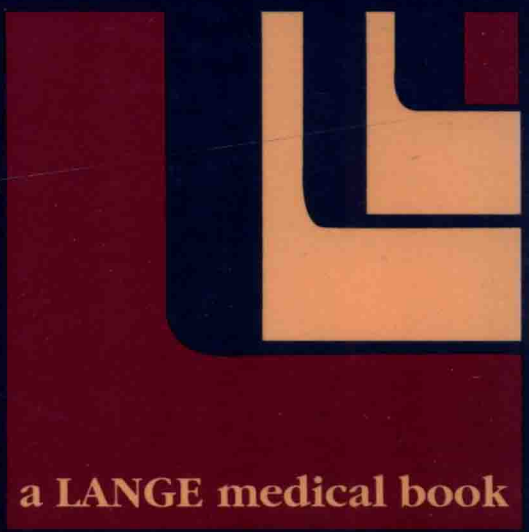
Examination & Board Review

Medical Biostatistics & Epidemiology

Diane Essex-Sorlie



LANGE medical book

Examination & Board Review

Medical Biostatistics & Epidemiology

First Edition

Diane Essex-Sorlie, PhD
Professor of Biometrics
University of Illinois College of Medicine
at Urbana-Champaign

APPLETON & LANGE
Norwalk, Connecticut

95 96 97 98 / 10 9 8 7 6 5 4 3 2 1

Prentice Hall International (UK) Limited, *London*
Prentice Hall of Australia Pty. Limited, *Sydney*
Prentice Hall Canada, Inc., *Toronto*
Prentice Hall Hispanoamericana, S.A., *Mexico*
Prentice Hall of India Private Limited, *New Delhi*
Prentice Hall of Japan, Inc., *Tokyo*
Simon & Schuster Asia Pte. Ltd., *Singapore*
Editora Prentice Hall do Brasil Ltda., *Rio de Janeiro*
Prentice Hall *Englewood Cliffs, New Jersey*

ISBN: 0-8385-6219-1
ISSN: 1079-4883

Acquisitions Editor: John Dolan
Production Editor: Chris Langan
Designer: Elizabeth Schmitz

ISBN 0-8385-6219-1

9 780838 562192 90000

PRINTED IN THE UNITED STATES OF AMERICA

This book is dedicated to individuals who have enriched my life immeasurably.
The late Elsie Bennett Essex, a strong courageous woman.
The late Anthony and Emma Marinangeli, devoted and loving, always.
Edward Minium, an exceptional teacher.
Marten Kernis, a colleague, friend, and mentor.
Daniel K. Bloomfield, a friend, colleague, and mentor.
Lillian and Harold Sorlie, kind, warm, and loving.
AND
William Sorlie, the most extraordinary man I have met.

Acknowledgments

I am indebted to the staff of Lange Medical Publications because they allowed me to turn a dream into reality. In particular, Stephen Foster, an Appleton & Lange sales representative, introduced me to John Dolan, senior medical editor at Appleton & Lange. I am most fortunate to have had John Dolan's advocacy, support, encouragement, and enthusiasm throughout this project. Gregory Huth and Chris Langan coordinated the production and helped turn my printed pages into a finished product. In addition, I am grateful to John Dinolfo, my developmental editor; words of thanks are inadequate to describe my gratitude for his excellent stylistic suggestions, encouragement, sense of humor, and commitment. If this book serves the audiences for whom it is intended, John Dinolfo deserves much of the credit.

I also wish to thank journal editors, such as *The New England Journal of Medicine,* and textbook publishers, such as *John Wiley, Inc.,* for allowing me to adapt and reproduce material. It is with the kind permission of the *Biometrika* trustees that I am able to adapt and reproduce statistical tables.

Dr. Edward Minium taught the first several statistics courses I took as an undergraduate and master's degree-level student. His enthusiasm for statistics and his dedication to teaching still serve today as my role model. He is a teacher extraordinaire. I am also indebted to my students who, over many years, encouraged me, challenged me to explain concepts clearly, and put up with my continuous experimentation with different teaching techniques. Further, several colleagues encouraged me throughout my writing. I thank Dr. Howard Bers and Dr. Marten Kernis for their constant support. Always in good cheer, Dr. Peter Imrey answered my questions, even when he struggled to cope with his own workload. Dr. H. Bruce Bosmann was generous in providing summer research awards that allowed me to concentrate on my writing. Jeffrey Loftiss in the Library of the Health Sciences helped me secure copies of many articles that were not available on the Urbana-Champaign campus of the University of Illinois. His good humor and efficiency are more than appreciated.

Writing this book has given me great personal and professional joy. I am privileged to share these joys with my husband, William Sorlie, who is my constant support. He quipped just recently, "I hope she finishes the book soon because I would like to see her again." He spent many nights and weekends alone while I wrote and wrote and wrote. . . . Thank you for your never-ending encouragement.

Finally, my beloved grandparents, Emma and Anthony Marinangeli died within two weeks of each other during the early stages of my writing. I owe them a debt I can never repay for their love and belief in me. Elsie Bennet Essex, my paternal grandmother, inspired me when she was alive; her memories continue to inspire me. Thank you for strength and courage.

Preface

Medical Biostatistics & Epidemiology: Examination & Board Review grew out of my intense love of teaching and my commitment to students, as well as to faculty who work diligently to teach biostatistics & epidemiology. It (1) covers many important concepts tested on the United States Licensing Examinations (USMLE) and provides a structured review for individuals preparing for these examinations; (2) presents a sound treatment of basic biostatistics and epidemiology that is highly accessible, even to readers who have not had a basic biostatistics course; (3) addresses the fundamental statistical and epidemiological concepts necessary to read and evaluate much of published medical literature; (4) serves as a resource for students, clinicians, and other health care professionals who would like to refresh their memory about basic statistical or epidemiological concepts or learn about a concept unfamiliar to them; and (5) can be used as a textbook for introductory medical statistics and epidemiology, with little or no supplementation by the instructor.

The key features of this book include the following:

- Each chapter begins with a clinical example that introduces statistical and epidemiological concepts.
- Objectives in each chapter help readers identify key concepts and reach key instructional goals.
- Clinical examples are used extensively to highlight the application and interpretation of statistical and epidemiological concepts.
- Many chapters conclude with a summary table that outlines a statistical test, including formulae, assumptions, relevant distributions and statistical tables.
- Flow charts are included in many chapters to help readers select the appropriate statistical test.
- Boxes summarize the analysis of clinical data and reinforce teaching concepts. These boxes provide a complete overview of an analysis, beginning with hypotheses and ending with interpreting results.
- Chapters include self-study questions with detailed solutions. These exercises are intended to help readers develop and sharpen diverse skills, including reading tables and graphs, interpreting results, formulating conclusions, selecting statistical tests, and conducting small-scale analyses and interpreting their results.
- The book concludes with a chapter on reading and evaluating the medical literature; an evaluation checklist is provided to help individuals use their reading time as effectively as possible.
- Finally, a 122-item comprehensive examination appears in the appendix. The answer key includes for each question: (a) a brief explanation of the correct answer, (b) the concept tested, and (c) reference to the chapter(s) in which the concept is presented. The examination format is similar to those used on course final examinations and licensure exams.

Diane Essex-Sorlie, Ph.D.

Urbana, Illinois
December 1994

Table of Contents

Defining Basic Concepts

1

Clinical Example: Primary Invasive Breast Cancer

> To determine whether a relationship exists between tumor angiogenesis and metastatic status (metastases present/metastases absent), oncologists studied tumor specimens from 90 randomly selected women with primary invasive breast cancer. Forty women had metastases and 50 women were without metastases ($n_{WITH} = 40$, $n_{WITHOUT} = 50$). Representative samples of invasive disease were selected from tumor specimen sections stained with hematoxylin and eosin. The number of microvessels per stained section was counted on a 200x field. In women with metastases, the mean number of microvessels per 200x field was 101, with a standard deviation of 49.3. In women without metastases, the mean number of microvessels was 45, with a standard deviation of 21.1. Based on these results, the clinicians concluded that tumor angiogenesis correlates directly with metastatic status in the patients evaluated.

Like all other disciplines, statistics has unique concepts and a specialized vocabulary that we must understand to comprehend the medical literature. Consider the research summarized above. In this hypothetical example based on very real clinical issues, can you distinguish the **independent variable** from the **dependent variable?** Are these variables **quantitative** or **qualitative?** To which patients can these results be extended (**generalized**)—that is, what is the **target population?** Which statistical analyses are appropriate, given the goals of the study and the way in which the variables are measured?

To answer these and other key questions, you must have a good understanding of the statistical and epidemiologic concepts discussed in this chapter. A solid grasp of the following concepts will help you evaluate published reports and determine whether—and to whom—findings may be extended.

OBJECTIVES

When you complete this chapter, you should be able to:

- Define key concepts in statistics and epidemiology.
- Read a brief clinical example and identify the type of study, independent variable, sample population, and other key features.
- Read a brief clinical example and determine if the conclusions extend to the individuals to whom they are being applied.

STATISTICS VERSUS EPIDEMIOLOGY

Statistics is "the science and art of dealing with variation in data through collection, classification, and analysis in such a way as to obtain reliable results" (Last, 1983, p 100). **Biostatistics** and **biometrics** refer to the application of statistics in medicine. For example, a biostatistician used statistical tools to analyze the data in the opening clinical example. The goal was to determine if tumor angiogenesis and metastatic status are related. **Epidemiology** is the study of the development, frequency, distribution, determinants, and consequences of disease in

human populations. Epidemiologic research typically seeks to (1) understand the causes of disease, (2) plan treatment, and (3) contribute to the development of public health policies. For example, an epidemiologist may study new cases of AIDS in a 12-month period and compare the findings to results from earlier studies, to detect changing disease patterns. An epidemiologist may also investigate an outbreak of bacterial meningitis in a university to identify disease determinants (eg, exposure and lifestyle patterns). Based on the findings, the university may decide to offer immunizations to all students.

DESCRIPTIVE VERSUS INFERENTIAL STATISTICS

The discipline of statistics is often divided into two branches: descriptive and inferential. The goal of **descriptive statistics** is to organize and summarize data. The goal of **inferential statistics** is to draw inferences and reach conclusions about data, when only a part of a population, or **sample**, has been studied. A **generalization** is a principle deduced from limited information (a sample) and extended to a larger collection of observations (the population). To **generalize** is to make deductions from limited information and extend those deductions to a larger collection of observations.

The **population** is the complete set of observations, patients, entities, measurements, and so forth, about which we would like to draw conclusions. The **sampled population** is the group from which the sample (or subset of the population) is drawn. The **target population** is the group to which we would like to extend our conclusions. Numerical characteristics of populations are **parameters.** Numerical characteristics of samples are **statistics.** For example, the mean microvessel count of the 40 women with primary invasive breast cancer with metastases is a statistic—in this case, 101. Lower-case "n" is frequently used to indicate sample size, that is, the number of entities observed—in this case, the number of patients. Upper-case N is generally used to designate population size.

Consider the research about breast cancer and tumor angiogenesis mentioned at the beginning of the chapter. Two populations were sampled. Forty randomly selected women with primary invasive breast cancer with metastases constitute the first population. Fifty randomly selected women with primary invasive breast cancer without metastases comprise the second population.

VARIABLES

Certain characteristics, attributes, and qualities are of interest to us; as we study them, we draw inferences and conclusions about them. Characteristics that can be measured are called **variables.** The term "variable" is used because its values may change, depending on related factors. Thus, we may measure an individual's blood pressure or heart rate on two separate occasions and find considerable variability, which can be due to many factors, such as the time of day or whether the person has just eaten a meal or smoked a cigarette.

Several types of variables exist in statistics. An **independent variable** is a characteristic or experimental intervention thought to influence a particular event or a manifestation of it. A **dependent variable** is the resulting outcome, event, or manifestation "whose variation we seek to explain or account for by the influences of independent variables" (Last, 1983, p 28). Thus, in our clinical example at the beginning of the chapter, metastatic status is the independent variable; microvessel counts per 200x field is the dependent variable. **Confounding variables** are factors that distort the degree to which the independent variable affects the dependent variable. The distortion may occur because the confounding factor is associated with the independent variable, dependent variable, or both (Masuner & Kramer, 1985, pp 159–160). If not taken into account by study design or statistical analysis, confounding variables may lead to erroneous conclusions.

The following example contains a confounding variable. In some men, workplace exposure to sulfur dioxide (SO_2) has been linked to the development of chronic cough. Cigarette

smoking is also associated with chronic cough in some men. If we examine the relationship between SO_2 exposure and chronic cough, we may develop incorrect conclusions, unless we include cigarette smoking history (a confounding variable) in our analysis.

MEASUREMENT OF VARIABLES

One of fours scales may be used to measure variables; these are the nominal, ordinal, interval, and ratio scales. A **nominal scale** uses names, labels, numbers, or other symbols to assign objects to a series of categories. A nominal scale contains a series of *unordered categories* or classes that are mutually exclusive and exhaustive. Labels, numbers, or other symbols are used to differentiate one class from another. Nominal variables are also called categorical or qualitative variables. Qualitative variables describe qualities or attributes that cannot be measured in the same sense as, for example, height, weight, cholesterol, triglyceride level, age, and blood pressure. Observations in a class are qualitatively the same; observations in different classes are qualitatively different. We determine nominal variables by identifying the categories or classes making up the scale and indicating the number or percentage of observations in each class or category. Examples of variables measured on a nominal scale include blood type, type of cancer, race, sex, and handedness. In our opening scenario, metastatic status (with/without) is a qualitative or nominal variable.

An **ordinal scale** contains distinct *ordered* or *ranked* qualitative categories. For this reason, ordinal scales are sometimes called **rank-order scales**. Labels, numbers, or other symbols are used to distinguish ordered categories. Observations that differ from category to category can be ranked according to whether the observation is more or less of some criterion. The categories are qualitative in the sense that the distance between and among them is not measurable numerically. However, the number or percentage of observations in each category can be described. Duke's classification of colorectal cancer (stages A through D) is an ordinal scale. Stage A represents cancer limited to the mucosa and submucosa. Stage B denotes cancer that extends into the muscularis or serosa. Stage C signifies cancer that involves the regional lymph nodes. Stage D represents cancer that has metastasized to the liver, bone, and/or lung. The categories are ordered because the extent of disease is more limited and prognosis more favorable in stage A than in stage D. In fact, approximately 90% of patients with cancer limited to the mucosa and submucosa (stage A) survive 5 years. In contrast, approximately 5% of patients whose cancer has metastasized to the liver, bone, and/or lung survive 5 years (stage D) (Glickman, 1987, p 1300).

Interval and ratio scales are used to describe quantitative variables. Similar statistical techniques can be applied to data from interval and ratio scales. A **quantitative variable** can be measured according to amount or quantity. Serum cholesterol, body weight, white blood cell count, body temperature, and patient age are quantitative variables. **Interval scales** involve assigning numbers at equal intervals from an arbitrary origin. Objects are ordered by the amount of the characteristic they possess. The arbitrarily selected origin (or zero point) does not imply a true absence of the measured characteristic. Fahrenheit and Celsius temperature scales are examples of an interval scale.

A **ratio scale** is an interval scale with a true zero point that reflects the absence of the measured characteristic. Blood pressure, body weight, time, age, Kelvin temperature, volume, and mass are ratio scales. Because the ratio scale has a true zero point, ratios between values are meaningful. For example, a patient who weighs 80 kg weighs twice as much as a patient who weighs 40 kg.

Interval and ratio scales, sometimes called **quantitative** or **numerical scales,** can be discrete or continuous. **Discrete scales** have integer values. Hospital census (number of beds occupied), number of new cases of AIDS in 12 months, number of patients seen in a family practice clinic in 12 months, and number of angioplasties performed in 6 months are examples of variables measured on a discrete scale. With **continuous scales**, values may have fractional components. Gestational age (eg, 17.6 weeks), body weight (eg, 50.5 kg), body temperature (eg, 97.4 °F), and survival postdiagnosis (eg, 6.3 months) are measured on a continuous scale. Depending on the precision required (as indicated by the number of places to the right of the decimal), continuous data may be reported as integer values. For example, many health

surveys report body weight to the nearest kilogram or patient age to the nearest year. Despite such simplification, these variables are still continuous because they can exhibit fractional components, at least in theory. Thus they can be measured on a continuum.

COMMON TYPES OF STUDIES

Investigators design studies in different ways in order to measure variables and use these measurements to learn and draw conclusions about phenomena, such as diseases and treatments. Because different types of studies are appropriate for studying different kinds of questions, it is important that you understand the key elements of common study designs. This understanding will help you read the medical literature, understand the type of conclusions that can be drawn from different types of studies, and evaluate the validity of conclusions.

Research can be divided into *observational studies* or *experiments,* based on whether patients are merely observed or whether some kind of intervention is performed (Dawson-Saunders & Trapp, 1994, p 6). Case reports and case series, case control, cohort (prospective and retrospective), and cross-sectional studies are observational, because one or more groups of patients are observed. Their characteristics are recorded for analysis—for example, exposure to pesticides and subsequent development of cancer, or augmentation mammoplasty with silicone gel breast implants and subsequent development of systemic sclerosis. Clinical trials are experiments because the investigator manipulates or controls an intervention, such as different drugs, drug doses, procedures, or treatments. For example, investigators may conduct a clinical trial to compare the use of zidovudine (AZT) alone to combination therapy with zidovudine, ddI, and pyridinone in patients with AIDS.

Case reports or **case series studies** are careful, detailed descriptions of interesting characteristics in a single patient (case report) or series of patients (case series). A case report and a case series study exclude patients without the characteristic or the disease under investigation. Case series studies may produce hypotheses that lead to more formalized research to identify causes of disease, diagnose disease, or treat disease.

In a **case-control study**, we use data from cases and controls to test theories derived from inferences about previously studied factors, past events, or experiences. *Cases* are persons who have the characteristic or disease of interest *at the outset of the investigation*; *control patients* are disease free. Case-control studies require knowledge of disease status at the beginning of the study. In this kind of study, we examine the experience and history of patients to identify factors present in the history of cases, but absent in the history of controls. Case-control studies are especially useful when investigating relatively rare diseases, such as certain types of brain tumors. Because disease status is known at the beginning of the investigation, researchers can select a sufficient number of diseased and nondiseased individuals in order to reach valid statistical conclusions about the illness. Because they "look back" from disease status to factors that may explain the occurrence of disease, case-control studies are sometimes called retrospective studies. However, an increasing number of investigators now reserve the term *retrospective* for a type of *cohort study* (Hennekens & Buring, 1987, p 23).

Cross-sectional studies provide a "snapshot" of what is happening at a particular point in time, rather than over a period of time. Cross-sectional studies are especially useful for evaluating a new diagnostic procedure and for estimating the frequency (prevalence) of a disease or a characteristic at a point in time. For example, a nationwide health survey may be conducted to describe the number of adults who attempt to control their fat intake, exercise strenuously 3 times per week, or use seatbelts regularly. An investigator might also use a cross-sectional study to evaluate the utility of a technetium-labeled granulocyte scan to diagnose active inflammatory bowel disease, and to compare the results to findings from conventional barium-contrast studies.

Cohort studies involve a carefully defined population that has been or may be exposed to a factor or factors thought to contribute to the occurrence of disease or other outcome. Subsets of the population, called cohorts, are followed over time to see if they develop disease. Thus, the starting point of all cohort studies is exposure status.

Cohort studies are either prospective or retrospective (Hennekens & Buring, 1987, p 23). In a **prospective cohort study,** exposed and unexposed individuals are followed at regular intervals to learn about the development and extent of disease. For example, clinicians might

conduct a prospective cohort study to compare the frequency of colds, bronchitis, and other respiratory problems in nonsmokers who are exposed to secondary cigarette smoke in the home, versus nonsmokers who are *not* exposed to secondary cigarette smoke in the home. In this type of study, the researchers would follow individuals for up to a year or more to monitor the incidence of respiratory infections.

In a **retrospective cohort study**, the investigation is initiated at a point in time after the exposure has occurred; however, the outcome of interest may or may not have occurred already by the time the study begins. Investigators begin with exposure status and use medical records, death certificates, and other available information to document disease development. For example, investigators may rely on patient charts to identify women who took diethylstilbestrol (DES) between 1947 and 1971 to treat threatened miscarriage. The objective is to use the medical records of offspring to determine whether fetal exposure to DES is associated with reproductive abnormalities, such as clear cell carcinoma of the vagina in fetal-exposed women and testicular abnormalities in fetal-exposed men. Cohort studies are an ideal way to investigate relatively common diseases or outcomes, such as how the development of breast cancer may differ in women with and without a family history of breast cancer.

Sometimes called intervention or experimental studies, **clinical trials** evaluate the effectiveness of a therapeutic procedure or agent (eg, a new drug). **Controlled clinical trials** compare a therapeutic agent or procedure with another agent or procedure. Patients in one group receive the new agent or procedure, while patients in the control (or reference) group receive a placebo or another drug or procedure. Investigations with one experimental treatment are **uncontrolled clinical trials;** because no control (or reference) group is included, the experimental treatment cannot be compared to another intervention. Without a control or reference group, it is often impossible to separate treatment effects from other factors, such as normal biologic variation within and between patients. Because controlled trials are more likely to permit researchers to decide whether differences are due to a treatment, clinical intervention, or some other factors, medical researchers and clinicians often consider controlled trials more useful than uncontrolled studies (Dawson-Saunders & Trapp, 1994, p 13).

UNDERSTANDING ASPECTS OF STUDY DESIGN

Certain terms are used frequently in the medical literature to describe features of the design of a study. An understanding of these terms will help you evaluate a study and assess the degree to which conclusions can be extended from the patients studied to other individuals.

A **placebo** is an inert medication or procedure intended to be indistinguishable from the experimental or active treatment. Patients are not told whether they receive an active medication or a placebo. The placebo sometimes has a therapeutic effect, which is due to the patient's expectation that he or she will feel better or that the disease will improve after taking medication. Improvement, then, is attributed to the patient's expectation that a treatment or drug is beneficial. This improvement is known as a **placebo** or **halo effect**.

To **replicate** is to administer the same drug or procedure to two or more patients under identical conditions. **Replication** is the act of repeating an investigation to confirm findings or to improve the accuracy of measurements.

To **control** is to minimize extraneous sources of variation, either by study design or statistical analysis. Control in a study design can be accomplished by:

Including a control (or reference) group.
Randomly assigning patients to treatment groups or experimental conditions.
Restricting patients who enroll in the study to reduce variation between patients.
Matching patients (described below).

Statistical control can be achieved through the selection of an appropriate analysis, by mathematical modeling, or by statistical adjustment.

Random assignment of patients to treatment groups or conditions minimizes extraneous variation by helping to ensure that all patients who enroll in the study have an equal chance of receiving the treatment (Fletcher et al, 1988, p 122).

In **matching,** often referred to as **pair matching,** the investigator matches patients as they

enter the study, based on characteristics thought to be associated with the outcome of interest. Pair matching is done so that the patients in each pair are as alike as possible with respect to certain factors, such as disease severity, age, smoking history, family history, or sex. One member of the pair is assigned randomly to one group; the pairmate is then placed automatically in the other group. Investigators often match patients according to age or disease severity. As noted earlier, matching is one way to control extraneous variation and increase the likelihood that differences between groups occur because of the treatment or intervention, rather than being due initial differences between the groups.

Bias is any error or effect that causes the results of a study to depart from true values (Last, 1983, p 10). Bias can occur because one or more of the following is inadequate: the measuring instrument, selection of patients, assignment of patients to treatments, or method of measuring outcomes of interest. Suppose researchers want to study the relationship between estrogen replacement therapy in women and subsequent development of breast cancer. For the estrogen replacement group, patients with a family history of breast cancer are selected; for the control group (who receive a placebo), patients without a family history of breast cancer are selected. This example illustrates a *confounding bias* in selecting patients, because the outcome of interest (breast cancer development) is confounded by family history. In this case, the development of breast cancer is more likely to occur in women with a family history of breast cancer than in women without a family history. Keep in mind that bias due to patient assignment can occur when patients with substantial disease are selected for the treatment group and patients whose disease is "mild" are assigned to the placebo group, or vice versa.

Confounding occurs when the effects of one or more variables cannot be separated from each other and from their effect on the dependent variable (Last, 1983, p 21). In the hypothetical research on breast cancer and estrogen use, the effects of family history and estrogen use are not separated to determine their respective impact on the development of breast cancer in the sampled population. Thus, the effect of treatment is confounded by family history. An excellent discussion of confounding and other flaws in the medical literature is presented by Michael and associates (1984).

Validity is the extent to which a device (eg, a test) measures what it purports to measure, or a clinical observation accurately describes a phenomenon. Validity is discussed most frequently in terms of internal and external validity. A study has **internal validity** if the differences between treatment groups can be attributed to the independent variable(s). Sampling procedures, definition of outcomes, measurement of outcomes, statistical analyses, and administration of treatments must all be appropriate for a study to be internally valid. Internal validity is required to extend a finding beyond the patients studied (Fletcher et al, 1988, p 12).

External validity refers to the degree to which research findings can be applied or *generalized* to individuals other than those who were studied. External validity applies only to a specified target population (Last, 1983, p 108). Suppose that after examining the effect of AZT to prolong the life span of 25 AIDS patients with full-blown disease, we try to generalize our findings to asymptomatic, HIV-positive individuals. Our conclusions could be misleading because they may lack sufficient external validity. However, if we generalize our conclusions to AIDS patients whose disease stage and other characteristics closely resemble those of patients in our study group, our findings could have high external validity.

Reliability is another important consideration when evaluating the results of a study. Reliability refers to the degree to which a measuring device or a procedure produces repeatable results on each subsequent use or occasion. **Reproducibility** and **repeatability** are synonyms for reliability. Results may be reliable but not valid, as when a blood pressure cuff produces a diastolic reading that is relatively repeatable for a patient but is substantially higher (or lower) than the patient's true pressure.

SUMMARY

To read and evaluate medical literature, including published studies as well as information distributed by pharmaceutical companies, you must have a good understanding of concepts and vocabulary unique to the discipline of statistics. In addition to discussing the goals of statistics, we define and explain a wide range of statistical concepts in this chapter, including