Susan White

BASIC & CLINICAL BIOSTATISTICS

5th Edition

Mc Graw Hill LANGE®

This page intentionally left blank

a LANGE medical book

Basic & Clinical Biostatistics

Fifth edition

Susan E. White, PhD
Associate Professor Clinical
School of Health and Rehabilitation Sciences
The Ohio State University
Administrator of Analytics
The James Cancer Hospital at
The Ohio State University Wexner Medical Center

New York Chicago San Francisco Athens London Madrid
Mexico City Milan New Delhi
Singapore Sydney Toronto

ISBN: 978-1-26-045537-3
MHID: 1-26-045537-8

The material in this eBook also appears in the print version of this title: ISBN: 978-1-26-045536-6,
MHID: 1-26-045536-X.

eBook conversion by codeMantra
Version 1.0

Notice

Medicine is an ever-changing science. As new research and clinical experience broaden our knowledge, changes in treatment and drug therapy are required. The authors and the publisher of this work have checked with sources believed to be reliable in their efforts to provide information that is complete and generally in accord with the standards accepted at the time of publication. However, in view of the possibility of human error or changes in medical sciences, neither the authors nor the publisher nor any other party who has been involved in the preparation or publication of this work warrants that the information contained herein is in every respect accurate or complete, and they disclaim all responsibility for any errors or omissions or for the results obtained from use of the information contained in this work. Readers are encouraged to confirm the information contained herein with other sources. For example and in particular, readers are advised to check the product information sheet included in the package of each drug they plan to administer to be certain that the information contained in this work is accurate and that changes have not been made in the recommended dose or in the contraindications for administration. This recommendation is of particular importance in connection with new or infrequently used drugs.

Contents

Preface

Basic & Clinical Biostatistics introduces the medical student, researcher, or practitioner to the study of statistics applied to medicine and other disciplines in the health field. The book covers the *basics* of biostatistics and quantitative methods in epidemiology and the *clinical* applications in evidence-based medicine and the decision-making methods. Particular emphasis is on study design and interpretation of results of research.

OBJECTIVE

The primary objective of this text is to provide the resources to help the reader become an informed user and consumer of statistics. This book should allow you to:

- Develop sound judgment about data applicable to clinical care.
- Read the clinical literature critically, understanding potential errors and fallacies contained therein, and apply confidently the results of medical studies to patient care.
- Interpret commonly used vital statistics and understand the ramifications of epidemiologic information for patient care and prevention of disease.
- Reach correct conclusions about diagnostic procedures and laboratory test results.
- Interpret manufacturers' information about drugs, instruments, and equipment.
- Evaluate study protocols and articles submitted for publication and actively participate in clinical research.
- Develop familiarity with well-known statistical software and interpret the computer output.

APPROACH & DISTINGUISHING FEATURES

The practitioner's interests, needs, and perspectives in mind during the preparation of this text. Thus, our approach embraces the following features:

- A genuine medical context is offered for the subject matter. After the introduction to different kinds of studies is presented in Chapter 2, subsequent chapters begin with several *Presenting Problems*—discussions of studies that have been published in the medical literature. These illustrate the methods discussed in the chapter and in some instances are continued through several chapters and in the exercises to develop a particular line of reasoning more fully.
- All example articles and datasets are available via open source access.
- Actual data from the *Presenting Problems* are used to illustrate the statistical methods.
- A focus on concepts is accomplished by using computer programs to analyze data and by presenting statistical calculations only to illustrate the logic behind certain statistical methods.
- The importance of sample size (power analysis) is emphasized, and computer programs to estimate sample size are illustrated.
- Information is organized from the perspective of the research question being asked.
- Terms are defined within the relevant text, whenever practical, because biostatistics may be a new language to you. In addition, a glossary of statistical and epidemiologic terms is provided at the end of the book.
- A table of all symbols used in the book is provided on the inside back cover.
- A simple classification scheme of study designs used in clinical research is discussed (Chapter 2). We employ this scheme throughout the book as we discuss the *Presenting Problems.*
- Flowcharts are used to relate research questions to appropriate statistical methods (inside front cover and Appendix C).
- A step-by-step explanation of how to read the medical literature critically (Chapter 13)—a necessity for the modern health professional—is provided.

- Evidence-based medicine and decision-making are addressed in a clinical context (Chapters 3 and 12). Clinicians will be called on increasingly to make decisions based on statistical information.
- Numerous end-of-chapter *Exercises* (Chapters 2 through 12) and their complete solutions (Appendix B) are provided.
- A posttest of multiple-choice questions (Chapter 13) similar to those used in course final examinations or licensure examinations is included.

SPECIAL FEATURES IN THIS EDITION

There are several important enhancements included in the fifth edition.

To facilitate and increase learning, each chapter (except Chapter 1) contains a set of *Key Concepts* to orient the reader to the important ideas discussed in the chapter.

- Many of the *Presenting Problems* have been updated with journal references that require the authors to provide access to the journal article and data through a creative commons license. The links to articles and datasets used for examples are detailed in the *Presenting Problem* summary at the beginning of each chapter.
- Material addressing best practices in data visualization is included in Chapter 3.
- All sample size calculations are now presented using G*Power, an open source program used widely for sample size calculation by researchers.
- Inclusion of output and exercise answers using R and R Commander—open source statistical applications that may be used across many computer operating systems (Windows, Mac, and Unix).

Susan E. White, PhD

Using R

R is a statistical computing package that is available via an open source license. R (R Core Team, 2019) may be downloaded from http://www.R-project.org.

The add-on R Commander provides new users with a graphical interface that makes using R far more intuitive.

R Commander (Fox and Bouchet-Valat 2019) may be downloaded from this site:
https://www.rcommander.com/
or
https://socialsciences.mcmaster.ca/jfox/Misc/Rcmdr/

There are also R Commander plug-ins that are used in the examples:
RcmdrPlugin.survival
RcmdrPlugin.aRnova

There are a number of excellent resources online to help you learn to use R and R Commander. Here is a short list:

R Commander an introduction: https://cran.r-project.org/doc/contrib/Karp-Rcommander-intro.pdf

Getting Started with R Commander: https://cran.r-project.org/web/packages/Rcmdr/vignettes/Getting-Started-with-the-Rcmdr.pdf

This page intentionally left blank

Introduction to Medical Research

1

The goal of this text is to provide you with the tools and skills you need to be a smart user and consumer of medical statistics. This goal has guided the selection of material and in the presentation of information. This chapter outlines the reasons physicians, medical students, and others in the health care field should know biostatistics. It also describes how the book is organized, what you can expect to find in each chapter, and how you can use it most profitably.

THE SCOPE OF BIOSTATISTICS & EPIDEMIOLOGY

The word "statistics" has several meanings: data or numbers, the process of analyzing the data, and the description of a field of study. It derives from the Latin word *status,* meaning "manner of standing" or "position." Statistics were first used by tax assessors to collect information for determining assets and assessing taxes—an unfortunate beginning and one the profession has not entirely lived down.

Everyone is familiar with the statistics used in baseball and other sports, such as a baseball player's batting average, a bowler's game point average, and a basketball player's free-throw percentage. In medicine, some of the statistics most often encountered are called means, standard deviations, proportions, and rates. Working with statistics involves using statistical methods that summarize data (to obtain, for example, means and standard deviations) and using statistical procedures to reach certain conclusions that can be applied to patient care or public health planning. The subject area of statistics is the set of all the statistical methods and procedures used by those who work with statistics. The application of statistics is broad indeed and includes business, marketing, economics, agriculture, education, psychology, sociology, anthropology, and biology, in addition to our special interest, medicine and other health care disciplines. The terms **biostatistics** and **biometrics** refer to the application of statistics in the health-related fields.

Although the focus of this text is biostatistics, some topics related to epidemiology are included as well. The term "epidemiology" refers to the study of health and illness in human populations, or, more precisely, to the patterns of health or disease and the factors that influence these patterns; it is based on the Greek words for "upon" (*epi*) and "people" (*demos*). Once knowledge of the epidemiology of a disease is available, it is used to understand the cause of the disease, determine public health policy, and plan treatment. The application of population-based information to decision-making about individual patients is often referred to as **clinical epidemiology** and, more recently, **evidence-based medicine.** The tools and methods of biostatistics are an integral part of these disciplines.

BIOSTATISTICS IN MEDICINE

Clinicians must evaluate and use new information throughout their lives. The skills you learn in this text will assist in this process because they concern modern knowledge acquisition methods. The most important reasons for learning biostatistics are listed in the following subsections. (The most widely applicable reasons are mentioned first.)

Evaluating the Literature

Reading the literature begins early in the training of health professionals and continues throughout their careers. They must understand biostatistics to decide whether they can rely on the results presented in the literature. Journal editors try to screen out articles that are improperly designed or analyzed, but few have formal statistical training and they naturally focus on the content of the research rather than the method. Investigators for large, complex studies almost always consult statisticians for assistance in project design and data analysis, especially research funded by the National Institutes of Health and other national agencies and foundations. Even then it is important to be aware of possible shortcomings in the way a study is designed and carried out. In smaller research projects, investigators consult with statisticians less frequently, either

because they are unaware of the need for statistical assistance or because the biostatistical resources are not readily available or affordable. The availability of easy-to-use computer programs to perform statistical analysis has been important in promoting the use of more complex methods. This same accessibility, however, enables people without the training or expertise in statistical methodology to report complicated analyses when they are not always appropriate.

The problems with studies in the medical literature have been amply documented. Sander Greenland's (2011) article on the misinterpretation in statistical testing in health risk assessment outlines errors in the reporting and interpretation of statistics in medical literature. The article includes a number of examples of erroneous conclusions surrounding the reporting of odds ratios and conclusions based on inadequate sample sizes. Much of the misinterpretation around the results of medical studies are in the reporting of statistical conclusions based on interferential methods such as hypothesis tests and *p*-values. Greenland's later work (2016) lists 25 misinterpretations of *p*-values, confidence intervals, and power commonly found in scientific literature.

The issue with misuse of *p*-values is so rampant that the American Statistical Association published a statement to guide the proper interpretation of *p*-values (Wasserstein and Lazar, 2016). The article outlines six principles that address the most common misconceptions around *p*-values:

1. *P*-values can indicate how incompatible the data are with a specified statistical model.
2. *P*-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
3. Scientific conclusions and business or policy decisions should not be based only on whether a *p*-value passes a specific threshold.
4. Proper inference requires full reporting and transparency.
5. A *p*-value, or statistical significance, does not measure the size of an effect or the importance of a result.
6. By itself, a *p*-value does not provide a good measure of evidence regarding a model or hypothesis.

Journals have also published a number of articles that suggest how practitioners could better report their research findings. Although these recommendations may result in improvements in the reporting of statistical results, the reader must assume the responsibility for determining whether the results of a published study are valid. The development of this book has been guided by the study designs and statistical methods found primarily in the medical literature, and topics were selected to provide the skills needed to determine whether a study is valid and should be believed. Chapter 13 focuses specifically on how to read the medical literature and provides checklists for flaws in studies and problems in analysis.

Applying Study Results to Patient Care

Applying the results of research to patient care is the major reason practicing clinicians read the medical literature. They want to know which diagnostic procedures are best, which methods of treatment are optimal, and how the treatment regimen should be designed and implemented. Of course, they also read journals to stay aware and up to date in medicine in general as well as in their specific area of interest. Chapters 3 and 12 discuss the application of techniques of evidence-based medicine to decisions about the care of individual patients.

Interpreting Vital Statistics: Physicians must be able to interpret vital statistics in order to diagnose and treat patients effectively. Vital statistics are based on data collected from the ongoing recording of vital events, such as births and deaths. A basic understanding of how vital statistics are determined, what they mean, and how they are used facilitates their use. Chapter 3 provides information on these statistics.

Understanding Epidemiologic Problems: Practitioners must understand epidemiologic problems because this information helps them make diagnoses and develop management plans for patients. Epidemiologic data reveal the prevalence of a disease, its variation by season of the year and by geographic location, and its relation to certain risk factors. In addition, epidemiology helps us understand how newly identified viruses and other infectious agents spread. This information helps society make informed decisions about the deployment of health resources, for example, whether a community should begin a surveillance program, whether a screening program is warranted and can be designed to be efficient and effective, and whether community resources should be used for specific health problems. Describing and using data in decision-making are highlighted in Chapters 3 and 12.

Interpreting Information about Drugs and Equipment: Physicians continually evaluate information about drugs and medical instruments and equipment. This material may be provided by company representatives, sent through the mail, or published in journals. Because of the high cost of developing drugs and medical instruments, companies do all they can to recoup their investments. To sell their products, a company must convince physicians that its products are

better than those of its competitors. To make its point, a company uses graphs, charts, and the results of studies comparing its products with others on the market. Every chapter in this text is related to the skills needed to evaluate these materials, but Chapters 2, 3, and 13 are especially relevant.

Using Diagnostic Procedures: Identifying the correct diagnostic procedure to use is a necessity in making decisions about patient care. In addition to knowing the prevalence of a given disease, physicians must be aware of the sensitivity of a diagnostic test in detecting the disease when it is present and the frequency with which the test correctly indicates no disease in a well person. These characteristics are called the sensitivity and specificity of a diagnostic test. Information in Chapters 4 and 12 relates particularly to skills for interpreting diagnostic tests.

Being Informed: Keeping abreast of current trends and being critical about data are more general skills and ones that are difficult to measure. These skills are also not easy for anyone to acquire because many responsibilities compete for a professional's time. One of the by-products of working through this text is a heightened awareness of the many threats to the validity of information, that is, the importance of being alert for statements that do not seem quite right.

Appraising Guidelines: The number of guidelines for diagnosis and treatment has increased greatly in recent years. Practitioners caution that guidelines should not be accepted uncritically; although some are based on medical evidence, many represent the collective opinion of experts. A review of clinical practices guidelines between 1980 and 2007 by Alonso-Coello and colleagues (2010) found that the quality scores of the guidelines as measured by the AGREE Instrument improved somewhat over time, but remained in the moderate to low range.

Evaluating Study Protocols and Articles: Physicians and others in the health field who are associated with universities, medical schools, or major clinics are often called on to evaluate material submitted for publication in medical journals and to decide whether it should be published. Health practitioners, of course, have the expertise to evaluate the content of a protocol or article, but they often feel uncomfortable about critiquing the design and statistical methods of a study. No study, however important, will provide valid information about the practice of medicine and future research unless it is properly designed and analyzed. Careful attention to the concepts covered in this text will provide physicians with many of the skills necessary for evaluating the design of studies.

Participating in or Directing Research Projects: Clinicians participating in research will find knowledge about biostatistics and research methods indispensable. Residents in all specialties as well as other health care trainees are expected to show evidence of scholarly activity, and this often takes the form of a research project. The comprehensive coverage of topics in this text should provide most of them with the information they need to be active participants in all aspects of research.

THE DESIGN OF THIS BOOK

This text is both *basic* and *clinical* because both the basic concepts of biostatistics and the use of these concepts in clinical decision-making are emphasized. This comprehensive text covers the traditional topics in biostatistics plus the quantitative methods of epidemiology used in research. For example, commonly used ways to analyze survival data are included in Chapter 9; illustrations of computer analyses in chapters in which they are appropriate, because researchers today use computers to calculate statistics; and applications of the results of studies to the diagnosis of specific diseases and the care of individual patients, sometimes referred to as medical decision-making or evidence-based medicine.

The presentations of techniques and examples are illustrated using the statistical program R (R Core Team, 2018). R is a cross platform software program that is freely distributed on the terms of a GNU General Public License. Since the software is cross platform, the examples presented in the text may be replicated using computers that run Windows, macOS, or UNIX.

This text deemphasizes calculations and uses computer programs to illustrate the results of statistical tests. In most chapters, the calculations of some statistical procedures are included, primarily to illustrate the logic behind the tests, not because you will need to be able to perform the calculations yourself. Some exercises involve calculations because some students wish to work through a few problems in detail so as to understand the procedures better. The major focus of the text, however, is on the interpretation and use of research methods.

A word regarding the accuracy of the calculations is in order. Many examples and exercises require several steps. The accuracy of the final answer depends on the number of significant decimal places to which figures are extended at each step of the calculation. Calculators and computers, however, use a greater number of significant decimal places at each step and often yield an answer different from that obtained using only two or three significant digits. The difference will usually be small, but do not be concerned if your calculations vary slightly from the examples.

The examples used are taken from studies published in the medical literature. Occasionally, a subset of the data is used to illustrate a more complex procedure. In addition, the focus of an example may be on only one aspect of the data analyzed in a published study in order to illustrate a concept or statistical test. To explain certain concepts, tables and graphs are reproduced as they appear in a published study. These reproductions may contain symbols that are not discussed until a later chapter in this book. Simply ignore such symbols for the time being. The focus on published studies is based on two reasons: First, they convince readers of the relevance of statistical methods in medical research; and second, they provide an opportunity to learn about some interesting studies along with the statistics.

The presentation of techniques in this text often refer to both previous and upcoming chapters to help tie concepts together and point out connections. This technique requires to use definitions somewhat differently from many other statistical texts; that is, terms are often used within the context of a discussion without a precise definition. The definition is given later. Several examples appear in the foregoing discussions (e.g., vital statistics, means, standard deviations, proportions, rates, validity). Using terms properly within several contexts helps the reader learn complex ideas, and many ideas in statistics become clearer when viewed from different perspectives. Some terms are defined along the way, but providing definitions for every term would inhibit our ability to point out the connections between the ideas. To assist the reader, boldface type is used for terms (the first few times they are used) that appear in the Glossary of statistical and epidemiologic terms provided at the end of the book.

THE ORGANIZATION OF THIS BOOK

Each chapter begins with two components: **key concepts** and an introduction to the examples (presenting problems) covered in the chapter. The key concepts are intended to help readers organize and visualize the ideas to be discussed and then to identify the point at which each is discussed. At the conclusion of each chapter is a summary that integrates the statistical concepts with the presenting problems used to illustrate them. When flowcharts or diagrams are useful, they are included to help explain how different procedures are related and when they are relevant. The flowcharts are grouped in Appendix C for easy reference.

Patients come to their health care providers with various health problems. In describing their patients, these providers commonly say, "The patient presents with …" or "The patient's presenting problem is …" This terminology is used in this text to emphasize the similarity between medical practice and the research problems discussed in the medical literature. Almost all chapters begin with presenting problems that discuss studies taken directly from the medical literature; these research problems are used to illustrate the concepts and methods presented in the chapter. In chapters in which statistics are calculated (e.g., the mean in Chapter 3) or statistical procedures are explained (e.g., the t test in Chapters 5 and 6), data from the presenting problems are used in the calculations. The selection of presenting problems is intended to represent a broad array of interests, while being sure that the studies use the methods discussed.

Exercises are provided with all chapters (2–13); answers are given in Appendix B, most with complete solutions. A variety of exercises are included to meet the different needs of students. Some exercises call for calculating a statistic or a statistical test. Some focus on the presenting problems or other published studies and ask about the design (as in Chapter 2) or about the use of elements such as charts, graphs, tables, and statistical methods. Occasionally, exercises extend a concept discussed in the chapter. This additional development is not critical for all readers to understand, but it provides further insights for those who are interested. Some exercises refer to topics discussed in previous chapters to provide reminders and reinforcements.

The **symbols** used in statistics are sometimes a source of confusion. These symbols are listed on the inside back cover for ready access. When more than one symbol for the same item is encountered in the medical literature, the most common one is used and points out the others. Also, a Glossary of biostatistics and epidemiologic terms is provided at the end of the book (after Chapter 13).

ADDITIONAL RESOURCES

References are provided to other texts and journal articles for readers who want to learn more about a topic. With the growth of the Internet, many resources have become easily available for little or no cost. A number of statistical programs and resources are available on the Internet. Some of the programs are freeware, meaning that anyone may use them free of charge; others, called shareware, charge a relatively small fee for their use. Many of the software vendors have free products or software you can download and use for a restricted period of time.

The American Statistical Association (ASA) has a number of sections with a special emphasis, such as Teaching Statistics in the Health Sciences, Biometrics

Section, Statistical Education, and others. Many of these Section homepages contain links to statistical resources. The ASA homepage is http://www.amstat.org.

Dartmouth University has links to the impressive Chance Database http://www.dartmouth.edu/%7Echance/index.html, which contains many teaching resources and, in turn, many useful links to other resources.

The Medical University of South Carolina has links to a large number of evidence-based-medicine sites, including its own resources https://musc.libguides.com/ebp.

Study Designs in Medical Research 2

KEY CONCEPTS

1. *Study designs in medicine fall into two categories: studies in which subjects are observed, and studies in which the effect of an intervention is observed.*
2. *Observational studies may be forward-looking (cohort), backward-looking (case–control), or looking at simultaneous events (cross-sectional). Cohort studies generally provide stronger evidence than the other two designs.*
3. *Studies that examine patient outcomes are increasingly published in the literature; they focus on specific topics, such as resource utilization, functional status, quality of life, patient satisfaction, and cost-effectiveness.*
4. *Studies with interventions are called experiments or clinical trials. They provide stronger evidence than observational studies.*
5. *The single best way to minimize bias is to randomly select subjects in observational studies or randomly assign subjects to different treatment arms in clinical trials.*
6. *Bias occurs when the way a study is designed or carried out causes an error in the results and conclusions. Bias can be due to the manner in which subjects are selected or data are collected and analyzed.*
7. *Clinical trials without controls (subjects who do not receive the intervention) are difficult to interpret and do not provide strong evidence.*
8. *Each study design has specific advantages and disadvantages.*

This chapter introduces the different kinds of studies commonly used in medical research. Knowing how a study is designed is important for understanding the conclusions that can be drawn from it. Therefore, considerable attention will be devoted to the topic of study designs.

If you are familiar with the medical literature, you will recognize many of the terms used to describe different study designs. If you are just beginning to read the literature, you should not be dismayed by all the new terminology; there will be ample opportunity to review and become familiar with it. Also, the glossary at the end of the book defines the terms used here. In the final chapter of this book, study designs are reviewed within the context of reading journal articles, and pointers are given on how to look for possible biases that can occur in medical studies. Bias can be due to the manner in which patients are selected, data are collected and analyzed, or conclusions are drawn.

CLASSIFICATION OF STUDY DESIGNS

There are several different schemes for classifying study designs. The one most relevant in clinical applications divides studies into those in which the subjects were merely observed, sometimes called **observational studies,** and those in which some intervention was performed, generally called **experiments.** This approach is simple and reflects the sequence an investigation

Table 2–1. Classification of study designs.

I. Observational studies
A. Descriptive or case–series
B. Case–control studies (retrospective)
1. Causes and incidence of disease
2. Identification of risk factors
C. Cross-sectional studies, surveys (prevalence)
1. Disease description
2. Diagnosis and staging
3. Disease processes, mechanisms
D. Cohort studies (prospective)
1. Causes and incidence of disease
2. Natural history, prognosis
3. Identification of risk factors
E. Historical cohort studies
II. Experimental studies
A. Controlled trials
1. Parallel or concurrent controls
a. Randomized
b. Not randomized
2. Sequential controls
a. Self-controlled
b. Crossover
3. External controls (including historical)
B. Studies with no controls
III. Meta-analyses

sometimes takes. With a little practice, you should be able to read medical articles and classify studies according to the outline in Table 2–1 with little difficulty.

Each study design in Table 2–1 is illustrated in this chapter, using some of the studies that are presenting problems in upcoming chapters. In observational studies, one or more groups of patients are observed, and characteristics about the patients are recorded for analysis. Experimental studies involve an **intervention**—an investigator-controlled maneuver, such as a drug, a procedure, or a treatment—and interest lies in the effect the intervention has on study subjects. Of course, both observational and experimental studies may involve animals or objects, but most studies in medicine involve people.

OBSERVATIONAL STUDIES

Observational studies are of four main types: case–series, case–control, cross-sectional (including surveys), and cohort studies. When certain characteristics of a group (or series) of patients (or cases) are described in a published report, the result is called a **case–series study;** it is the simplest design in which the author describes some interesting or intriguing observations that occurred for a small number of patients.

Case–series studies frequently lead to the generation of hypotheses that are subsequently investigated in a case–control, cross-sectional, or cohort study. These three types of studies are defined by the period of time the study covers and by the direction or focus of the research question. Cohort and case–control studies generally involve an extended period of time defined by the point when the study begins and the point when it ends; some process occurs, and a certain amount of time is required to assess it. For this reason, both cohort and case–control studies are sometimes also called **longitudinal studies.** The major difference between them is the direction of the inquiry or the focus of the research question: Cohort studies are forward-looking, from a risk factor to an outcome, whereas case–control studies are backward-looking, from an outcome to risk factors. The cross-sectional study analyzes data collected on a group of subjects at one time. If you would like a more detailed discussion of study designs used in medicine, a book by Hulley et al (2013) is devoted entirely to the design of clinical research. Garb (1996) and Burns and Grove (2014) discuss study design in medicine and nursing, respectively.

Case–Series Studies

A case–series report is a simple descriptive account of interesting characteristics observed in a group of patients. For example, Glazer et al (2016) presented information on a series of 21 patients with acinar cell carcinoma of the pancreas. The authors wanted to compare two treatments, a combination of surgery and adjuvant chemotherapy versus surgery only, to see which resulted in longer survival in both metastatic and nonmetastatic cancers. They concluded that a multidisciplinary approach to treat the disease may result in longer survival.

Case–series reports generally involve patients seen over a relatively short time. Generally, case–series studies do not include **control subjects,** persons who do not have the disease or condition being described. Some investigators would not include case–series in a list of types of studies because they are generally not planned studies and do not involve any research hypotheses. On occasion, however, investigators do include control subjects. We mention case–series studies because of their important descriptive role as a precursor to other studies.

Case–Control Studies

Case–control studies begin with the absence or presence of an outcome and then look backward in time to try to detect possible causes or risk factors that may have

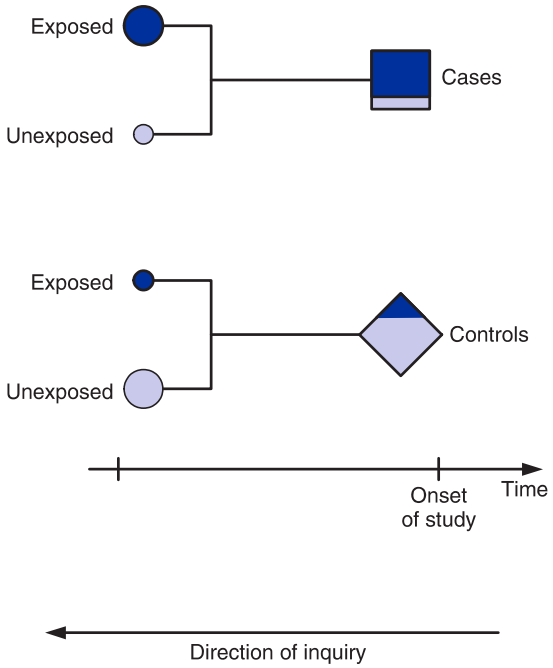

Figure 2–1. Schematic diagram of case–control study design. Shaded areas represent subjects exposed to the antecedent factor; unshaded areas correspond to unexposed subjects. Squares represent subjects with the outcome of interest; diamonds represent subjects without the outcome of interest. (Adapted with permission from Ilango K, Vijayakumar TM, Dubey GP, et al: An Enlarged Vision on Various Types of Study Design in Human Subjects, Global J Pharm 2012 Jan;6(3):216-221.)

been suggested in a case–series report. The *cases* in case–control studies are individuals selected on the basis of some disease or outcome; the *controls* are individuals without the disease or outcome. The history or previous events of both cases and controls are analyzed in an attempt to identify a characteristic or risk factor present in the cases' histories but not in the controls' histories.

Figure 2–1 illustrates that subjects in the study are chosen at the onset of the study after they are known to be either cases with the disease or outcome (squares) or controls without the disease or outcome (diamonds). The histories of cases and controls are examined over a previous period to detect the presence (shaded areas) or absence (unshaded areas) of predisposing characteristics or risk factors, or, if the disease is infectious, whether the subject has been exposed to the presumed infectious agent. In case–control designs, the nature of the inquiry is backward in time, as indicated by the arrows pointing backward in Figure 2–1 to illustrate the backward, or retrospective, nature of the research process. We can characterize case–control studies as studies that ask "What happened?" In fact, they are sometimes called **retrospective studies** because of the direction of inquiry. Case–control studies are longitudinal as well, because the inquiry covers a period of time.

Cai and colleagues (2014) compared patients who had a surgical site infection (SSI) following total joint arthroplasty (cases) with patients who developed no infection (controls). The investigators found that Aquacel dressing use was associated with a lower rate of infection. The study found a number of variables that increased the odds of an SSI, including: age, body mass index, smoking history, thyroid and/or liver disease, and a history of steroid treatment.

Investigators sometimes use **matching** to associate controls with cases on characteristics such as age and sex. If an investigator feels that such characteristics are so important that an imbalance between the two groups of patients would affect any conclusions, they should employ matching. This process ensures that both groups will be similar with respect to important characteristics that may otherwise cloud or confound the conclusions.

Deciding whether a published study is a case–control study or a case–series report is not always easy. Confusion arises because both types of studies are generally conceived and written after the fact rather than having been planned. The easiest way to differentiate between them is to ask whether the author's purpose was to describe a phenomenon or to attempt to explain it by evaluating previous events. If the purpose is simple description, chances are the study is a case–series report.

Cross-Sectional Studies

The third type of observational study goes by all of the following names: cross-sectional studies, surveys, epidemiologic studies, and prevalence studies. We use the term "cross-sectional" because it is descriptive of the timeline and does not have the connotation that the terms "surveys" and "prevalence" do. **Cross-sectional studies** analyze data collected on a group of subjects at one time rather than over a period of time. Cross-sectional studies are designed to determine "What is happening?" right now. Subjects are selected and information is obtained in a short period of time (Figure 2–2; note the short timeline). Because they focus on a point in time, they are sometimes also called **prevalence** studies. Surveys and polls are generally cross-sectional studies, although surveys can be part of a cohort or case–control study if the survey data is collected from a subset of the subjects. Cross-sectional studies may be designed to address research questions raised by a case–series, or they may be done without a previous descriptive study.

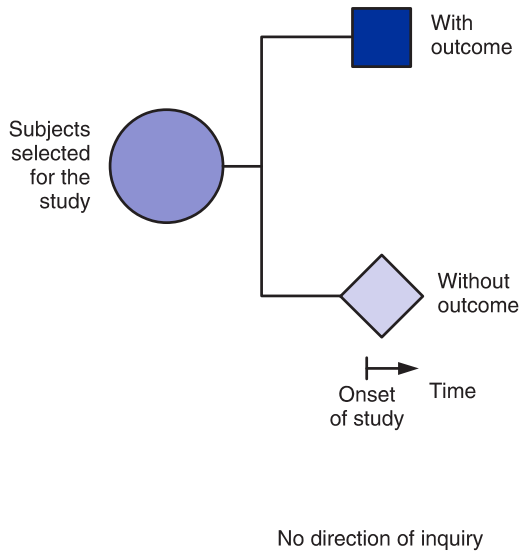

Figure 2–2. Schematic diagram of cross-sectional study design. Squares represent subjects with the outcome of interest; diamonds represent subjects without the outcome of interest. (Adapted with permission from Ilango K, Vijayakumar TM, Dubey GP, et al: An Enlarged Vision on Various Types of Study Design in Human Subjects, Global J Pharm 2012 Jan;6(3):216-221.)

Diagnosing or Staging a Disease: Anderson et al (2018) studied predictors of influenza in over 4,500 patients presenting to a hospital with flu-like symptoms from 2009 to 2014. They found that the most important symptoms for predicting influenza were cough, runny nose, chills, and body aches. They formulated a predictive model that was able to predict the presence/absence of the flu virus. Further, they tested the predictive value of a rapid influenza test versus virologically confirmed influenza cases.

Studying the Relationship Between Diseases: Poblador-Plou and her coinvestigators (2014) were interested in learning more about the relationship between dementia and other chronic diseases. Using electronic health records for patients identified with dementia, they were able to identify relationships with other chronic diseases such as Parkinson's disease, congestive heart failure, and others using a variety of statistical methods.

Establishing Norms: Knowledge of the range within which most patients fit is very useful to clinicians. Laboratories, of course, establish and then provide the normal limits of most diagnostic tests when they report the results for a given patient. Often these limits are established by testing people who are known to have normal values. We would not, for example, want to use people with diabetes mellitus to establish the norms for serum glucose levels. The results from the people known to have normal values are used to define the range that separates the lowest 2.5% of the values and the highest 2.5% of the values from the middle 95%. These values are called normal values, or norms.

Outside of the laboratory, there are many qualities for which normal ranges have not been established. This was true cognitive norms for Alzheimer's patients. Cognitive scores are an important tool used to detect patients with dementia, but may only be used if the distribution of normative scores is available. Kornak and colleagues (2018) analyzed data from the National Alzheimer's Coordinating Center (NACC). The investigators determined norms by exploring the relationships between age, sex, and other covariates to the cognitive scores for both normal subjects and those with dementia.

Surveys: Surveys are especially useful when the goal is to gain insight into a perplexing topic or to learn how people think and feel about an issue. Surveys are generally cross-sectional in design, but they can be used in case–control and cohort studies as well.

Monitoring the Future (MTF) is a longitudinal study that examines substance abuse in adolescents, college students, and adult high school graduates through age 55. Johnston et al (2018) compiled a summary of the data collected through 2017. They examined the trends in drug use including marijuana, bath salts, narcotics, tobacco, and alcohol based on 43,700 students in 360 secondary schools.

Interviews are sometimes used in surveys, especially when it is important to probe reasons or explanations more deeply than is possible with a written questionnaire. Interview surveys are also useful when the questions include topics that may require explanation due to complex topics or recalling particular events. The National Health Interview Survey (NHIS) has been conducted since 1962. The content and methodology of the survey has evolved over time to remain relevant and useful for research and investigation. The NHIS is an extensive survey that contains data regarding access to health care, cancer screening, health status, Internet, and email use as well as extensive sociodemographic data.

Many countries and states collect data on a variety of conditions to develop tumor registries, trauma, and databases of cases of infectious disease. Chaudhry and colleagues (2018) studied the number of cancer survivors based on the Ontario Cancer Registry (OCR) and health care administrative data. As cancer treatments advance, the number of survivors is increasing.

Understanding the number of survivors and their health status is an important public health question. The researchers included subjects with malignant cancer recorded in the OCR from 1964 to 2017. They found that 3% of the Ontario population were cancer survivors.

Cohort Studies

A **cohort** is a group of people who have something in common and who remain part of a group over an extended time. In medicine, the subjects in **cohort studies** are selected by some defining characteristic (or characteristics) suspected of being a precursor to or risk factor for a disease or health effect. Cohort studies ask the question "What will happen?" and thus, the direction in cohort studies is forward in time. Figure 2–3 illustrates the study design. Researchers select subjects at the onset of the study and then determine whether they have the risk factor or have been exposed. All subjects are followed over a certain period to observe the effect of the risk factor or exposure. Because the events of interest transpire after the study has begun, these studies are sometimes called **prospective studies.**

Typical Cohort Studies: A classical cohort study with which most of you are probably familiar is the Framingham study of cardiovascular disease. This study was begun in 1948 to investigate factors associated with the development of atherosclerotic and hypertensive cardiovascular disease, for which Gordon and Kannel (1970) reported a comprehensive 20-year follow-up. More than 6,000 citizens in Framingham, Massachusetts, agreed to participate in this long-term study that involved follow-up interviews and physical examinations every 2 years. Many journal articles have been written about this cohort, and some of the children of the original subjects are now being followed as well.

Cohort studies often examine what happens to the disease over time—the natural history of the disease. Many studies have been based on the Framingham cohort; hundreds of journal articles are indexed by **MEDLINE.** Many studies deal with

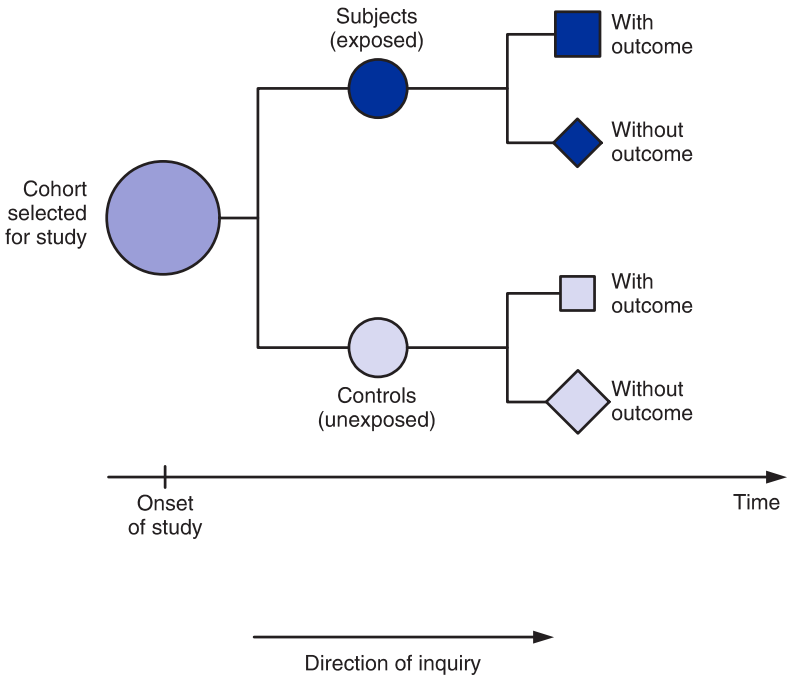

Figure 2–3. Schematic diagram of cohort study design. Shaded areas represent subjects exposed to the antecedent factor; unshaded areas correspond to unexposed subjects. Squares represent subjects with the outcome of interest; diamonds represent subjects without the outcome of interest. (Adapted with permission from Ilango K, Vijayakumar TM, Dubey GP, et al: An Enlarged Vision on Various Types of Study Design in Human Subjects, Global J Pharm 2012 Jan;6(3):216-221.)

cardiovascular-related conditions for which the study was designed, such as investigating cardiovascular biomarkers with heart failure (de Boer et al, 2018), but this very rich source of data is being used to study many other conditions as well. For instance, two recent articles examined treatable vascular disease and cognitive performance (van Eersel et al, 2019) and the relation of bone mass to hip fractures in women (McLean et al, 2018).

Although the Framingham Heart Study is very long term, many cohort studies follow subjects for a much shorter period. A presenting problem in Chapters 5 describes a cohort study to determine the effect of cholecystectomy on bowel habits and bile acid absorption (Dittrich et al, 2018). Thirteen subjects undergoing hypnosis were evaluated in three sessions at least 72 hours apart to detect changes such as EMG signals, peak muscle contraction, and M-wave amplitude.

Outcome Assessment: Increasingly, studies that assess medical outcomes are reported in the medical literature. Patient outcomes have always been of interest to health care providers; physicians and others in the health field are interested in how patients respond to different therapies and management regimens. There continues to be a growing focus on the ways in which patients view and value their health, the care they receive, and the results or outcomes of this care. The reasons for the increase in patient-focused health outcomes are complex, and some of the major ones are discussed later in this chapter.

Interest in outcome assessment was spurred by the Medical Outcomes Study (MOS), designed to determine whether variations in patient outcomes were related to the system of care, clinician specialty, and the technical and interpersonal skill of the clinician (Tarlov et al, 1989). Many subsequent studies looked at variations in outcomes in different geographic locations or among different ethnic groups that might result from access issues. In a cross-sectional study, Priede and colleagues (2018) studied models of social support in recently diagnosed cancer patients using the social support survey component of the MOS (MOS-SSS). They examined the results of the MOS-SSS and the Hospital Anxiety and Depression Scale (HADS) using factor analysis. The method allowed them to measure the structure of the survey and segment the questions into a five-factor model including: emotional, informational, tangible support, positive social interaction, and affection.

Functional status refers to a person's ability to perform their daily activities. Some researchers subdivide functional status into physical, emotional, mental, and social components (Gold et al, 1996). The 6-minute walk test (how far a person can walk in 6 minutes) was studied by Enright and colleagues (2003), and they recommended that the standards be adjusted for age, gender, height, and weight. Many instruments used to measure physical functional status have been developed to evaluate the extent of a patient's rehabilitation following injury or illness. These instruments are commonly called measures of activities of daily living (ADL). Cornelis and colleagues (2017) used the ADLS to aid in the early diagnosis of Alzheimer's disease.

Quality of life (QOL) is a broadly defined concept that includes subjective or objective judgments about all aspects of an individual's existence: health, economic status, environmental, and spiritual. Interest in measuring QOL was heightened when researchers realized that living a long time does not necessarily imply living a good life. QOL measures can help determine a patient's preferences for different health states and are often used to help decide among alternative approaches to medical management (Prigerson et al, 2015).

Patient satisfaction has been discussed for many years and has been shown to be highly associated with whether patients remain with the same physician provider and the degree to which they adhere to their treatment plan (Weingarten et al, 1995).

Patient satisfaction with medical care is influenced by a number of factors, not all of which are directly related to quality of care. The factors that influence patient satisfaction are often dependent on the reason for the contact. For example, Jacobs et al (2014) found that the most important factors driving patient satisfaction after total knee arthroplasty were extent of procedure and pain level post procedure as well as some demographic factors including race of the patient.

Cost-effectiveness and **cost–benefit analysis** are methods used to evaluate economic outcomes of interventions or different modes of treatment. Bagwell et al (2018) studied the effectiveness of intracapsular tonsillectomy and total tonsillectomy to treat pediatric obstructive sleep apnea (OSA). They used a decision tree model to simulate a model of choosing each of the two treatments. They found that when the recurrence rate of OSA was low (3.12%), partial tonsillectomy was more cost-effective. Cost-effectiveness analysis gives policy makers and health providers critical data needed to make informed judgments about interventions (Gold et al, 1996). A large number of questionnaires or instruments have been developed to measure outcomes. For quality of life, the most commonly used general-purpose instrument is the Medical Outcomes Study MOS 36-Item Short-Form Health Survey (SF-36). Originally developed at the RAND Corporation (Stewart et al, 1988), a refinement of the instrument has been validated and is now used worldwide to provide baseline measures and to monitor the results of medical care. The SF-36 provides a way to collect valid

data and does not require very much time to complete. The 36 items are combined to produce a patient profile on eight concepts in addition to summary physical and mental health measures.

Many instruments are problem-specific. Cramer and Spilker (1998) provide a broad overview of approaches to QOL assessment, evaluations of outcomes, and pharmacoeconomic methods—both general purpose and disease-specific.

Some outcome studies address a whole host of topics, and we have used several as presenting problems in upcoming chapters. As efforts continue to contain costs of medical care while maintaining a high level of patient care, we expect to see many additional studies focusing on patient outcomes. The journal *Medical Care* is devoted exclusively to outcome studies.

Historical Cohort Studies: Many cohort studies are prospective; that is, they begin at a specific time, the presence or absence of the risk factor is determined, and then information about the outcome of interest is collected at some future time, as in the two studies described earlier. One can also undertake a cohort study by using information collected in the past and kept in records or files.

For example, St. Sauver and colleagues (2015) studied the risk of developing multimorbidity using data from 123,716 residents of Olmsted County Minnesota. They defined multimorbidity as the development of at least 2 of the 20 chronic conditions selected by HHS. They found that the incidence of multimorbidity increased with age, but the number of people with more than one chronic condition was greater for those under 65 than 65 and older.

Some investigators call this type of study a **historical cohort study** or **retrospective cohort study** because historical information is used; that is, the events being evaluated actually occurred before the onset of the study (Figure 2–4). Note that the direction of the inquiry is still forward in time, from a possible cause or risk factor to an outcome. Studies that merely describe an investigator's experience with a group of patients and attempt to identify features associated with a good or bad outcome fall into this category, and many such studies are published in the medical literature.

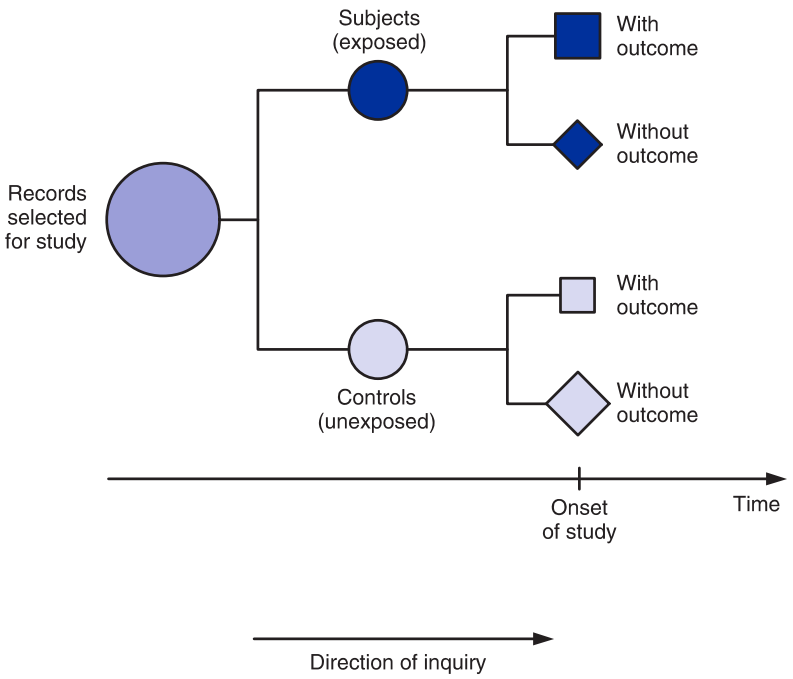

Figure 2–4. Schematic diagram of historical cohort study design. Shaded areas represent subjects exposed to the antecedent factor; unshaded areas correspond to unexposed subjects. Squares represent subjects with the outcome of interest; diamonds represent subjects without the outcome of interest. (Adapted with permission from Ilango K, Vijayakumar TM, Dubey GP, et al: An Enlarged Vision on Various Types of Study Design in Human Subjects, Global J Pharm 2012 Jan;6(3):216-221.)

The time relationship among the different observation study designs is illustrated in Figure 2–5. The figure shows the timing of surveys, which have no direction of inquiry, case–control designs, which look backward in time, and cohort studies, which look forward in time.

Comparison of Case–Control and Cohort Studies

Both case–control and cohort studies evaluate risks and causes of disease, and the design an investigator selects depends in part on the research question.

Moore and colleagues (2016) undertook a matched case–control study to look at the effectiveness of pneumonia vaccines in children. They examined 722 children with pneumonia and 2,991 controls. They found that 13-valent pneumococcal conjugate vaccine (PCV13) was highly effective against the disease.

As this illustration shows, a case–control study takes the outcome as the starting point of the inquiry and looks for precursors or risk factors; while a cohort study starts with a risk factor or exposure and looks at consequences.

Generally speaking, results from a well-designed cohort study carry more weight in understanding a disease than do results from a case–control study. A large number of possible biasing factors can play a role in case–control studies, and several of them are discussed at greater length in Chapter 13.

In spite of their shortcomings with respect to establishing causality, case–control studies are frequently used in medicine and can provide useful insights if well designed. They can be completed in a much shorter time than cohort studies and are correspondingly less expensive to undertake. Case–control studies are especially useful for studying rare conditions or diseases that may not manifest themselves for many years. In addition, they are valuable for testing an original premise; if the results of the case–control study are promising, the investigator can design and undertake a more involved cohort study.

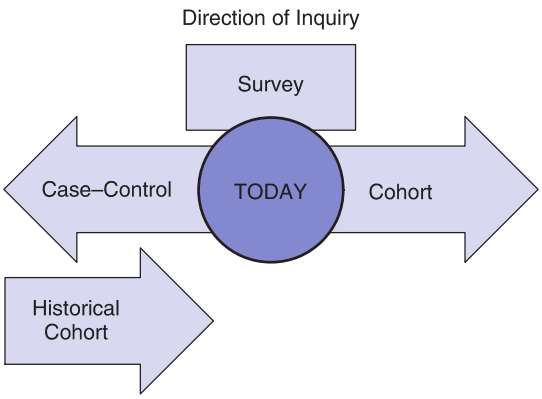

Figure 2–5. Schematic diagram of the time relationship among different observational study designs. The arrows represent the direction of the inquiry.

EXPERIMENTAL STUDIES OR CLINICAL TRIALS

Experimental studies are generally easier to identify than observational studies in the medical literature. Authors of medical journal articles reporting experimental studies tend to state explicitly the type of study design used more often than do authors reporting observational studies. Experimental studies in medicine that involve humans are called **clinical trials** because their purpose is to draw conclusions about a particular procedure or treatment. Table 2–1 indicates that clinical trials fall into two categories: those with and those without controls.

Controlled trials are studies in which the experimental drug or procedure is compared with another drug or procedure, sometimes a placebo and sometimes the previously accepted treatment. Uncontrolled trials are studies in which the investigators' experience with the experimental drug or procedure is described, but the treatment is not compared with another treatment, at least not formally. Because the purpose of an experiment is to determine whether the intervention (treatment) makes a difference, studies with controls are much more likely than those without controls to detect whether the difference is due to the experimental treatment or to some other factor. Thus, controlled studies are viewed as having far greater validity in medicine than uncontrolled studies. The consolidated standard of reporting trials (CONSORT) guidelines reflect an effort to improve the reporting of clinical trials. The CONSORT statement was last updated in 2010 and may be found on the CONSORT Web site (www.consort-statement.org).

Trials with Independent Concurrent Controls

One way a trial can be controlled is to have two groups of subjects: one that receives the experimental procedure (the experimental group) and the other that receives the placebo or standard procedure (the control group; Figure 2–6). The experimental and control groups should be treated alike in all ways except for the procedure itself so that any differences between the groups will be due to the procedure and not to other

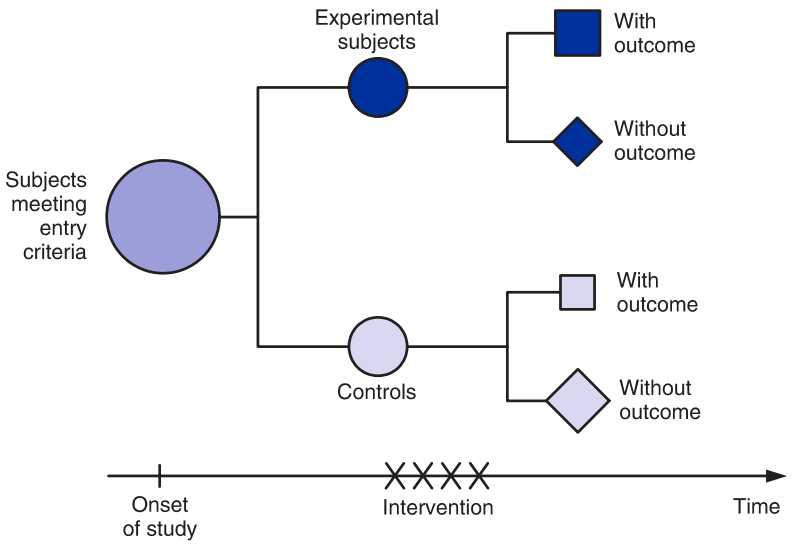

Figure 2–6. Schematic diagram of randomized controlled trial design. Shaded areas represent subjects assigned to the treatment condition; unshaded areas correspond to subjects assigned to the control condition. Squares represent subjects with the outcome of interest; diamonds represent subjects without the outcome of interest.

factors. The best way to ensure that the groups are treated similarly is to plan interventions for both groups for the same time period in the same study. In this way, the study achieves **concurrent control.** To reduce the chances that subjects or investigators see what they expect to see, researchers can design **double-blind trials** in which neither subjects nor investigators know whether the subject is in the treatment or the control group. When only the subject is unaware, the study is called a **blind trial.** In some unusual situations, the study design may call for the investigator to be blinded even when the subject cannot be blinded. Blindedness is discussed in detail in Chapter 13. Another issue is how to assign some patients to the experimental condition and others to the control condition; the best method of assignment is random assignment. Methods for randomization are discussed in Chapter 4.

Randomized Controlled Trials: The **randomized controlled trial** is the epitome of all research designs because it provides the strongest evidence for concluding causation; it provides the best insurance that the result was due to the intervention.

One of the more noteworthy randomized trials is the Physicians' Health Study (Steering Committee of the Physicians' Health Study Research Group, 1989), which investigated the role of aspirin in reducing the risk of cardiovascular disease. One purpose was to learn whether aspirin in low doses reduces the mortality rate from cardiovascular disease. Participants in this clinical trial were over 22,000 healthy male physicians who were randomly assigned to receive aspirin or placebo and were followed over an average period of 60 months. The investigators found that fewer physicians in the aspirin group experienced a myocardial infarction during the course of the study than did physicians in the group receiving placebo.

Nonrandomized Trials: Subjects are not always randomized to treatment options. Studies that do not use randomized assignment are generally referred to as **nonrandomized trials** or simply as clinical trials or comparative studies, with no mention of randomization. Many investigators believe that studies with nonrandomized controls are open to so many sources of bias that their conclusions are highly questionable. Studies using nonrandomized controls are considered to be much weaker because they do nothing to prevent bias in patient assignment. For instance, perhaps it is the stronger patients who receive the more aggressive treatment and the higher risk patients who are treated conservatively. An example of a nonrandomized study is a study comparing traditional lecture versus case-based learning and simulation in nurse education (Raurell-Torredà et al, 2014). The investigators studied 66 undergraduates enrolled in a traditional lecture and discussion course and 35 enrolled in a course that also included a case-based learning component. These two groups were then compared to

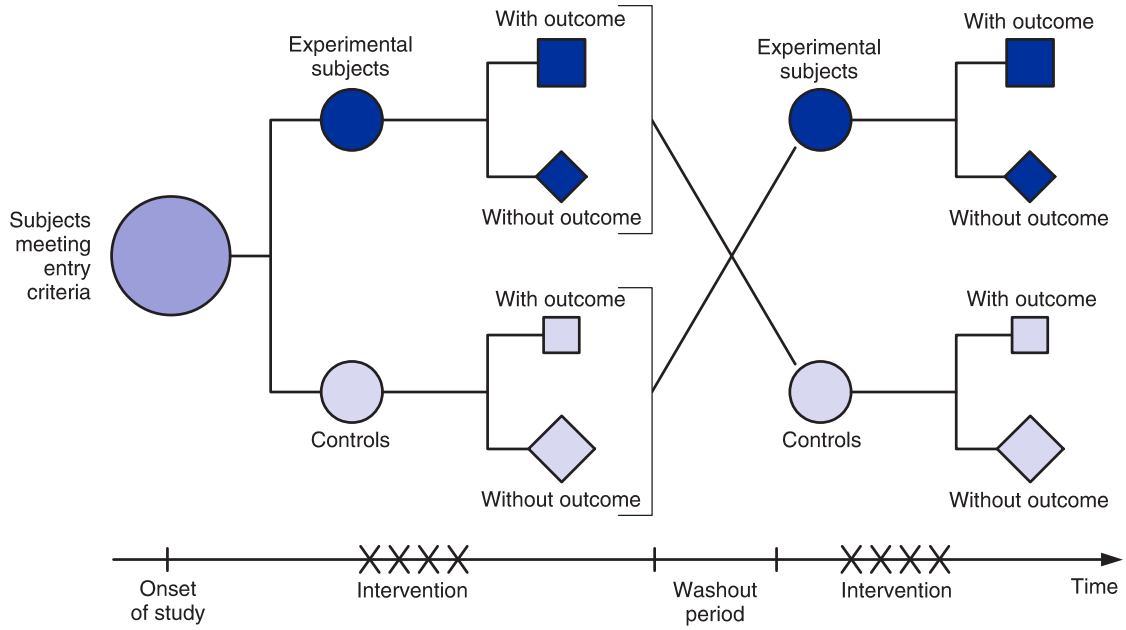

Figure 2–7. Schematic diagram of trial with crossover. Shaded areas represent subjects assigned to the treatment condition; unshaded areas correspond to subjects assigned to the control condition. Squares represent subjects with the outcome of interest; diamonds represent subjects without the outcome of interest.

59 continuing professional education (CPE) nurses with clinical experience. After administering a simulated clinical exam test, they found the intervention group (case-based learning) performed better than the traditional learning group.

Trials with Self-Controls

Moderate level of control can be obtained by using the same group of subjects for both experimental and control options. The study by Goto and colleagues (2018) examined the risk of acute exacerbation of COPD after bariatric surgery. They followed obese adults with COPD that underwent bariatric surgery. They compared the risk of an acute exacerbation in the 12-month period after surgery to months 13 to 24 before surgery. This type of study uses patients as their own controls and is called a **self-controlled study.** Studies with self-controls and no other control group are still vulnerable to the well-known Hawthorne effect, described by Roethlisberger and colleagues (1946), in which people change their behavior and sometimes improve simply because they receive special attention by being in a study and not because of the study intervention. These studies are similar to cohort studies except for the intervention or treatment that is involved.

The self-controlled study design can be modified to provide a combination of concurrent and self-controls. This design uses two groups of patients: one group is assigned to the experimental treatment, and the second group is assigned to the placebo or control treatment (Figure 2–7). After a time, the experimental treatment and placebo are withdrawn from both groups for a "washout" period. During the washout period, the patients generally receive no treatment. The groups are then given the alternative treatment; that is, the first group now receives the placebo, and the second group receives the experimental treatment. This design, called a **crossover study,** is powerful when used appropriately.

Trials with External Controls

The third method for controlling experiments is to use controls external to the study. Sometimes, the result of another investigator's research is used as a comparison. On other occasions, the controls are patients the investigator has previously treated in another manner, called **historical controls.** The study design is illustrated in Figure 2–8.

Historical controls are frequently used to study diseases for which cures do not yet exist and are used in oncology studies, although oncologic studies use concurrent controls when possible. In studies involving

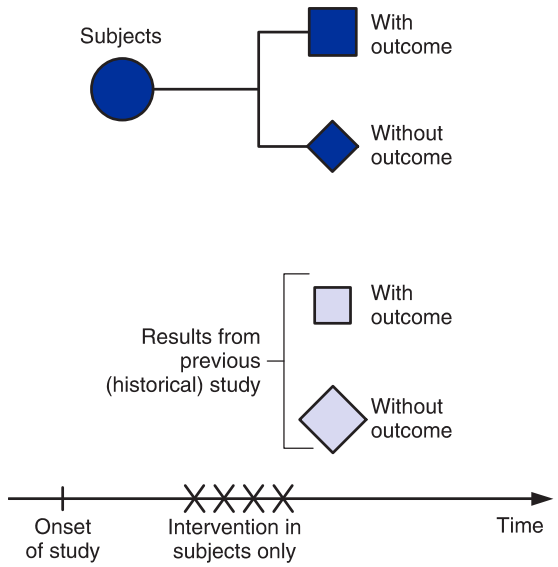

Figure 2–8. Schematic diagram of trial with external controls. Shaded areas represent subjects assigned to the treatment condition; unshaded areas correspond to patients cared for under the control condition. Squares represent subjects with the outcome of interest; diamonds represent subjects without the outcome of interest.

historical controls, researchers should evaluate whether other factors may have changed since the time the historical controls were treated; if so, any differences may be due to these other factors and not to the treatment.

Uncontrolled Studies

Not all studies involving interventions have controls, and by strict definition they are not really experiments or trials. For example, Bottegoni and associates (2016) reported the results of a trial of administering homologous platelet-rich plasma in elderly patients with knee osteoarthritis. Subjects were followed for a 2-month and 6-month visit after administration. The researchers found that there was some short-term clinical improvement after the treatment and that 90% of the patients were satisfied with the results 6 months after treatment. This study was an **uncontrolled study** because there were no comparisons with patients treated in another manner.

Uncontrolled studies are more likely to be used when the comparison involves a procedure than when it involves a drug. The major shortcoming of such studies is that investigators assume that the procedure used and described is the best one. The history of medicine is filled with examples in which one particular treatment is recommended and then discontinued after a controlled clinical trial is undertaken. One significant problem with uncontrolled trials is that unproved procedures and therapies can become established, making it very difficult for researchers to undertake subsequent controlled studies. Another problem is finding a significant difference when it may be unfounded. Guyatt and colleagues (2000) identified 13 randomized trials and 17 observational studies in adolescent pregnancy prevention. Six of eight outcomes they examined showed a significant intervention effect in the observational studies, whereas the randomized studies showed no benefit.

META-ANALYSIS & REVIEW PAPERS

A type of study that does not fit specifically in either category of observation studies or experiments is called **meta-analysis**. Meta-analysis uses published information from other studies and combines the results so as to permit an overall conclusion. Meta-analysis is similar to review articles, but additionally includes a quantitative assessment and summary of the findings. It is possible to do a meta-analysis of observational studies or experiments; however, a meta-analysis should report the findings for these two types of study designs separately. This method is especially appropriate when the studies that have been reported have small numbers of subjects or come to different conclusions.

Finnerup and colleagues (2015) performed a meta-analysis of neuropathic pain in adults. The investigators wanted to know if topical or oral medications were more effective in treating pain. They found 229 studies that had addressed this question and combined the results in a statistical manner to reach an overall conclusion about their effectiveness—mainly that the evidence supporting the use of oral medications was stronger.

ADVANTAGES & DISADVANTAGES OF DIFFERENT STUDY DESIGNS

The previous sections introduced the major types of study designs used in medical research, broadly divided into experimental studies, or clinical trials, and observational studies (cohort, case–control, cross-sectional, and case–series designs). Each study design has certain advantages over the others as well as some specific disadvantages, which we discuss in the following sections.

Advantages & Disadvantages of Clinical Trials

The randomized clinical trial is the gold standard, or reference, in medicine; it is the design against which others are judged—because it provides the greatest

justification for concluding causality and is subject to the least number of problems or biases. Clinical trials are the best type of study to use when the objective is to establish the efficacy of a treatment or a procedure. Clinical trials in which patients are randomly assigned to different treatments, or "arms," are the strongest design of all. One of the treatments is the experimental condition; another is the control condition. The control may be a placebo or a sham procedure; often, it is the treatment or procedure commonly used, called the standard of care or reference standard. A number of published articles have shown the tendency for nonrandomized studies, especially those using historical controls, to be more likely to show a positive outcome, compared with randomized studies. In some situations, however, historical controls can and should be used. For instance, historical controls may be useful when preliminary studies are needed or when researchers are dealing with late treatment for an intractable disease, such as advanced cancer. Although clinical trials provide the greatest justification for determining causation, obstacles to using them include their great expense and long duration. For instance, a randomized trial comparing various treatments for carcinoma requires the investigators to follow the subjects for a long time. Another potential obstacle to using clinical trials occurs when certain practices become established and accepted by the medical community, even though they have not been properly justified. As a result, procedures become established that may be harmful to many patients, as evidenced by the controversy over silicone breast implants and the many different approaches to managing hypertension, many of which have never been subjected to a clinical trial that includes the most conservative treatment, diuretics.

Advantages & Disadvantages of Cohort Studies

Cohort studies are the design of choice for studying the causes of a condition, the course of a disease, or the risk factors because they are longitudinal and follow a group of subjects over a period of time. Causation generally cannot be proved with cohort studies because they are observational and do not involve interventions. However, because they follow a cohort of patients forward through time, they possess the correct time sequence to provide strong evidence for possible causes and effects, as in the smoking and lung cancer controversy. In well-designed cohort studies, investigators can control many sources of bias related to patient selection and recorded measurements.

The length of time required in a cohort study depends on the problem studied. With diseases that develop over a long period of time or with conditions that occur as a result of long-term exposure to some causative agent, many years are needed for study. Extended time periods make such studies costly. They also make it difficult for investigators to argue causation because other events occurring in the intervening period may have affected the outcome. For example, the long time between exposure and effect is one of the reasons it is difficult to study the possible relationship between environmental agents and various carcinomas. Cohort studies that require a long time to complete are especially vulnerable to problems associated with patient follow-up, particularly patient attrition (patients stop participating in the study) and patient migration (patients move to other communities). This is one reason that the Framingham study, with its rigorous methods of follow-up, is such a rich source of important information.

Advantages & Disadvantages of Case–Control Studies

Case–control studies are especially appropriate for studying rare diseases or events, for examining conditions that develop over a long time, and for investigating a preliminary hypothesis. They are generally the quickest and least expensive studies to undertake and are ideal for investigators who need to obtain some preliminary data prior to writing a proposal for a more complete, expensive, and time-consuming study. They are also a good choice for someone who needs to complete a clinical research project in a specific amount of time.

The advantages of case–control studies lead to their disadvantages. Of all study methods, they have the largest number of possible biases or errors, and they depend completely on high-quality existing records. Data availability for case–control studies sometimes requires compromises between what researchers wish to study and what they are able to study. One of the previous edition authors was involved in a study of elderly burn patients in which the goal was to determine risk factors for survival. The primary investigator wanted to collect data on fluid intake and output. He found, however, that not all of the existing patient records contained this information, and thus it was impossible to study the effect of this factor.

One of the greatest problems in a case–control study is selection of an appropriate control group. The cases in a case–control study are relatively easy to identify, but deciding on a group of persons who provide a relevant comparison is more difficult. Because of the problems inherent in choosing a control group in a case–control study, some statisticians have recommended the use of two control groups: one control group similar in some ways to the cases (e.g., having been hospitalized during the same period of time) and another control group of healthy subjects.

Advantages & Disadvantages of Cross-Sectional Studies

Cross-sectional studies are best for determining the status quo of a disease or condition, such as the prevalence of HIV in given populations, and for evaluating diagnostic procedures. Cross-sectional studies are similar to case–control studies in being relatively quick to complete, and they may be relatively inexpensive as well. Their primary disadvantage is that they provide only a "snapshot in time" of the disease or process, which may result in misleading information if the research question is really one of disease process. For example, clinicians used to believe that diastolic blood pressure, unlike systolic pressure, does not increase as patients grow older. This belief was based on cross-sectional studies that had shown mean diastolic blood pressure to be approximately 80 mm Hg in all age groups. In the Framingham cohort study, however, the patients who were followed over a period of several years were observed to have increased diastolic blood pressure as they grew older (Gordon et al, 1959).

This apparent contradiction is easier to understand if we consider what happens in an aging cohort. For example, suppose that the mean diastolic pressure in men aged 40 years is 80 mm Hg, although there is individual variation, with some men having a blood pressure as low as 60 mm Hg and others having a pressure as high as 100 mm Hg. Ten years later, there is an increase in diastolic pressure, although it is not an even increase; some men experience a greater increase than others. The men who were at the upper end of the blood pressure **distribution** 10 years earlier and who had experienced a larger increase have died in the intervening period, so they are no longer represented in a cross-sectional study. As a result, the mean diastolic pressure of the men still in the cohort at age 50 is about 80 mm Hg, even though individually their pressures are higher than they were 10 years earlier. Thus, a cohort study, not a cross-sectional study, provides the information leading to a correct understanding of the relationship between normal aging and physiologic processes such as diastolic blood pressure.

Surveys are generally cross-sectional studies. Most of the voter polls done prior to an election are one-time samplings of a group of citizens, and different results from week to week are based on different groups of people; that is, the same group of citizens is not followed to determine voting preferences through time. Similarly, consumer-oriented studies on customer satisfaction with automobiles, appliances, health care, and so on are cross-sectional.

A common problem with survey research is obtaining sufficiently large response rates; many people asked to participate in a survey decline because they are busy, not interested, and so forth. The conclusions are, therefore, based on a subset of people who agree to participate, and these people may not be **representative** of or similar to the entire population. The problem of representative participants is not confined to cross-sectional studies; it can be an issue in other studies whenever subjects are selected or asked to participate and decline or drop out. Another issue is the way questions are posed to participants; if questions are asked in a leading or emotionally inflammatory way, the responses may not truly represent the participants' feelings or opinions. We discuss issues with surveys more completely in Chapter 11.

Advantages & Disadvantages of Case–Series Studies

Case–series reports have two advantages: They are easy to write, and the observations may be extremely useful to investigators designing a study to evaluate causes or explanations of the observations. But as we noted previously, case–series studies are susceptible to many possible biases related to subject selection and characteristics observed. In general, you should view them as hypothesis-generating and not as conclusive.

SUMMARY

This chapter illustrates the study designs most frequently encountered in the medical literature. In medical research, subjects are observed or experiments are undertaken. Experiments involving humans are called trials. Experimental studies may also use animals and tissue, although we did not discuss them as a separate category; the comments pertaining to clinical trials are relevant to animal and tissue studies as well.

Each type of study discussed has advantages and disadvantages. Randomized, controlled clinical trials are the most powerful designs possible in medical research, but they are often expensive and time-consuming. Well-designed observational studies can provide useful insights on disease causation, even though they do not constitute proof of causes. Cohort studies are best for studying the natural progression of disease or risk factors for disease; case–control studies are much quicker and less expensive. Cross-sectional studies provide a snapshot of a disease or condition at one time, and we must be cautious in inferring disease progression from them. Surveys, if properly done, are useful in obtaining current opinions and practices. Case–series studies should be used only to raise questions for further research.

We have used several presenting problems from later chapters to illustrate different study designs. We will

point out salient features in the design of the presenting problems as we go along, and we will return to the topic of study design again after all the prerequisites for evaluating the quality of journal articles have been presented.

EXERCISES

Read the descriptions of the following studies and determine the study design used.

1. *Researchers wanted to determine if adding vancomycin to the protocol for shunt insertion would reduce the infection rate (van Lindert et al, 2018). The researchers compared patients with shunt insertions prior to the protocol change (263 procedures from January 2010 to December 2011) with those after the addition of vancomycin to the protocol (499 procedures from April 2012 to December 2015).*
2. *Priede and coworkers (2018) studied the level of psychological stress in newly diagnosed cancer patients using the MOS-SSS survey. Patients were recruited from December 2011 to October 2013.*
3. *The Prostate Cancer Outcomes Study was designed to investigate the patterns of cancer care and effects of treatment on quality of life. Hoffman and coworkers (2017) identified eligible cases from one SEER tumor registry. They surveyed 934 known survivors to assess treatment decision regret. Multivariate logistic regression was used to investigate the factors related to regret.*
4. *The relationship between exposure to benzodiazepine and Alzheimer's disease was investigated by Billioti de Gage and colleagues (2014). Subjects with Alzheimer's disease were matched with controls based on sex age group and duration of follow-up.*
5. *A study to determine whether radiation treatment with or without anti-androgen therapy in recurrent prostate cancer (Shipley et al, 2017). The primary outcome was overall survival.*
6. *Eckel et al (2018) reported on the relationship between transition from metabolic healthy to unhealthy status and association with cardiovascular disease. Subjects in the study were selected from the Nurses' Health Study originally completed in 1976; the study included 120,000 married female registered nurses, aged 30–55. The original survey provided information on the subjects' age, parental history of myocardial infarction, smoking status, height, weight, use of oral contraceptives or postmenopausal hormones, and history of myocardial infarction or angina pectoris, diabetes, hypertension, or high serum cholesterol levels. Follow-up surveys were every 2 years thereafter.*
7. ***Group Exercise.** The abuse of phenacetin, a common ingredient of analgesic drugs, can lead to kidney disease. There is also evidence that use of salicylate provides protection against cardiovascular disease. How would you design a study to examine the effects of these two drugs on mortality due to different causes and on cardiovascular morbidity?*
8. ***Group Exercise.** Select a study with an interesting topic, either one of the studies referred to in this chapter or from a current journal. Carefully examine the research question and decide which study design would be optimal to answer the question. Is that the study design used by the investigators? If so, were the investigators attentive to potential problems identified in this chapter? If not, what are the reasons for the study design used? Do they make sense?*

Summarizing Data & Presenting Data in Tables & Graphs

3

KEY CONCEPTS

1. *All observations of subjects in a study are evaluated on a scale of measurement that determines how the observations should be summarized, displayed, and analyzed.*
2. *Nominal scales are used to categorize discrete characteristics.*
3. *Ordinal scales categorize characteristics that have an inherent order.*
4. *Numerical scales measure the amount or quantity of something.*
5. *Means measure the middle of the distribution of a numerical characteristic.*
6. *Medians measure the middle of the distribution of an ordinal characteristic or a numerical characteristic that is skewed.*
7. *The standard deviation is a measure of the spread of observations around the mean and is used in many statistical procedures.*
8. *The coefficient of variation is a measure of relative spread that permits the comparison of observations measured on different scales.*
9. *Percentiles are useful to compare an individual observation with a norm.*
10. *Stem-and-leaf plots are a combination of frequency tables and histograms that are useful in exploring the distribution of a set of observations.*
11. *Frequency tables show the number of observations having a specific characteristic.*
12. *Histograms, box plots, and frequency polygons display distributions of numerical observations.*
13. *Proportions and percentages are used to summarize nominal and ordinal data.*
14. *Rates describe the number of events that occur in a given period.*
15. *Prevalence and incidence are two important measures of morbidity.*
16. *Rates must be adjusted when populations being compared differ in an important confounding factor.*
17. *The relationship between two numerical characteristics is described by the correlation.*
18. *The relationship between two nominal characteristics is described by the risk ratio, odds ratio, and event rates.*
19. *Number needed to treat is a useful indication of the effectiveness of a given therapy or procedure.*
20. *Scatterplots illustrate the relationship between two numerical characteristics.*
21. *Poorly designed graphs and tables mislead in the information they provide.*