CHAPTER 6

Overview of Epidemiological Study Designs

LEARNING OBJECTIVES

By the end of this chapter the reader will be able to:

- Distinguish between experimental and observational studies.
- Describe the key characteristics of experimental, cohort, case–control, cross-sectional, and ecological studies regarding subject selection, data collection, and analysis.
- Identify the design of a particular study.
- Discuss the factors that determine when a particular design is indicated.

CHAPTER 6

Overview of Epidemiological Study Designs

LEARNING OBJECTIVES

By the end of this chapter the reader will be able to:

- Distinguish between experimental and observational studies.
- Describe the key characteristics of experimental, cohort, case–control, cross-sectional, and ecological studies regarding subject selection, data collection, and analysis.
- Identify the design of a particular study.
- Discuss the factors that determine when a particular design is indicated.

Introduction

Epidemiology is the study of the distribution and determinants of disease frequency in human populations and the application of this study to control health problems.[1(p1),2(p95)] The term ***study*** includes both **surveillance**, whose purpose is to monitor aspects of disease occurrence and spread that are pertinent to effective control,[3(p95)] and epidemiological research, whose goal is to harvest valid and precise information about the causes, preventions, and treatments for disease. The term ***disease*** refers to a broad array of health-related states and events, including diseases, injuries, disabilities, and death.

Epidemiological research encompasses several types of study designs, including experimental studies and observational studies, such as cohort and case–control studies. Each type of epidemiological study design simply represents a different way of harvesting information. The selection of one design over another depends on the particular research question, concerns about validity and efficiency, and practical and ethical considerations. For example, experimental studies, also known as trials, investigate the role of some factor or agent in the prevention or treatment of a disease. In this type of study, the investigator assigns individuals to two or more groups that either receive or do not receive the preventive or therapeutic agent. Because experimental studies closely resemble controlled laboratory investigations, they are thought to produce the most scientifically rigorous data of all the designs.

However, experimental studies are often infeasible because of difficulties enrolling participants, high costs, and thorny ethical issues, most epidemiological research is conducted using an **observational study**, which is considered a “natural” experiment because the investigator lets nature take its course. Observational studies take advantage of the fact that people are exposed to noxious and/or healthy substances through their personal habits, occupation, place of residence, and so on. The studies provide information on exposures that occur in natural settings, and they are not limited to preventions and treatments. Furthermore, they do not suffer from the ethical and feasibility issues of experimental studies. For example, although it is unethical to conduct an experimental study of the effect of drinking alcohol on the developing fetus by assigning newly pregnant women to either a drinking or nondrinking group, it is perfectly ethical to conduct an observational study by comparing women who choose to drink during pregnancy with those who decide not.

The two principal types of observational studies are cohort and case–control studies. A classic cohort study examines one or more health effects of exposure to a single agent. Subjects are defined according to their exposure status and followed over time to determine the incidence of health outcomes. In contrast, a classic case–control study examines a single disease in relation to exposure to one or more agents. Cases that have the disease of interest and controls who are a sample of the population that produced the cases are defined and enrolled. The purpose of the control group is to provide information on the exposure distribution in the population that gave rise to the cases. Investigators obtain and compare exposure histories of cases as well as of controls.

Additional observational study designs include cross-sectional studies and ecological studies. A cross-sectional study examines the relationship between a disease and an exposure among individuals in a defined population at a point in time. Thus, it takes a snapshot of a population and usually measures the exposure prevalence in relation to the disease prevalence. An **ecological study** evaluates an association

using the population rather than the individual as the unit of analysis. The rates of disease are examined in relation to factors described on the population level. Both the cross-sectional and ecological designs have important limitations that make them less scientifically rigorous than cohort and case–control studies. These limitations are discussed later in this chapter.

An overview of these study designs is provided in **TABLE 6-1**. The goal of all these studies is to determine the relationship between an exposure and a disease with validity and precision using minimal resources. **Validity** is defined as the lack of bias and confounding. Bias is an error committed by the investigator in the design or conduct of a study that leads to a false association between the exposure and disease. Confounding, on the other hand, is not the fault of the investigator but rather reflects the fact that epidemiological research is conducted among free-living humans with unevenly distributed characteristics. As a result, epidemiological studies that try to determine the relationship between an exposure and a disease are susceptible to the disturbing influences of extraneous factors known as confounders. Precision is the lack of random error, which leads to a false association between the exposure and disease just by "chance," an uncontrollable force that seems to have no assignable cause.[4(p95)]

TABLE 6.1 Main Types of Epidemiological Studies

Type of study	Characteristics
Experimental	Studies preventions and treatments for diseases; investigator actively manipulates which groups receive the agent under study.
Observational	Studies causes, preventions, and treatments for diseases; investigator passively observes as nature takes its course.
Cohort	Typically examines multiple health effects of an exposure; subjects are defined according to their exposure levels and followed for disease occurrence.
Case–control	Typically examines multiple exposures in relation to a disease; subjects are defined as cases and controls, and exposure histories are compared.
Cross-sectional	Typically examines the relationship between exposure and disease prevalence in a defined population at a single point in time
Ecological	Examines the relationship between exposure and disease with population-level rather than individual-level data

Several factors help epidemiologists determine the most appropriate study design for evaluating a particular association, including the hypothesis being tested, state of knowledge, and frequency of the exposure and the disease and expected strength of the association between the two. This chapter provides (1) an overview of epidemiological research designs—experimental, cohort, case–control, case–crossover, ecological, and agent-based modeling—and (2) a description of the settings in which the three main study designs—experimental, cohort, and case–control—are most appropriate.

Overview of Experimental Studies

Definitions and Classification

An experimental study, also known as a trial, investigates the role of some agent in the prevention or treatment of a disease. In this type of study, the investigator assigns individuals to two or more groups that either receive or do not receive the preventive or therapeutic agent. The group that is allocated the agent under study is generally called the **treatment group**, and the group that is not allocated the agent under study is called the **comparison group**. Depending on the purpose of the trial, the comparison group may receive no treatment at all, an inactive treatment such as a placebo, or another active treatment.

The active manipulation of the agent by the investigator is the hallmark that distinguishes experimental from observational studies. In the latter, the investigator acts as a passive observer, merely letting nature take its course. Because experimental studies more closely resemble controlled laboratory investigations, most epidemiologists believe that experimental studies produce more scientifically rigorous results than do observational studies.

Experimental studies are commonly classified by their objective, that is, by whether they investigate a measure that prevents disease occurrence or a measure that treats an existing condition. The former is known as a preventive or prophylactic trial, and the latter is known as a therapeutic or clinical trial. In preventive trials, agents such as vitamins or behavioral modifications such as smoking cessation are studied to determine whether they are effective in preventing or delaying the onset of disease among healthy individuals. In therapeutic trials, treatments such as surgery, radiation, and drugs are tested among individuals who already have a disease. A schematic representation of a typical experimental study is presented in **FIGURE 6-1**.

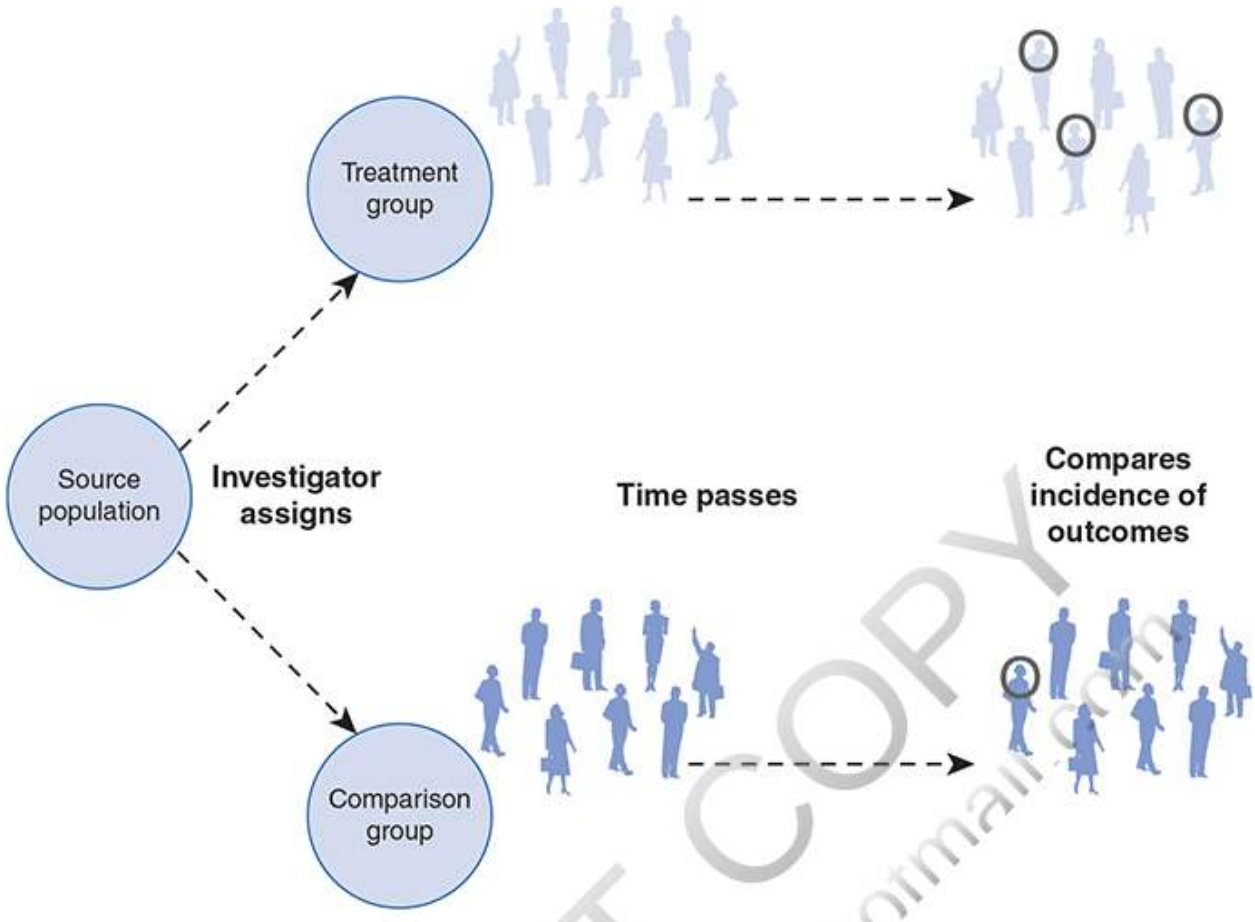

FIGURE 6.1 Schematic representation of experimental study implementation.

Selection of Study Population

During the recruitment phase of an experimental study, the study population, which is also called the experimental population, is enrolled on the basis of eligibility criteria that reflect the purpose of the trial as well as scientific, safety, and practical considerations. For example, healthy or high-risk individuals are enrolled in prevention trials, whereas individuals with specific diseases are enrolled in therapeutic trials. Additional inclusion and exclusion criteria may be used to restrict the study population by factors such as gender and age.

The study population must include an adequate number of individuals to determine whether there is a true difference between the treatment and comparison groups. An investigator determines how many subjects to include by using formulas that take into account the anticipated difference between the groups, the background rate of the outcome, and

recurrence, symptom improvement, length of survival, or side effects. The length of follow-up depends on the particular outcome under study. It can range from a few months to a few decades.

Usually, all reported outcomes under study are confirmed to guarantee their accuracy. Confirmation is typically done by masked investigators who gather corroborating information from objective sources, such as medical records and laboratory tests. High and comparable follow-up rates are needed to ensure the quality of the outcome data. Follow-up is adversely affected when participants withdraw from the study (these individuals are called dropouts) or cannot be located or contacted by the investigator (these individuals are termed lost to follow-up). Reasons for dropouts and losses include relocation, waning interest, and adverse reactions to the treatment.

Analysis

The classic analytic approach for an experimental study is known as an intent-to-treat or treatment assignment analysis. In this analysis, all individuals who were randomly allocated to a treatment are analyzed regardless of whether they completed the regimen or received the treatment.[8] An intent-to-treat analysis gives information on the effectiveness of a treatment under everyday practice conditions. The alternative to an intent-to-treat analysis is known as an efficacy analysis, which determines the treatment effects under ideal conditions, such as when participants take the full treatment exactly as directed.

Overview of Cohort Studies

Definitions

A **cohort** is defined as a group of people with a common characteristic or experience. In a cohort study, healthy subjects are defined according to their exposure status and followed over time to determine the incidence of symptoms, disease, or death. The common characteristic for grouping subjects is their exposure level. Usually, two groups are compared: an **exposed** and an **unexposed** group. The unexposed group is called the reference, referent, or comparison group (see **FIGURE 6-2**).

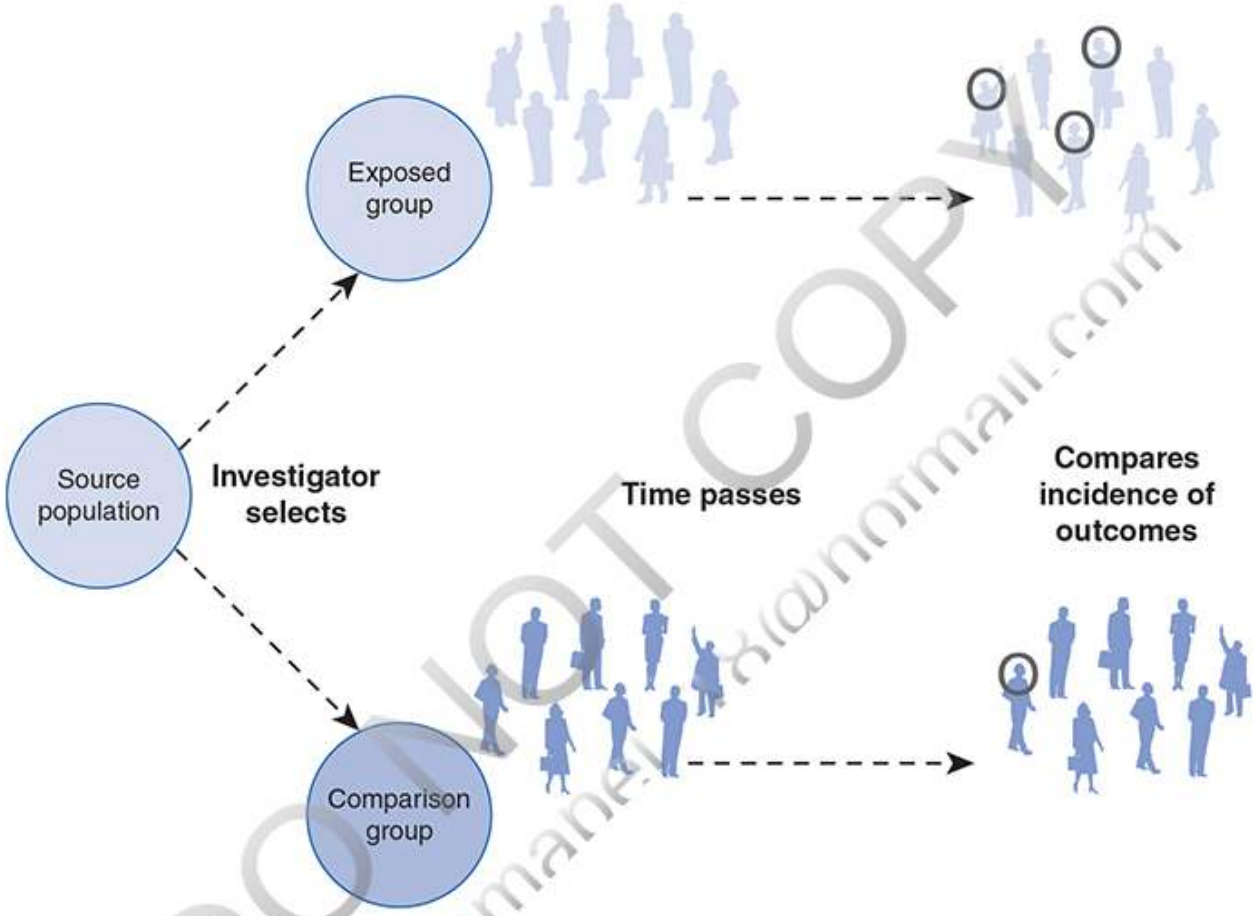

FIGURE 6.2 Schematic representation of cohort study implementation.

Cohort study is the term that is typically used to describe an epidemiological investigation that follows groups with common characteristics. Other expressions that are used include follow-up, incidence, or longitudinal study. There are several additional terms for describing cohort studies that depend on the characteristics of the population from which the cohort is derived, whether the exposure changes over time and whether there are losses to follow-up. The term **fixed cohort** is used when the cohort is formed on the basis of an irrevocable event, such as undergoing a medical procedure. Thus, an individual's exposure in a fixed cohort does not change over time. The term **closed cohort** is used to describe a fixed cohort with no losses to follow-up. In contrast, a cohort study conducted in an **open population**, also known as a **dynamic population**, is defined by exposures that can change over time, such as cigarette smoking. Cohort studies in open populations may also experience losses to follow-up.

Timing of Cohort Studies

Three terms are used to describe the timing of events in a cohort study in relation to the initiation of the study: prospective, retrospective, and ambidirectional. At the initiation of a **prospective cohort study**, participants are grouped on the basis of past or current exposure and are followed into the future to observe the outcomes of interest. When the study commences, the outcomes have not yet developed, and the investigator must wait for them to occur. At the initiation of a **retrospective cohort study**, both the exposures and outcomes have already occurred when the study begins. Thus, this type of investigation studies only prior and not future outcomes. An **ambidirectional cohort study** has both prospective and retrospective components. The decision whether to conduct a retrospective, a prospective, or an ambidirectional study depends on the research question, practical constraints such as time and money, and the availability of suitable study populations and records.

Selection of the Exposed Population

The choice of the exposed group in a cohort study depends on the hypothesis being tested; the exposure frequency; and feasibility considerations, such as the availability of records and ease of follow-up. Special cohorts are used to study the health effects of rare exposures, such as uncommon workplace chemicals, unusual diets, and uncommon lifestyles. Special cohorts are often selected from occupational groups (such as automobile manufacturing workers) or religious groups (such as Mormons) in which the exposures are known to occur. General cohorts are typically assembled for common exposures, such as cigarette smoking and alcohol consumption. These cohorts are often selected from professional groups, such as nurses, or from well-defined geographic areas to facilitate follow-up and accurate ascertainment of the outcomes under study.

Selection of Comparison Group

There are three sources for the comparison group in a cohort study: an internal comparison group, the general population, and a comparison cohort. An internal comparison group consists of unexposed members of the same cohort. An internal comparison group should be used whenever possible because its characteristics will be the most similar to the exposed group. The general population is used for comparison when it is not possible to find a comparable internal comparison group. The general population comparison is based on preexisting population data on disease incidence and mortality. A comparison cohort consists of members of another cohort. It is the least desirable option because the comparison cohort, although not exposed to the exposure under study, is often exposed to other potentially harmful substances and therefore the results can be difficult to interpret.

Sources of Information

Cohort study investigators typically rely on many sources for information on exposures, outcomes, and other key variables. They include medical and employment records, interviews, direct physical examinations, laboratory tests, biological specimens, and environmental monitoring. Some of these sources are preexisting, and others are designed specifically for the study. Because each type of source has advantages and disadvantages, investigators often use several sources to piece together all the necessary information.

Healthcare records are used to describe a participant's exposure history in studies of possible adverse health effects stemming from medical procedures. The advantages of these records include low expense and a high level of accuracy and detail regarding a disease and its treatment. Their main disadvantage is that information on many other key characteristics, apart from basic demographic characteristics, is often missing.

Employment records are used to identify individuals for studies of occupational exposures. Typical employment record data include job title, department of work, years of employment, and basic demographic characteristics. Like medical records, they usually lack details on exposures and other important variables.

Because existing records, such as healthcare and employment records, often have limitations, many studies are based on data collected specifically for the investigation. They include interviews, physical examinations, and laboratory tests. Interviews and self-administered questionnaires are particularly useful for obtaining information on lifestyle characteristics (such as use of cigarettes or alcohol), which are not consistently found in records. Whatever the source of information, it is important to use comparable procedures for obtaining information on the exposed and unexposed groups. Biased results may occur if different sources and procedures are used. Thus, all resources used for one group must be used for the other. In addition, it is a good idea to mask investigators to the exposure status of a subject so they make unbiased decisions when assessing the outcomes. Standard outcome definitions are also recommended to guarantee both accuracy and comparability.

Approaches to Follow-Up

Loss to follow-up occurs either when the participant no longer wishes to take part in the study or he or she cannot be located. Because high rates of follow-up are critical to the success of a cohort study, investigators have developed many methods to maximize retention and trace study participants.[9] For prospective cohort studies, strategies include collection of information (such as full name, Social Security number, and date of birth) that helps locate participants as the study progresses. In addition, regular contact is recommended for participants in prospective studies. These contacts might involve requests for up-to-date outcome information or newsletters describing the study's progress and findings.[9] The best strategy to use when participants do not initially respond is to send additional mailings.

When participants are truly lost to follow-up, investigators employ a number of strategies.[9] They include sending letters to the last known address with "Address Correction requested"; checking telephone directories; directory assistance; Internet resources, such as **whitepages.com**; vital statistics records; driver's license rosters; and voter registration records and contacting relatives, friends, and physicians identified at baseline.

Analysis

The primary objective of analyzing cohort study data is to compare the occurrence of symptoms, disease, and death in the exposed and unexposed groups. If it is not possible to find a completely unexposed group to serve as the comparison, then the least exposed group is used.

The occurrence of the outcome is usually measured using cumulative incidence or incidence rates, and the relationship between the exposure and outcome is quantified using absolute or relative difference between the risks or rates.

Overview of Case–Control Studies

The case–control study has traditionally been viewed as an inferior alternative to the cohort study. In the traditional view, subjects are selected on the basis of whether they have or do not have the disease. An individual who has the disease is termed a **case**, and someone who does not have the disease is termed a **control**. The exposure histories of cases and controls are then obtained and compared. Thus, the central feature of the traditional view is the comparison of the exposure histories of the cases and controls. This differs from the logic of experimental and cohort study designs in which the key comparison is disease incidence between the exposed and unexposed (or least exposed) groups.

Over the past 3 decades, the traditional view that a case–control study is a backward cohort study has been supplanted by a modern view that asserts that it is merely an efficient way to learn about the relationship between an exposure and a disease.[10] More specifically, a case–control study is a method of sampling a population in which researchers identify and enroll cases of disease and a sample of the source population that gave rise to the cases. The sample of the source population is known as the control group (see **FIGURE 6-3**). Its purpose is to provide information on the exposure distribution in the population that produced the cases so that the rates of disease in exposed and unexposed groups can be compared. Thus, the key comparison in the modern view is the same as that of a cohort study.

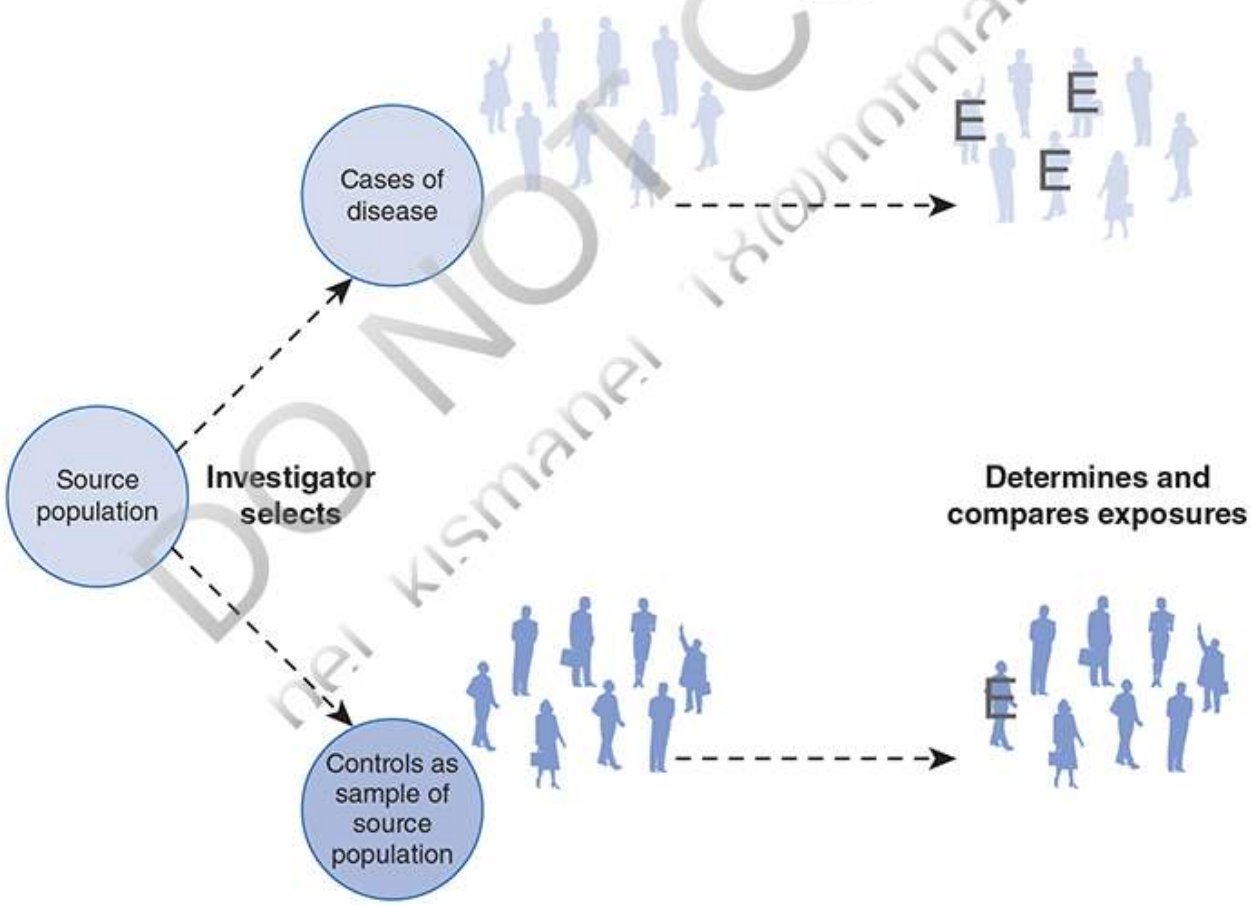

FIGURE 6.3 Schematic representation of case–control study implementation.

Selection of Cases

The first step in the selection of cases for a case–control study is the formulation of a disease or case definition. A case definition is usually based on a combination of signs and symptoms, physical and pathological examinations, and results of diagnostic tests. It is best to use all available evidence to define with as much accuracy as possible the true cases of disease.

Once investigators have created a case definition, they can begin case identification and enrollment. Typical sources for identifying cases are hospital or clinic patient rosters; death certificates; special surveys; and reporting systems, such as cancer or birth defects registries. Investigators consider both accuracy and efficiency in selecting a particular source for case identification. The goal is to identify as many true cases of disease as quickly and cheaply as possible.

Another important issue in selecting cases is whether they should be incident or prevalent. Researchers who study the causes of disease prefer incident cases because they are usually interested in the factors that lead to developing a disease rather than factors that affect its duration. However, sometimes epidemiologists have no choice but to rely on prevalent cases (e.g., when studying the causes of insidious diseases whose exact onset is difficult to pinpoint). Studies using prevalent cases must be interpreted cautiously because it is impossible to determine whether the exposure is related to the inception of the disease, its duration, or a combination of the two.

Selection of Controls

Controls are a sample of the population that produced the cases. The guiding principle for the valid selection of controls is that they come from the same base population as the cases. If this condition is met, then a member of the control group who gets the disease under study would end up as a case in the study. This concept is known as “the would criterion,” and its fulfillment is crucial to the validity of a case–control study. Another important principle is that controls must be sampled independently of exposure status. In other words, exposed and unexposed controls should have the same probability of selection.

Epidemiologists use several sources for identifying controls in case–control studies. They may sample (1) individuals from the general population, (2) individuals attending a hospital or clinic, (3) friends or relatives identified by the cases, or (4) individuals who have died. Population controls are typically selected when cases are identified from a well-defined population, such as residents of a geographic area. These controls are usually identified using voter registration lists, driver’s license rosters, telephone directories, and random digit dialing (a method for identifying telephone subscribers living in a defined geographic area).

Population controls have one principal advantage that makes them preferable to other types of controls. Because of the manner in which population controls are identified, investigators are usually assured that the controls come from the same population as the cases. Thus, investigators are usually confident that population controls are comparable to the cases with respect to demographic and other important variables. However, population controls have several disadvantages. First, they are time consuming and expensive to identify. Second, these individuals do not have the same level of interest in participating as do cases and controls identified from other sources. Third, because they are generally healthy, their recall may be less accurate than that of cases, who are likely reviewing their history in search of a “reason” for their illness.

Epidemiologists usually select hospital and clinic controls when they identify cases from these healthcare facilities. Thus, these controls have diseases or have experienced events (such as a car accident) for which they have sought medical care. The most difficult aspect of using these types of controls is determining which diseases or events are suitable for inclusion. In this regard, investigators should follow two general principles. First, the illnesses in the control group should, on the basis of current knowledge, be unrelated to the exposure under study. For example, a case–control study of cigarette smoking and emphysema should not use lung cancer patients as controls because lung cancer is known to be caused by smoking cigarettes. Second, the control’s illness should have the same referral pattern to the healthcare facility as the case’s illness. For example, a case–control study of acute appendicitis should use patients with other acute conditions as controls. Following this principle will help ensure that the cases and controls come from the same source population.

There are several advantages to the use of hospital and clinic controls. Because they are easy to identify and have good participation rates, hospital and clinic controls are less expensive to identify than population controls. In addition, because they come from the same source population, they will have characteristics comparable to the cases. Finally, their recall of prior exposures will be similar to the cases’ recall because they are also ill. The main disadvantage of this type of control is the difficulty in determining appropriate illnesses for inclusion.

In rare circumstances, deceased and “special” controls are enrolled. Deceased controls are occasionally used when some or all of the cases are deceased by the time data collection begins. Researchers usually

identify these controls by reviewing death records of individuals who lived in the same geographic area and died during the same time period as the cases. The main rationale for selecting dead controls is to ensure comparable data collection procedures between the two groups. For example, if researchers collect data via interview, they would conduct proxy interviews with subjects' spouses, children, relatives, or friends for both the dead cases and dead controls.

However, many epidemiologists discourage the use of dead controls because they may not be a representative sample of the source population that produced the cases, which by definition, consists of living people. Furthermore, the investigator must consider the study hypothesis before deciding to use dead controls because they are more likely than living controls to have used tobacco, alcohol, or drugs.[11] Consequently, dead controls may not be appropriate if the study hypothesis involves one of these exposures.

In unusual circumstances, a friend, spouse, or relative (usually a sibling) is nominated by a case to serve as his or her control. These "special" controls are used because, if they are related to the cases, they are likely to share the cases' socioeconomic status, race, age, educational level, and genetic characteristics. However, cases may be unwilling or unable to nominate people to serve as their controls. In addition, biased results are possible if the study hypothesis involves a shared activity among the cases and controls.

Methods for Sampling Controls

Epidemiologists use three main strategies for sampling controls in a case–control study. Investigators can select controls from the "noncases" or "survivors" at the end of the case diagnosis and accrual period. This method of selection, which is known as survivor sampling, is the predominant method for selecting controls in traditional case–control studies. In case–base or case–cohort sampling, investigators select controls from the population at risk at the beginning of the case diagnosis and accrual period. In risk set sampling, controls are selected from the population at risk as the cases are diagnosed.

When case–base and risk set sampling methods are used, the control group may include future cases of disease. Although this may seem incorrect, modern epidemiological theory supports it. Recall that both diseased and nondiseased individuals contribute to the denominators of the risks and rates in cohort studies. Thus, it is reasonable for the control group to include future cases of disease because it is merely an efficient way to obtain the denominator data for the risks and rates.

Sources of Exposure Information

Case–control studies are used to investigate the risk of disease in relation to a wide variety of exposures, including those related to lifestyle, occupation, environment, genes, diet, reproduction, and the use of medications.[12] Most exposures that are studied are complex; therefore, investigators must attempt to obtain sufficiently detailed information on the nature, sources, frequency, and duration of these exposures. Sources available for obtaining exposure data include in-person and telephone

interviews; self-administered questionnaires; preexisting medical, pharmacy, registry, employment, insurance, birth, death, and environmental records; and biological specimens.[12] When selecting a particular source, investigators consider its availability and accuracy and the logistics and cost of data collection. Accuracy is a particular concern in case–control studies because exposure data are retrospective. In fact, the relevant exposures may have occurred many years before data collection, making it difficult to gather correct information.

Analysis

As described earlier, controls are a sample of the population that produced the cases. However, in most instances, the sampling fraction is not known; therefore, the investigator cannot fill in the total population in the margin of a two-by-two table or obtain the rates and risks of disease. Instead, the researcher obtains a number called an odds, which functions as a rate or risk. An **odds** is defined as the probability that an event will occur divided by the probability that it will not occur. In a case–control study, epidemiologists typically calculate the odds of being a case among the exposed (a/b) compared to the odds of being a case among the nonexposed (c/d). The ratio of these two odds is expressed as follows:

$$\frac{a/b}{c/d} \text{ or } \frac{ad}{bc}$$

This ratio, known as the disease **odds ratio**, provides an estimate of the relative risk just as the incidence rate ratio and cumulative incidence ratio do. Risk or rate differences are not usually obtainable in a case–control study. However, it is possible to calculate the attributable proportion among the exposed and the attributable proportion in the total population using the odds ratio and the proportion of exposed controls.

Case–Crossover Study

The case–crossover study is a variant of the case–control study that was developed for settings in which the risk of the outcome is increased for only a brief time following the exposure.[13] The period of increased risk following the exposure is termed the **hazard period**.[14] In the case–crossover study, cases serve as their own controls, and the exposure frequency during the hazard period is compared with that from a control period. Because cases serve as their own controls, this design has several advantages, including the elimination of confounding by characteristics such as gender and race and the elimination of a type of bias that results from selecting unrepresentative controls. In addition, because variability is reduced, this design requires fewer subjects than does the traditional case–control study.

When Is It Desirable to Use a Particular Study Design?

The goal of every epidemiological study is to gather correct and sharply defined data on the relationship between an exposure and a health-related state or an event in a population. The three main study designs represent different ways of gathering this information. Given the strengths and weaknesses of each design, there are circumstances for which a particular type of study is clearly indicated. These situations are described in the following paragraphs.

Experimental Studies

Investigators conduct an experimental study when they wish to learn about a prevention or treatment for a disease. In addition, they conduct this type of study when they need data with a high degree of validity that is simply not possible in an observational study. The high degree of validity in an experimental study stems mainly from investigators' ability to randomize subjects to either the treatment group or the comparison group and thereby control for distortions produced by confounding variables. A high level of validity may be needed for studying a prevention or treatment that is expected to have a small effect, usually defined as a difference of 20% or less between groups. A difference of this size is difficult to detect using an observational study because of uncontrolled bias and confounding. When the difference between groups is small, even a small degree of bias or confounding can create or mask an effect.

Although most scientists agree that well-conducted experimental studies produce more scientifically rigorous data than do observational studies, several thorny issues make it difficult to conduct experimental studies. These issues include noncompliance, the need to maintain high follow-up rates, high costs, physician and patient reluctance to participate, and numerous ethical issues. Investigators must address all these issues when considering this design. In particular, it is ethical to conduct experimental studies only when there is a state of equipoise within the expert medical community regarding the treatment. **Equipoise** is a "state of mind characterized by legitimate uncertainty or indecision as to choice or course of action."[6(p95)] In other words, there must be genuine confidence that a treatment may be worthwhile to administer it to some individuals and genuine reservations about the treatment to withhold it from others.

Observational Studies

Observational studies can be used to study the effects of a wider range of exposures than experimental studies, including preventions, treatments, and possible causes of disease. For example, observational studies provide information to explain the causes of disease incidence and the determinants of disease progression to predict the future healthcare needs of a population and to control disease by studying ways to prevent disease and prolong life with disease. The main limitation of observational studies is investigators' inability to have complete control over disturbing influences or extraneous factors. As Susser states, "Observational studies have a place in the epidemiological armament no less necessary and valid than controlled trials; they take second place in the hierarchy of rigor but not in practicability and generalizability. . . . Even when trials are possible, observational studies may yield more of the truth than randomized trials."[15](p95)

Once an investigator has decided to conduct an observational study, the next decision is usually whether to select a cohort or case–control design. Because a cohort study can provide information on a large number of possible health effects, this type of study is preferable when little is known about the health consequences of an exposure. A cohort study is also efficient for investigating a rare exposure, which is usually defined as a frequency of less than 20%.

Case–control studies are preferable when little is known about the etiology of a disease because they can provide information on a large number of possible risk factors. Case–control studies take less time and cost less money than do cohort studies primarily because the control group is a sample of the source population. Case–control studies are also more efficient than cohort studies for studying rare diseases because fewer subjects are needed and for studying diseases with long induction and latent periods because long-term prospective follow-up is avoided. (A long induction and latent period means that there is a long time between the causal action of an exposure and the eventual diagnosis of disease.[16]) Because of their relatively smaller sample size, case–control studies are preferred when the exposure data are difficult or expensive to obtain. Finally, they are desirable when the population under study is dynamic because it is difficult to keep track of a population that is constantly changing. Tracing is required for a typical cohort study but not for a typical case–control study.

Case–control studies have a few important disadvantages. First, because of the retrospective nature of the data collection, there is a greater chance of bias. Some epidemiologists have argued that case–control studies are not well suited for detecting weak associations (those with odds ratios less than 1.5) because of the likelihood of bias.[17] Second, because data collection is retrospective, it may be difficult to establish the correct temporal relationship between the exposure and disease.

If an investigator has decided to conduct a cohort study, he or she must make one more choice: Should it be a retrospective or prospective cohort study? This decision depends on the particular research question, the practical constraints of time and money, and the availability of suitable study populations and records. For example, a retrospective design must be used to study historical exposures. In making this decision, the investigator must also take into account the complementary advantages and disadvantages of retrospective and prospective cohort studies. For example, retrospective cohort studies are more efficient than prospective studies for studying diseases with long induction and latent periods. However, minimal information is usually available on the exposure, outcome, confounders, and contacts for follow-up because

retrospective cohort studies typically rely on existing records that were not designed for research purposes. In addition, the use of retrospective data makes it more difficult to establish the correct temporal relationship between the exposure and disease.

In prospective cohort studies, investigators can usually obtain more detailed information on exposures and confounders because they have more control of the data collection process and can gather information directly from the participants. Follow-up may be easier because the investigator can obtain tracing information from participants and maintain periodic contact with subjects. Prospective cohort studies are considered less vulnerable to bias than retrospective studies because the outcomes have not occurred when the cohort is assembled and the exposures are assessed. In addition, it is easier for investigators to establish a clear temporal relationship between exposure and outcome. A decision tree depicting the choices between the three main study designs is shown in **FIGURE 6-4**.

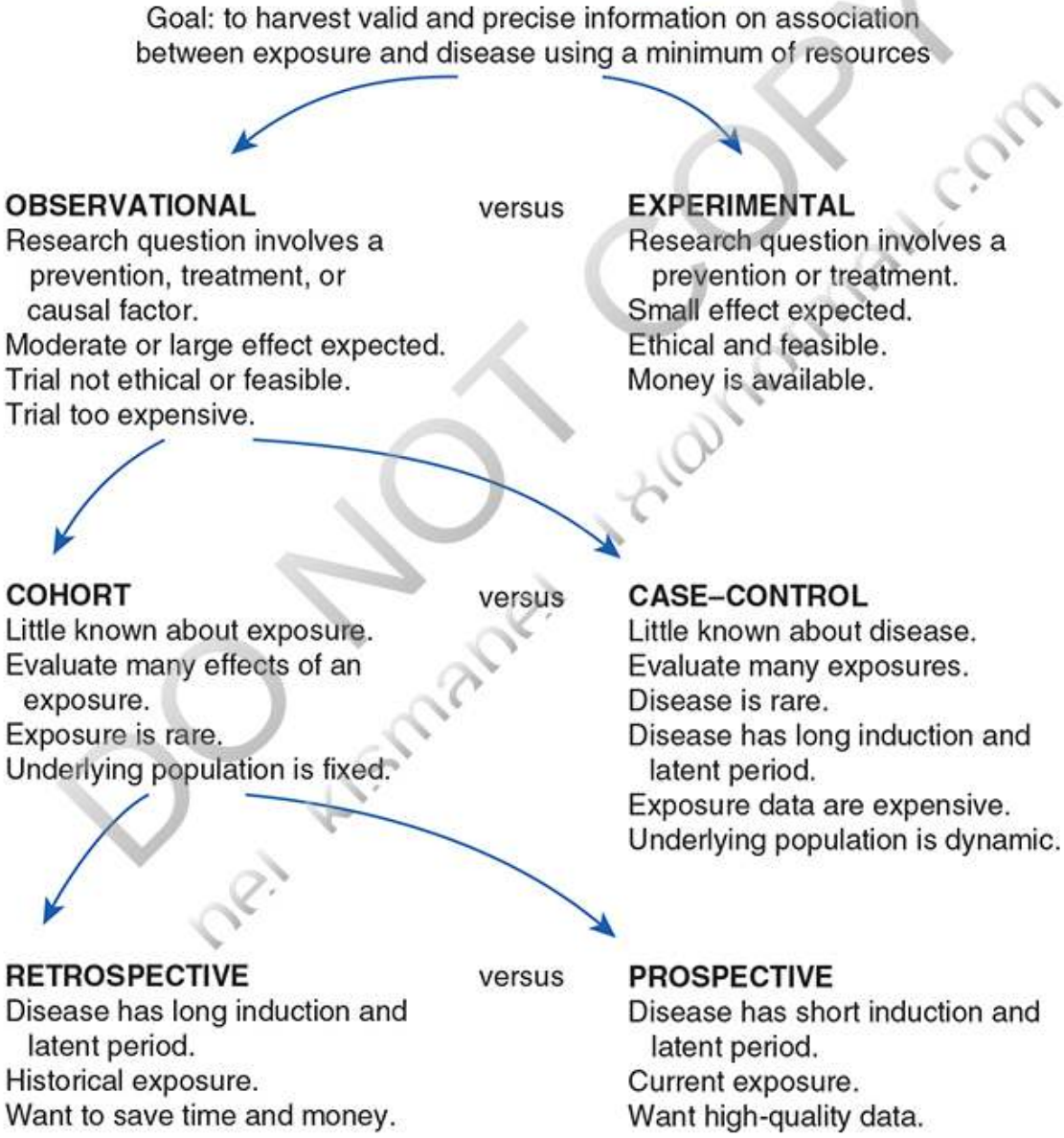

FIGURE 6.4 Decision tree for choosing among study designs.

also take into account the complementary advantages and disadvantages of retrospective and prospective cohort studies. For example, retrospective cohort studies are more efficient than prospective studies for studying diseases with long induction and latent periods. However, minimal information is usually available on the exposure, outcome, confounders, and contacts for follow-up because retrospective cohort studies typically rely on existing records that were not designed for research purposes. In addition, the use of retrospective data makes it more difficult to establish the correct temporal relationship between the exposure and disease.

In prospective cohort studies, investigators can usually obtain more detailed information on exposures and confounders because they have more control of the data collection process and can gather information directly from the participants. Follow-up may be easier because the investigator can obtain tracing information from participants and maintain periodic contact with subjects. Prospective cohort studies are considered less vulnerable to bias than retrospective studies because the outcomes have not occurred when the cohort is assembled and the exposures are assessed. In addition, it is easier for investigators to establish a clear temporal relationship between exposure and outcome. A decision tree depicting the choices between the three main study designs is shown in **FIGURE 6-4**.

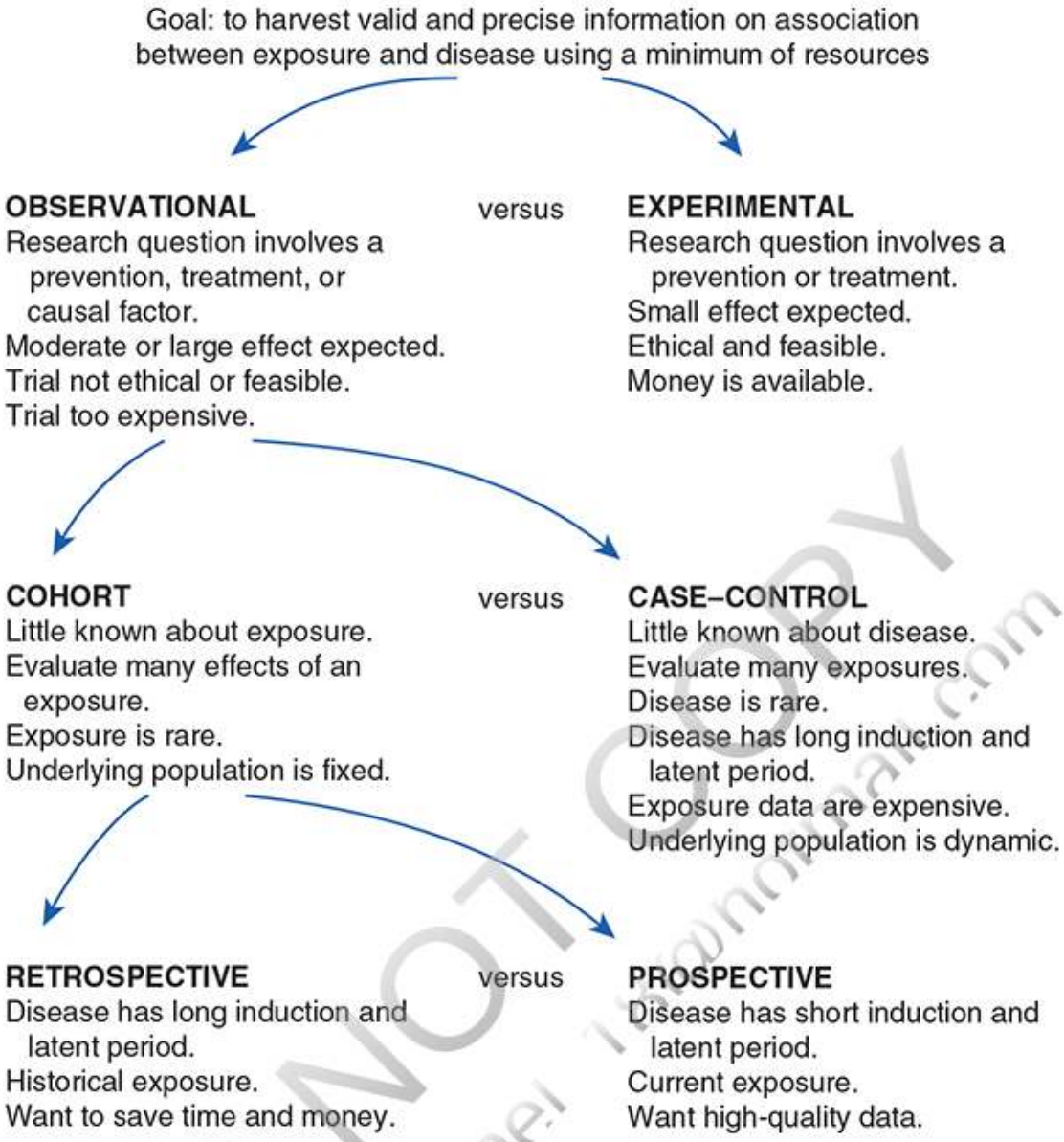

FIGURE 6.4 Decision tree for choosing among study designs.

Other Types of Studies

In addition to the three main study designs described in the previous sections, two other types of studies are commonly conducted in epidemiological research: cross-sectional and ecological studies (see **TABLE 6-2**). Although both studies are popular, these designs have important limitations that are not present in the other observational designs. Lastly, agent-based modeling is a new form of research in epidemiology that is gaining popularity.

TABLE 6-2 Key Features of Cross-Sectional and Ecological Studies

Cross-sectional studies

- Examine association at a single point in time, and therefore measure exposure prevalence in relation to disease prevalence.
- Cannot infer temporal sequence between exposure and disease if exposure is a changeable characteristic.
- Other limitations may include preponderance of prevalent cases of long duration and healthy worker effect.
- Advantages include generalizability and low cost.

Ecological studies

- Examine rates of disease in relation to a population-level factor.
- Population-level factors include summaries of individual population members, environmental measures, and global measures.
- Study groups are usually identified by place, time, or a combination of the two.
- Limitations include the ecological fallacy and lack of information on confounding variables.
- Advantages include low cost, wide range of exposure levels, and the ability to examine contextual effects on health.

Cross-Sectional Studies

A **cross-sectional study** "examines the relationship between diseases (or other health-related characteristics) and other variables of interest as they exist in a defined population at one particular time."[2(p95)] Unlike populations studied in cohort and case–control studies, cross-sectional study populations are commonly selected without regard to exposure or disease status. Cross-sectional studies typically take a snapshot of a population at a single point in time and therefore usually measure the disease prevalence in relation to the exposure prevalence. In other words, current disease status is usually examined in relation to current exposure level. However, it is possible for cross-sectional studies to examine disease prevalence in relation to past exposures if the dates of the exposures are ascertained.

Cross-sectional studies are carried out for public health planning and etiologic research. Most governmental surveys conducted by the National Center for Health Statistics are cross-sectional in nature. For example, the National Survey of Family Growth is a periodic population-based survey focusing on factors that affect family health and fertility. Its most recent cycle was based on a national probability sample of men and women aged 15 to 49 years. In-person interviews gathered information on marriage and divorce, pregnancy, infertility, and contraception.[18]

Cross-sectional studies are fairly common in occupational settings using data from preemployment physical examinations and company health insurance plans.[19(p95)] For example, investigators conducted a cross-sectional study to determine the relationship between low back pain and sedentary work among crane operators, straddle-carrier drivers, and office workers.[20] All three groups had sedentary jobs that required prolonged sitting. Company records were used to identify approximately 300 currently employed male workers aged 25 through 60 years who had been employed for at least 1 year in their current job. Investigators assessed the "postural load" by observing workers' postures (such as straight upright position and forward or lateral flexion) and movements (such as sitting, standing, and walking). The investigators found that the prevalence of current and recent low back pain was more common among crane operators and straddle-carrier drivers than office workers. The crane operators and straddle-carrier drivers had two to three times the risk of low back pain than did the office workers. The authors postulated that these differences resulted from crane operators' and straddle-carrier drivers' more frequent adoption of "non-neutral" trunk positions involving back flexion and rotation while on the job.[20]

Unfortunately, when epidemiologists measure the exposure prevalence in relation to disease prevalence in cross-sectional studies, they are not able to infer the temporal sequence between the exposure and disease. In other words, they cannot tell which came first—the exposure or the disease. This occurs when the exposure under study is a *changeable characteristic*, such as a place of residence or a habit such as cigarette smoking. Consider, for example, a hypothetical cross-sectional study of stress levels and the risk of ovarian infertility conducted among patients seeking treatment at an infertility clinic. The current stress levels of women who have a diagnosis of ovarian infertility is compared with that of fertile women whose husbands are the source of the infertility. If the stress level is three times greater among the infertile women, one could conclude that there is a moderately strong association between stress and ovarian infertility. However, it is difficult to know whether stress caused the infertility because the women may have become stressed after they began having difficulties achieving a pregnancy. This is quite possible given that precise onset of infertility is difficult to determine and that medical treatment for infertility usually does not begin until a couple

Agent-based modeling is an increasingly popular form of research in epidemiology. It is not a study design but rather a *method of analysis* that uses computer simulations to study the complex interactions among individuals, their physical and social environments, and time.[29] For example, an agent-based model was recently developed to determine the best strategy for reducing the prevalence of violence-related posttraumatic stress disorder (PTSD). In particular, the model contrasted the effect of hot-spot policing (a population-level intervention expected to prevent neighborhood violence) and cognitive behavioral therapy (an individual-level intervention expected to shorten disease duration). The study found that the combination of both interventions produced the greatest reduction in PTSD prevalence.

Summary

Epidemiologists use both experimental and observational study designs to answer research questions. Each type of design represents a different way of harvesting the necessary information. The selection of one design over another depends on the research question and takes into account validity, efficiency, and ethical concerns.

For ethical reasons, experimental studies can be used to investigate only preventions and treatments for diseases. The hallmark of an experimental study is the investigator's active manipulation of the agent under study. Here, the investigator assigns subjects (usually at random) to two or more groups that either receive or do not receive the preventive or therapeutic agent. Investigators select this study design when they need data with a high degree of validity that is simply not possible to obtain in an observational study. However, experimental studies are expensive and often infeasible and unethical, and so most epidemiological research consists of observational studies.

Observational studies can be used to investigate a broader range of exposures, including causes, preventions, and treatments for diseases. The two most important types of observational studies are the cohort study and the case–control study. Epidemiologists use a cohort study when little is known about an exposure because this type of study allows investigators to examine many health effects in relation to an exposure. In a cohort study, subjects are defined according to their exposure levels and followed for disease occurrence. In contrast, investigators use a case–control study when little is known about a disease because this type of study allows researchers to examine many exposures in relation to a disease. In a case–control study, cases with the disease and controls are defined and their exposure histories are collected and compared.

Cross-sectional and ecological studies and agent-based modeling are three other types of observational research. Cross-sectional studies examine exposure prevalence in relation to disease prevalence in a defined population at a single point in time. Ecological studies examine disease rates in relation to a population-level factor. Both types of designs have important limitations absent from the other observational studies. An unclear temporal relationship between exposure and disease arises in cross-sectional studies of changeable exposures. Problems making cross-level inferences from the group to the individual (known as the ecological fallacy) occur in ecological studies. Agent-based modeling is not a study design but rather a method of analysis that uses computer simulations to study complex interactions between individuals, their environment, and time.

Cross-sectional and ecological studies and agent-based modeling are three other types of observational research. Cross-sectional studies examine exposure prevalence in relation to disease prevalence in a defined population at a single point in time. Ecological studies examine disease rates in relation to a population-level factor. Both types of designs have important limitations absent from the other observational studies. An unclear temporal relationship between exposure and disease arises in cross-sectional studies of changeable exposures. Problems making cross-level inferences from the group to the individual (known as the ecological fallacy) occur in ecological studies. Agent-based modeling is not a study design but rather a method of analysis that uses computer simulations to study complex interactions between individuals, their environment, and time.

References

1. MacMahon B, Trichopoulos D. *Epidemiology Principles and Methods*. 2nd ed. Boston, MA: Little, Brown and Company; 1996.
2. Porta, M. *A Dictionary of Epidemiology*. 6th ed. New York, NY: Oxford University Press; 2014.
3. Heymann, DL. *Control of Communicable Diseases Manual*. 20th ed. Washington, DC: American Public Health Association; 2015.
4. Pickett, JP (exec. ed.). *The American Heritage Dictionary of the English Language*. 4th ed. Boston, MA: Houghton Mifflin; 2000.
5. Colton, T. *Statistics in Medicine*. Boston, MA: Little, Brown and Company; 1974.
6. Meinert, CL. *Clinical Trials Dictionary: Terminology and Usage Recommendations*. Baltimore, MD: The Johns Hopkins Center for Clinical Trials; 1996.
7. Friedman LM, Furberg CD, Demets DL. *Fundamentals of Clinical Trials*. 2nd ed. Littleton, MA: PSG Publishing Co.; 1985.
8. Newell, DJ. Intention-to-treat analysis: implications for quantitative and qualitative research. *Int J Epidemiol*. 1992;21:837-841.
9. Hunt JR, White E. Retaining and tracking cohort study members. *Epidemiol Rev*. 1998;20:57-70.
10. Miettinen OS. The "case-control" study: valid selection of study subjects. *J Chron Dis*. 1985;38:543-548.
11. McLaughlin JK, Blot WJ, Mehl ES, Mandel JS. Problems in the use of dead controls in case–control studies. I. general results. *Am J Epidemiol*. 1985;121:131-139.
12. Correa A, Stewart WF, Yeh HC, Santos-Burgoa C. Exposure measurement in case-control studies: reported methods and recommendations. *Epidemiol Rev*. 1994;16:18-32.
13. Maclure, M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol*. 1991;133:144-153.
14. Mittleman MA, Maclure M, Robins JM. Control sampling strategies for case-crossover studies: an assessment of relative efficiency. *Am J Epidemiol*. 1995;142:91-98.
15. Susser M. Editorial: the tribulations of trials—interventions in communities. *Am J Public Health*. 1995;85:156-158.
16. Rothman KJ. Induction and latent periods. *Am J Epidemiol*. 1981;114:253-259.
17. Austin H, Hill HA, Flanders WD, Greenberg RS. Limitations in the application of case-control methodology. *Epidemiol Rev*. 1994;16:65-76.
18. Key statistics from the National Survey of Family Growth. **cdc.gov**. **https://www.cdc.gov/nchs/nsfg/key_statistics.htm**. Accessed October 2016.
19. Monson RR. *Occupational Epidemiology*. 2nd ed. Boca Raton, FL: CRC Press; 1990.
20. Burdorf A, Naaktgeboren B, De Groot H. Occupational risk factors for low back pain among sedentary workers. *J Occup Med*. 1993;35:1213-1220.
21. Hu H, Shih R, Rothenberg S, Swartz BS. The epidemiology of lead toxicity in adults: measuring dose and consideration of other methodologic issues. *Environ Health Perspect*. 2007;115:455-462.
22. Morgenstern H. Ecologic studies. In: Rothman KJ, Greenland S, Lash TL (eds.). *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams and Wilkins; 2008.

23. Zhang J, Zhao Z, Berkel HJ. Egg consumption and mortality from colon and rectal cancers: an ecological study. *Nutr Cancer*. 2003;46:158-165.

24. Hurley SF, Jolley DJ, Kaldor JM. Effectiveness of needle-exchange programmes for prevention of HIV infection. *Lancet*. 1997;349:1797-1800.

25. Shridhar K, Rajaraman P, Koyande S, et al. Trends in mouth cancer incidence in Mumbai, India (1995-2009): an age-period-cohort analysis. *Cancer Epidemiol.* 2016; 42:66-71.

26. Magnus P, Jaakkola JJK, Skrondal A, et al. Water chlorination and birth defects. *Epidemiology*. 1999;10:513-517.

27. Susser M. The logic in ecological: I. the logic of analysis. *Am J Public Health*. 1994;84:825-829.

28. Koopman JS, Longini IM. The ecological effects of individual exposures and nonlinear disease dynamics in populations. *Am J Public Health*. 1994;84:836-842.

29. Cerda M, Tracy M, Keyes KM, Galea S. To treat or to prevent? Reducing the population burden of violence-related post-traumatic stress disorder. *Epidemiology.* 2015;26:681-689.